Contribution ID: d7cd46e2-036c-4431-9360-02e030e3143d
Date: 11/12/2024 11:34:22

MULTI-STAKEHOLDER CONSULTATION FOR COMMISSION GUIDELINES ON THE APPLICATION OF THE DEFINITION OF AN AI SYSTEM AND THE PROHIBITED AI PRACTICES ESTABLISHED IN THE AI ACT

Fields marked with * are mandatory.

MULTI-STAKEHOLDER CONSULTATION FOR COMMISSION GUIDELINES ON THE APPLICATION OF THE DEFINITION OF AN AI SYSTEM AND THE PROHIBITED AI PRACTICES ESTABLISHED IN THE AI ACT

Disclaimer: This document is a working document for consultation and does not prejudge the final decision that the Commission may take on the final guidelines. The responses to this consultation paper will provide important input to the Commission when preparing the guidelines.

The European AI Office is launching this multi-stakeholder consultation on the application of the definition of an AI system and the prohibited AI practices established in the AI Act. This consultation is **targeted to stakeholders of different categories**, including providers and deployers of AI systems such as businesses, authorities (including local public authorities) and other organisations, academia and research institutions, trade unions and other workers' representatives, civil society organisations, public supervisory authorities, and the general public.

As not all questions may be relevant for all stakeholders, respondents may reply only to the section(s) and the questions they consider relevant. Respondents are

encouraged to provide **explanations and concrete cases** as part of their responses to support the practical usefulness of the guidelines.

The targeted consultation is available in English only and will be **open for 4 weeks starting on 13 November until 11 December 2024 (till 23:59). We strongly encourage early submissions.**

The questionnaire for this consultation is structured along 2 sections with several questions.

1. Questions in relation to the definition of an AI system

2. Questions in relation to prohibited AI practices

We **welcome collective answers from organisations.** You have the option to indicate if you a submitting such a collective answer in the end of the first section and identify the organisations on whose behalf the submission is made.

We welcome full or partial replies from all respondents based on their expertise and perspective.

All contributions to this consultation may be made publicly available. Therefore, please do not share any confidential information in your contribution. Individuals can request to have personal information removed from their contribution.

The Commission may publish a summary of the results of the consultation. In that case, results will be based on aggregated data and respondents will not be directly quoted.

Please allow enough time to submit your application before the deadline to avoid any issues. In case you experience technical problems which prevent you from submitting your application within the deadline, please take screenshots of the issue and the time it occurred.

In case you face any technical difficulties or would like to ask a question, please contact: CNECT-AIOFFICE@ec.europa.eu

---

General Introduction

The Artificial Intelligence Act (Regulation (EU) 2024/1689, hereinafter 'the AI Act'), which entered into force on 1 August 2024, improves the internal market by laying down harmonised rules for trustworthy and human-centric Artificial Intelligence (AI) in the EU (Article 1 AI Act). It aims to promote innovation and uptake of AI, while ensuring a high level of protection of health, safety and fundamental rights, including democracy and the rule of law.

The AI Act establishes a common definition of an AI system, aligned with the OECD definition (OECD Recommendation on Artificial Intelligence (OECD /LEGAL/0449, 2019, amended 2023)), as a central element of the scope of the AI Act (Article 3(1) AI Act and Recital 12). The AI Act follows a risk-based approach to regulating AI systems, by classifying such systems into different risk categories. One of which are the prohibited AI practices covering AI systems posing unacceptable risks to fundamental rights and European values (Article 5 AI Act).

Pursuant to Article 96(1) AI Act, the Commission must develop guidelines on the practical implementation of the Regulation, *inter alia*, on the prohibited AI practices referred to in Article 5 AI Act and the application of the definition of an AI system as set out in Article 3(1).

The purpose of the present targeted stakeholder consultation is to collect input from a wide range of stakeholders on concrete examples of AI systems and issues with the practical application of the relevant AI Act provisions that could be clarified in the Commission's **guidelines** on the **definition of an 'AI system'** as well as guidelines on the **prohibited AI practices**. The definitions and prohibitions are applicable six months after the entry into force of the AI Act, as from 2 February 2025. The input from this consultation will feed into the Commission guidelines to be adopted in early 2025. It should be noted that the

legal concepts in relation to the AI system definition and the prohibitions are already set out in the AI Act. The Commission launches the present consultation to seek additional practical examples from stakeholders to feed into the guidelines and provide further clarity on practical aspects and use cases.

The objective of the guidelines is to provide consistent interpretation and practical guidance to assist competent authorities in their enforcement actions as well as providers and deployers subject to the AI Act in their compliance actions with a view to ensuring consistent, effective and uniform application of the prohibitions and understanding of what constitutes an AI system within the scope of the AI Act.

About you

* 1. Do you represent one or more organisations (e.g., industry organisation or civil society organisation) or act in your personal capacity (e.g., independent expert)?

- (●) Organisation(s)
- ( ) In a personal capacity

If you are organisation(s), please specify the name(s):

European Center for Not-for-Profit Law Stichting

If you would like to share any affiliation, please specify:

* First name

Karolina

* Surname

Iwanska

* E-Mail address (this won't be published)

karolina@ecnl.org

* Are you headquartered/residing in the EU?

- (●) Yes
- ( ) No
- ( ) Other (e.g. multiple organisations)

* Headquarter / Country of residence

- ( ) AF - Afghanistan
- ( ) AL - Albania
- ( ) DZ - Algeria
- ( ) AD - Andorra
- ( ) AO - Angola
- ( ) AG - Antigua and Barbuda
- ( ) AR - Argentina
- ( ) AM - Armenia
- ( ) AU - Australia
- ( ) AT - Austria
- ( ) AZ - Azerbaijan
- ( ) BS - Bahamas
- ( ) BH - Bahrain
- ( ) BD - Bangladesh
- ( ) BB - Barbados
- ( ) BY - Belarus
- ( ) BE - Belgium
- ( ) BZ - Belize
- ( ) BJ - Benin
- ( ) BT - Bhutan
- ( ) BO - Bolivia
- ( ) BA - Bosnia and Herzegovina
- ( ) BW - Botswana
- ( ) BR - Brazil
- ( ) BN - Brunei Darussalam
- ( ) BG - Bulgaria
- ( ) BF - Burkina Faso
- ( ) BI - Burundi
- ( ) CV - Cabo Verde

- KH - Cambodia
- CM - Cameroon
- CA - Canada
- CF - Central African Republic
- TD - Chad
- CL - Chile
- CN - China
- CO - Colombia
- KM - Comoros
- CG - Congo
- CR - Costa Rica
- CI - Côte D'Ivoire
- HR - Croatia
- CU - Cuba
- CY - Cyprus
- CZ - Czechia
- CD - Democratic Republic of the Congo
- DK - Denmark
- DJ - Djibouti
- DM - Dominica
- DO - Dominican Republic
- EC - Ecuador
- EG - Egypt
- SV - El Salvador
- GQ - Equatorial Guinea
- ER - Eritrea
- EE - Estonia
- SZ - Eswatini
- ET - Ethiopia
- FJ - Fiji
- FI - Finland
- FR - France
- GA - Gabon
- GM - Gambia

- GE - Georgia
- DE - Germany
- GH - Ghana
- GR - Greece
- GD - Grenada
- GT - Guatemala
- GN - Guinea
- GW - Guinea Bissau
- GY - Guyana
- HT - Haiti
- HN - Honduras
- HU - Hungary
- IS - Iceland
- IN - India
- ID - Indonesia
- IR - Iran
- IQ - Iraq
- IE - Ireland
- IL - Israel
- IT - Italy
- JM - Jamaica
- JP - Japan
- JO - Jordan
- KZ - Kazakhstan
- KE - Kenya
- KI - Kiribati
- KW - Kuwait
- KG - Kyrgyzstan
- LA - Laos
- LV - Latvia
- LB - Lebanon
- LS - Lesotho
- LR - Liberia
- LY - Libya

- LI - Liechtenstein
- LT - Lithuania
- LU - Luxembourg
- MG - Madagascar
- MW - Malawi
- MY - Malaysia
- MV - Maldives
- ML - Mali
- MT - Malta
- MH - Marshall Islands
- MR - Mauritania
- MU - Mauritius
- MX - Mexico
- FM - Micronesia
- MC - Monaco
- MN - Mongolia
- ME - Montenegro
- MA - Morocco
- MZ - Mozambique
- MM - Myanmar
- NA - Namibia
- NR - Nauru
- NP - Nepal
- (●) NL - Netherlands
- NZ - New Zealand
- NI - Nicaragua
- NE - Niger
- NG - Nigeria
- KP - North Korea
- MK - North Macedonia
- NO - Norway
- OM - Oman
- PK - Pakistan
- PW - Palau

- PA - Panama
- PG - Papua New Guinea
- PY - Paraguay
- PE - Peru
- PH - Philippines
- PL - Poland
- PT - Portugal
- QA - Qatar
- MD - Republic of Moldova
- RO - Romania
- RU - Russian Federation
- RW - Rwanda
- KN - Saint Kitts and Nevis
- LC - Saint Lucia
- VC - Saint Vincent and the Grenadines
- WS - Samoa
- SM - San Marino
- ST - Sao Tome and Principe
- SA - Saudi Arabia
- SN - Senegal
- RS - Serbia
- SC - Seychelles
- SL - Sierra Leone
- SG - Singapore
- SK - Slovakia
- SI - Slovenia
- SB - Solomon Islands
- SO - Somalia
- ZA - South Africa
- KR - South Korea
- SS - South Sudan
- ES - Spain
- LK - Sri Lanka
- SD - Sudan

- SR - Suriname
- SE - Sweden
- CH - Switzerland
- SY - Syrian Arab Republic
- TJ - Tajikistan
- TZ - Tanzania
- TH - Thailand
- TL - Timor-Leste
- TG - Togo
- TO - Tonga
- TT - Trinidad and Tobago
- TN - Tunisia
- TR - Turkey
- TM - Turkmenistan
- TV - Tuvalu
- UG - Uganda
- UA - Ukraine
- AE - United Arab Emirates
- GB - United Kingdom
- US - United States of America
- UY - Uruguay
- UZ - Uzbekistan
- VU - Vanuatu
- VE - Venezuela
- VN - Viet Nam
- YE - Yemen
- ZM - Zambia
- ZW - Zimbabwe

* Do you have an office or other kind of representation in the EU?

- (●) Yes, we have a subsidiary, branch office or similar in the EU
- ( ) Yes, other
- ( ) No
- ( ) Not applicable

If applicable, please specify

* If you are an organisation, what is the size of your organisation and does it qualify as a small or medium sized enterprise according to the EU recommendation 2003 /361, if applicable ?

- ○ Small
- ○ Medium
- ○ Large
- ○ Other (e.g. multiple organisations, local authorities)
- ◉ Not applicable

If other, please specify

* Which stakeholder category would you consider yourself in?

- ○ Provider of an AI system
- ○ Deployer of an AI system
- ○ Other industry organisation, or acting on behalf of such organisations
- ○ Academia
- ◉ Civil Society Organisation
- ○ Public authority
- ○ Citizen
- ○ Others

If other, please specify

* In which sector do you operate?

- ☐ Information technology
- ☐ Public sector
- ☐ Law enforcement
- ☐ Security
- ☐ Media

- [ ] Healthcare
- [ ] Employment
- [ ] Education
- [ ] Consumer services
- [ ] Business services
- [ ] Banking and finance
- [ ] Manufacturing
- [ ] Energy
- [ ] Transport
- [ ] Telecommunications
- [ ] Retail
- [ ] E-commerce
- [ ] Advertising
- [ ] Arts & Entertainment
- [x] Others
- [ ] Not applicable

If other, please specify

human rights

* Please briefly describe the activities of your organisation or yourself:

1000 character(s) maximum

ECNL is a civil society organisation working to protect and promote civil society and civic freedoms.

Is your organisation submitting a collective answer on behalf of other organisations?

- ( ) Yes
- (•) No
- ( ) Not applicable

Please specify

All contributions to this consultation may be made publicly available. Therefore, please do not share any confidential information in your contribution. For organisations, their organisation details would be published while respondent details can be requested to be anonymised. Individuals can request

to have their contribution fully anonymised. Your e-mail address will never be published.

Please select the privacy option that best suits you. Privacy options default based on the type of respondent selected.

* **For natural persons: Contribution publication privacy settings**

If you act in your personal capacity: All contributions to this consultation may be made publicly available. You can choose whether you would like your details to be made public or to remain anonymous.

- ○ **Anonymous.** The type of respondent that you responded to this consultation as, your answer regarding residence, and your contribution may be published as received. Your name will not be published. Please do not include any personal data in the contribution itself.
- ○ **Public.** Your name, the type of respondent that you responded to this consultation as, your answer regarding EU nationality, and your contribution may be published.
- ◉ Not applicable

* **For organisations: Contribution publication privacy settings**

If you represent one or more organisations: All contributions to this consultation may be made publicly available. You can choose whether you would like respondent details to be made public or to remain anonymous.

- ○ **Anonymous.** Only organisation details may be published: The type of respondent that you responded to this consultation as, the name of the organisation on whose behalf you reply as well as its size, its presence in or outside the EU and your contribution may be published as received. Your name will not be published. Please do not include any personal data in the contribution itself if you want to remain anonymous.
- ◉ **Public.** Organisation details and respondent details may be published: The type of respondent that you responded to this consultation as, the name of the organisation on whose behalf you reply as well as its size, its presence in or outside the EU and your contribution may be published as received. Your name will also be published.
- ○ Not applicable

Privacy statement

☑ I acknowledge the attached privacy statement.

Privacy_Statement.pdf

Questionnaire

Section 1. Questions in relation to the definition of an AI system

The **definition of an AI system** is key to understanding the scope of application of the AI Act. It is a first step in the assessment whether an AI system falls into the scope of the AI Act.

The definition of an 'AI system' as provided in Article 3(1) AI Act is aligned with the OECD definition: '*AI system means a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments.*'

Recital 12 provides further clarifications on the definition of an AI system.

The following seven elements can be extracted from the definition:

1) 'a machine-based system'
2) 'designed to operate with varying levels of autonomy'
3) 'may exhibit adaptiveness after deployment',
4) 'for explicit or implicit objectives',
5) 'infers, from the input it receives, how to generate outputs'
6) 'predictions, content, recommendations, or decisions'
7) 'can influence physical or virtual environments'

Question 1: Elements of the definition of an AI system

The definition of the AI system in Article 3(1) AI Act can be understood to include the above mentioned main elements. The key purpose of the definition of an AI system is to provide characteristics that distinguish AI systems from

‘simpler traditional software systems or programming approaches’. A key distinguishing characteristic of an AI system is its capability to infer, from the input it receives how to generate outputs. This capability of inference, covers both the process of obtaining output in the post-deployment phase of an AI system as well as the capability of an AI system to derive models or algorithms or both from inputs or data at the pre-deployment phase. Other characteristics of an AI system definition such as the system’s level of autonomy, type of objectives, and degree of adaptiveness, help to define main elements of the AI system as well as to provide clarity on the nature of the AI system but are not decisive for distinguishing between AI systems and other type of software systems. In particular, AI systems that are built on one of the AI techniques but remain static after deployment triggered questions related to the scope of the AI Act, understanding of the concept of inference and the interplay between the different characteristics of the AI system definition. The guidelines are expected to provide explanation on the main elements of the AI system definition.

1.1: Based on Article 3(1) and Recital 12 AI Act, what elements of the definition of an AI system, in particular, require further clarification in addition to the guidance already provided in Recital 12?

Elements of an AI system - please rate the importance of further clarification from 1 to 10, 10 indicating 'most important':

'a machine based system'

Only values between 1 and 10 are allowed

1

'designed to operate with varying levels of autonomy'

Only values between 1 and 10 are allowed

5

'may exhibit adaptiveness after deployment'

Only values between 1 and 10 are allowed

1

'for explicit or implicit objectives'

Only values between 1 and 10 are allowed

1

'infers, from the input it receives, how to generate outputs'

Only values between 1 and 10 are allowed

10

'predictions, content, recommendations, or decisions'

Only values between 1 and 10 are allowed

1

'can influence physical or virtual environments'

Only values between 1 and 10 are allowed

1

Explain why one or more of these elements require further clarification and what part of this element needs further practical guidance for application in real world applications?

1500 character(s) maximum

Autonomy: In AI systems, autonomy manifests across multiple dimensions and operational domains, with systems exhibiting varying degrees of autonomous functionality in different aspects of their operation. Any attempt to in the context of AI Act accountability requirements would inevitably lead to arbitrary distinctions between systems made by providers with vested interest to escape AIA requirements. This would harm both fundamental rights (FR) and the internal market's functioning. Hence, we urge not to further define/narrow down autonomy in the guidelines.

Inference: This is defined as the process of deriving conclusions from an input using any valid reasoning method, including machine learning algorithms or logical rules employed in expert systems. We assert that no alternative definition exists that aligns consistently with established principles of legal interpretation.

Importance of context: The suggested definition overlooks risks to FR. Addressing high-risk elements is crucial to ensure effective FR protection. Hence, the elements of the definition relating to impact of systems (i.e. 'predictions, content, recommendations, or decisions' and ‘'can influence physical or virtual environments') should be given more weight than their technical aspects in determining whether or not an AI system is in the scope of the Act. This is in line with the OECD guidance on importance of context in determining scope of the AI definition.

Question 2: Simple software systems out of scope of the definition of an AI system

The AI Act does not apply to all software systems but only to systems defined as 'AI systems' in accordance with Article 3(1) AI Act. According to recital 12, the notion of AI system should be distinguished from ‘simpler traditional software systems or programming approaches and should not cover systems that are based on the rules defined solely by natural persons to automatically execute

operations'. In particular the use of statistical methods, such as logistic regression, triggered questions related to the conditions under which certain software systems should be considered out of the scope of AI system definition. The Commission guidelines are expected to provide methodology for distinguishing AI systems from simpler traditional software systems or programming approaches and thus would help define systems that are outside the scope of the AI Act.

Please provide examples of software systems or programming approaches that **does not fall** under the scope of the AI system definition in Article 3(1) AI Act and explain why, in your opinion, the examples are not covered by one or more of the seven main elements of the definition of an AI system in Article 3(1) AI Act.

1500 character(s) maximum

The fact that the question excludes some systems a priori from AIA based on technical implementation raises significant concerns about potential loopholes, enforcement and alignment with international legal norms. Research shows that neural networks can be converted into functionally equivalent decision trees or rule-based systems (see TREPAN). This poses a fundamental challenge: developers could bypass regulation by converting AI systems into rule-based versions with the same functionality and risks. Hence, regulation must focus on potential harm, not just technical methods. The OECD guidelines support this by advocating for a flexible, inclusive definition of AI, covering systems from simple to complex. Referring to the explanatory memorandum: specific techniques may raise particular policy considerations, while certain use contexts may warrant heightened scrutiny. Therefore, we recommend to presume that all algorithmic and predictive systems fall within the scope of the Act unless proven otherwise, and only on a case-by-case basis. This aligns with international legal norms, placing the burden on relevant actors to demonstrate their qualification for any exemptions, see for example Article 52.1 of the EU Charter of Fundamental Rights. The harm caused by simple systems, like the SyRI system in the Netherlands, highlights the need for comprehensive regulation. Technical implementation should not serve as a basis for automatic exclusion from oversight.

Section 2. Questions in relation to the prohibitions (Article 5 AI Act)

Article 5 AI Act prohibits the placing on the EU market, putting into service, or the use of certain AI systems that can be misused and provide novel and powerful tools for manipulative, exploitative, social control and/or surveillance practices.

The Commission guidelines are expected to include an introductory section explaining the general interplay of the prohibitions with other Union legal acts, the high-risk category and general-purpose AI systems as well as relevant specifications of some horizontal concepts such as provider and deployer of AI

systems, 'placement on the market', 'putting into service' and 'use' and relevant exceptions and exclusions from the scope of the AI Act (e.g. research, testing and development; military, defense and national security, personal non-professional activity).

Pursuant to Article 5(1) AI Act, the following practices are prohibited in relation to AI systems:

Article 5(1)(a) – Harmful subliminal, manipulative and deceptive techniques

Article 5(1)(b) – Harmful exploitation of vulnerabilities

Article 5(1)(c) – Unacceptable social scoring

Article 5(1)(d) – Individual crime risk assessment and prediction (with some exceptions)

Article 5(1)(e) – Untargeted scraping of internet or CCTV material to develop or expand facial recognition databases

Article 5(1)(f) – Emotion recognition in the areas of workplace and education (with some exceptions)

Article 5(1)(g) – Biometric categorisation to infer certain sensitive categories (with some exceptions)

Article 5(1)(h) – Real-time remote biometric identification (RBI) in publicly accessible spaces for law enforcement purposes (with some exceptions)

This section includes questions on each of the aforementioned prohibitions separately and one final question pertaining to all prohibitions alike and the interplay with other acts of Union law.

A. Questions in relation to harmful subliminal, manipulative or deceptive practices

The prohibition under Article 5(1)(a) AI Act targets AI systems that deploy subliminal techniques, purposefully manipulative or deceptive techniques that materially influence behaviour of people or aim to do so in significantly harmful ways. The underlying rationale of this prohibition is to protect individual autonomy and well-being from manipulative, deceptive and exploitative AI practices that can subvert and impair individuals' autonomy, decision-making, and free choice.

Proposed structure of the guidelines

It is proposed that the Commission guidelines would cover the following aspects regarding Article 5(1)(a) AI Act:

- *Rationale and objectives of the prohibition*
- *Main elements of the prohibition*
    - *AI systems deploying subliminal, purposefully manipulative and deceptive techniques*
    - *with the objective or the effect of materially distorting behaviour*
    - *in a manner (reasonably likely to) cause significant harm*
- *AI systems out of scope of the prohibition*
- *Interplay with other Union law (e.g. data protection, consumer protection, digital services regulation, criminal law)*

Main elements of the prohibition

*Several **cumulative elements must be in place** at the same time for the prohibition in Article 5(1)(a) AI Act to apply:*

*1) The activity must constitute **'placing on the market'** (Article 3(9) AI Act), **'putt***

ing into service' *(Article 3(11) AI Act), or* ***'use'*** *of an AI system (Article 3(1) AI Act). The prohibition applies to both providers and deployers of AI systems, each within their own responsibilities.*

2) The AI system must 'deploy ***subliminal techniques*** *beyond a person's consciousness (e.g. deploying imperceptible images or audio sounds),* ***purposef ully manipulative*** *(e.g. exploiting cognitive biases, emotional or other manipulative techniques) or* ***deceptive techniques****' (e.g. presenting false and misleading information to deceive individuals and influence their decisions in a manner that undermines their free choices). These techniques are alternative, but they can also apply in combination.*

3) The techniques deployed by the AI system should have the ***objective or the effect of materially distorting the behaviour of a person or a group of persons****. The distortion must* ***appreciably impair their ability to make an informed decision, resulting in a decision that the person or the group of persons would not have otherwise made.*** *This requires a substantial impact whereby the technique deployed by the AI system does not merely influence a person's (or group of persons) decision, but should be capable of effectively undermining their individual autonomy and ability to make an informed and independent free choice. This suggests that 'material distortion' involves a degree of coercion, manipulation or deception that goes beyond lawful persuasion that falls outside the ban.*

4) The distorted behaviour must ***cause or be reasonably likely to cause significant harm*** *to that person, another person, or a group of persons. In this context, important concepts that will be examined in the guidelines are the types of harms covered, the threshold of significance of the harm and its reasonable likelihood from the perspective of the provider and/or the deployer. 'Significant harms' implies sufficiently important adverse impacts on physical, psychological health or financial interests of persons and groups of persons that can be compound with broader group and societal harms. The determination of 'significant harm' is fact and context specific, necessitating careful consideration of each case's individual circumstances.*

For the prohibition to apply, all elements must be in place and there must be a causal link between the techniques deployed, the material distortion of the behaviour of the person and the significant harm that has resulted or is reasonably likely to result from that behaviour.

Question 3: Taking into account the provisions of the AI Act, what elements of the prohibition of harmful manipulation and deception do you think require further clarification in the Commission guidelines?

Please select all relevant options from the list

- [ ] *placement on the market, putting into service or use of an AI system*
- [x] *deploying subliminal, purposefully manipulative or deceptive techniques*
- [x] *with the objective or the effect of materially distorting behaviour of a person or groups of persons*
- [x] *in a manner that causes or is reasonably likely to cause significant harm*
- [ ] *none of the above*

Please explain why the elements selected above require further clarification and what needs to be further clarified in the Commission guidelines?

1500 character(s) maximum

We are concerned that if the Commission guidelines fail to provide clarifications on the notions of 'subliminal, manipulative and deceptive techniques' this prohibition will be toothless in practice. Subliminal techniques have long been studied in psychology and discussed in the marketing literature as a way to potentially influence consumer behavior, but legally speaking is not a well-established concept. Recitals 28 and 29 offer some useful insights and may play an important role in determining what kinds of techniques fall under the definition.
For instance, references have been made to the legal marketing practices that fall out of the scope of Article 5(1)(a). Further clarification and examples are needed to understand when the use of subliminal, manipulative or deceptive techniques by AI systems will render such advertising practices illegal.
Additionally, such AI systems are often used to create and disseminate disinformation with the explicit objective of materially influencing people's voting behaviours. The guidelines must clarify how to establish that the use of such techniques, including via synthetic media (e.g. "deepfakes"), falls in scope of the provision.
Finally, the cumulative and very restrictive element of 'significant harm' needs to be thoroughly examined and analysed. This is the key in rendering this article an empty shell or actual protection against practices that undermine and breach the fundamental rights and values of the EU.

Question 4: Do you have or know concrete examples of AI systems that in your opinion fulfil all elements of the prohibition described above?

- (•) Yes
- ( ) No

Please specify the concrete AI system, how it is used in practice and how all the necessary elements described above are fulfilled

1500 character(s) maximum

The manipulation of public opinion through social media remains a growing threat to democracies around the world, according to the 2020 media manipulation survey from the Oxford Internet Institute, part of the University of Oxford. To better understand the gravity of deception and manipulation, especially on social media, it is worth mentioning cases like Cambridge Analytica, where user data from Facebook was used to manipulate and influence voting behavior during U.S. elections, and Russia's interference in the 2016 U.S. presidential election, where social media accounts powered by AI systems were employed to spread disinformation, or the recent TikTok case in Romania.
Organized social media manipulation campaigns operate in 81 countries, up from 70 countries in 2019, with global misinformation being produced on an industrial scale by major governments, public relations firms and political parties. It describes how disinformation has become a common strategy of cyber manipulation, with more than 76 of the 81 countries deploying disinformation as part of political communication.
This is just one example demonstrating the power of new technologies in understanding how to shift our opinion, removing our freedom to make informed decisions, going against Union's universal values of human dignity, freedom, equality, solidarity, the principles of democracy and the rule of law.

Question 5: Do you have or know concrete examples of AI systems where you need further clarification regarding certain elements of this prohibition to determine whether the AI system is in the scope of the prohibition or not?

- ○ Yes
- ◉ No

Please specify the concrete AI system, how it is used in practice as well as the specific elements you would need further clarification in this regard

1500 character(s) maximum

B. Questions in relation to harmful exploitation of vulnerabilities

The prohibition under Article 5(1)(b) AI Act targets AI systems that exploit vulnerabilities of certain persons or groups of persons that materially influence behaviour of people or aim to do so in a significantly harmful way. The underlying rationale of the prohibition is to protect individual autonomy and well-being from exploitative AI practices that can subvert and impair individuals' autonomy, decision-making, and free choice similar. This prohibition in particular aims to protect those that are most vulnerable and susceptible to manipulation and exploitation because of their specific characteristics that make them

particularly vulnerable due to their age, disability and or specific socio-economic situation.

Proposed structure of the guidelines

It is proposed that the Commission guidelines would cover the following aspects regarding Article 5(1)(b) AI Act:

- *Rationale and objectives of the prohibition*
- *Main elements of the prohibition*
    - *AI system exploiting vulnerabilities due to age, disability or specific socio-economic situation*
    - *with the objective or the effect of materially distorting behaviour*
    - *in a manner (reasonably likely to) cause significant harm*
- *Interplay between the prohibitions in Article 5(1)(a) and (b) AI Act, with the latter acting as lex specialis in case of overlap*
- *AI systems out of scope of the prohibition*
- *Interplay with other Union law (e.g. data protection, non-discrimination law, digital services regulation, criminal law)*

Main elements of the prohibition

*Several **cumulative elements must be in place** at the same time for the prohibition in Article 5(1)(b) AI Act to apply:*

*1) The activity must constitute **'placing on the market'** (Article 3(9) AI Act), **'putting into service'** (Article 3(11) AI Act), or **'use'** of an AI system (Article 3(1) AI Act). The prohibition applies to both providers and deployers of AI systems, each within their own responsibilities.*

*2) The AI system must exploit **vulnerabilities due to age** (covering both children as well as elderly), **disability** (as defined in EU equality law encompassing a wide range of physical, mental, intellectual and sensory impairments that hinder full participation of individuals in the society), or **specific socio-economic situations** (e.g. persons living in extreme poverty, ethnic or religious minorities). Vulnerabilities of these persons should be understood to*

encompass a broad spectrum of categories, including cognitive, emotional, physical and other forms of susceptibility that can affect the ability of an individual or a group of persons pertaining to those groups to make informed decisions or otherwise influence their behaviour. 'Exploitation' should be understood as objectively making use of such vulnerabilities in a manner which is harmful for the exploited vulnerable (groups of) persons and/or other persons.

3. The techniques deployed by the AI system should have the ***objective or the effect of materially distorting the behaviour*** *of a person or a group of persons. Article 5(1)(a) and (b) AI Act make use of the same concept and should therefore be interpreted in the same way to the extent they overlap.*

4. The distorted behaviour must ***cause or be reasonably likely to cause significant harm*** *to that person, another person, or a group of persons. Article 5 (1)(a) and (b) AI Act make use of the same concept and should therefore be interpreted in the same way, while taking into account that the harms that can be suffered by vulnerable groups can be particularly severe and multifaceted due to their heightened susceptibility to exploitation.*

For the prohibition to apply, all elements must be in place and there must be a causal link between the vulnerability exploitation by the AI system, the material distortion of the behaviour of the person and the significant harm that has resulted or is reasonably likely to result from that behaviour.

Question 6: Taking into account the provisions of the AI Act, what elements of the prohibition of harmful exploitation of vulnerabilities do you think require further clarification in the Commission guidelines?

Please select all relevant options from the list

- [ ] *placement on the market, putting into service or use of an AI system*
- [x] *exploiting vulnerabilities due to age, disability or specific socio-economic situation*
- [x] *with the objective or the effect of materially distorting behaviour of a person or groups of persons*
- [x] *in a manner that causes or is reasonably likely to cause significant harm*
- [ ]

none of the above

Please explain why the elements selected above require further clarification and what needs to be further clarified in the Commission guidelines?

1500 character(s) maximum

Despite the fact that this paragraph provides a restrictive list of what vulnerability means in the context of AI systems, we believe that there are some overlaps with regards to the prohibited practices in Article 5 paragraph 1a, especially when it comes to distorting someone's behaviour.

The guidelines should provide clarifications on the notions used and especially the term 'specific social or economic situation', in particular on whether this is linked with a person's(individual) social and economic status or if it is associated with a status of social groups such as immigrants, refugees' their families or groups that face social exclusion.

As mentioned above, the notion of "significant harm" needs to be further explained as it is key in rendering a practice prohibited or not. Examples or a tool similar to the risk assessment ones (severity vs likelihood) can be used to measure harm or the likelihood of this harm occurring.

Question 7: Do you have or know concrete examples of AI systems that in your opinion fulfil all elements of the prohibition described above?

- (●) Yes
- ( ) No

Please specify the concrete AI system, how it is used in practice and how all the necessary elements described above are fulfilled

1500 character(s) maximum

- Advanced Analytics for Targeted Advertising
The use of profiling techniques by Facebook to exploit vulnerabilities related to mental health: https://en.panoptykon.org/algorithms-of-trauma-2-how-facebook-feeds-on-your-fears
- Chatbots Spreading Propaganda and Hate Speech
Russian Bots (2016–2020): During various elections, including the 2016 U.S. presidential election, bot accounts amplified divisive content on platforms like Facebook and Twitter, targeting specific racial and ethnic groups to polarize societies.
- AI Agents Promoting Risky Financial Decisions
High-Risk Lending Algorithms: AI-powered lending platforms have been shown to disproportionately offer high-interest loans to economically disadvantaged individuals. ZestFinance, for example, faced criticism for using opaque criteria that potentially perpetuated exploitative lending practices.

Question 8: Do you have or know concrete examples of AI systems where you need further clarification regarding certain elements of this prohibition to determine whether the AI system is in the scope of the prohibition or not?

- ( ) Yes
- (●) No

Please specify the concrete AI system, how it is used in practice as well as the specific elements you would need further clarification in this regard

1500 character(s) maximum

C. Questions in relation to unacceptable social scoring practices

The prohibition under Article 5(1)(c) AI Act aims to prevent 'social scoring' practices that evaluate persons over a certain period of time based on their social behaviour or personal characteristics leading to detrimental and unfair outcomes for certain individuals and groups. The prohibition applies in principle to both the public and the private sector. The underlying rationale of this prohibition is to prevent such unacceptable 'social scoring' practices that may lead to discriminatory and unfair outcomes for certain individuals and groups, including their exclusion from society. The prohibition of 'social scoring' aims to protect in particular the right to human dignity and other fundamental rights, including the right to non-discrimination and equality, to data protection and to private and family life. It also aims to safeguard and promote the European values of democracy, equality and justice.

Proposed structure of the guidelines

It is proposed that the Commission guidelines would cover the following aspects regarding Article 5(1)(c) AI Act:

- *Rationale and objectives of the prohibition*
- *Main elements of the prohibition*
  - *'Social scoring': evaluation or classification based on social behaviour or personal or personality characteristics over a certain period of time*
  - *Whether provided or used by public or private entities*
  - *Leading to detrimental or unfavourable treatment in unrelated social contexts and/or unjustified or disproportionate treatment*
- *AI systems out of scope of the prohibition*
- *Interplay with other Union law (e.g. data protection, non-discrimination)*

Main elements of the prohibition

*Several **cumulative elements must be in place** at the same time for the prohibition in Article 5(1)(c) AI Act to apply:*

*1) The activity must constitute **'placing on the market'** (Article 3(9) AI Act), **'putting into service'** (Article 3(11) AI Act), or **'use'** of an AI system (Article 3(1) AI Act). The prohibition applies to both providers and deployers of AI systems, each within their own responsibilities.*

*2) The AI systems must be intended or used for the **evaluation or classification** of natural persons or groups of persons over a certain period of time based on:*
*(i)their **social behaviour**; or*
*(ii) known, inferred or predicted personal or personality **characteristics**;*

*3) The social score created with the assistance of the AI system must lead to the **detrimental or unfavourable treatment** in one or more of the following scenarios:*
(i) in social contexts unrelated to those in which the data was originally generated or collected; and/or
(ii)treatment that is unjustified or disproportionate to their social behaviour or its gravity.

The detrimental or unfavourable treatment must be the consequence of the score, and the score the cause of the treatment. It is not necessary for the evaluation performed by the AI system to be 'solely' leading to the detrimental or unfavourable treatment (covering thus AI-enabled scoring practices that may be also subject to or combined with other human assessments). At the same time, the AI output has to play a sufficiently important role in the formation of the social score. For the prohibition to apply all elements described above must be in place at the same time.

Question 9: Taking into account the provisions of the AI Act, what elements of the prohibition of social scoring do you think require further clarification in the Commission guidelines?

Please select all relevant options from the list

- [ ] *placement on the market, putting into service or use of an AI system*
- [x] *for the evaluation or classification of natural persons or groups of persons over a certain period of time based on their social behaviour, or known, inferred or predicted personal or personality characteristics*
- [x] *with the social score leading to the detrimental or unfavourable treatment of the person or groups of persons*
- [x] *in social contexts unrelated to those in which the data was originally generated or collected*
- [x] *treatment that is unjustified or disproportionate to their social behaviour or its gravity*
- [ ] *none of the above*

Please explain why the elements selected above require further clarification and what needs to be further clarified in the Commission guidelines?

1500 character(s) maximum

‘Over a certain period of time’ specify that the duration of use is not a consequential factor for application of the ban and human rights implications should be prioritized.
‘Social behaviour’ clarify it encompasses wide-ranging elements. For the Danish Welfare Automation "unusual" living arrangement was one of the main indicators of fraud algorithms, which led to rights violations especially for migrant families and persons with disabilities.
‘Personal or personality characteristics’ must include both personal and non personal data, as the latter can be a proxy for indirect discrimination. In the Dutch child welfare scandal, postal code was a proxy that led to discrimination for people living in poverty and from migrant backgrounds.
‘Social scoring’ must apply to wide-ranging social scoring systems including, but not limited to, employment, education, housing, welfare benefits, health, migration, administration of justice.
‘Social contexts’ must be interpreted in a way that leads to ban excessive and unlawful data collection, as merging of public and private databases, data sharing among authorities, or scraping of open source data.
In assessing ‘unjustified or detrimental treatment’, specify that there must be a high threshold for deployers to argue that treatment is proportionate to social behaviour, prioritizing the protection of fundamental rights and maintaining the burden of proof on employers to demonstrate the lawfulness of the use of the AI system.

Question 10: Do you have or know concrete examples of AI systems that in your opinion fulfil all elements of the prohibition described above?

- (•) Yes
- ( ) No

Please specify the concrete AI system, how it is used in practice and how all the necessary elements described above are fulfilled

1500 character(s) maximum

Netherlands Childcare benefit scandal – welfare fraud detection. Racial profiling was baked into the design of the algorithmic system used to determine whether claims for childcare benefit were flagged as incorrect and potentially fraudulent. The evaluation is based on social behaviour such as being a single mother, have a low income, or personal characteristics that lead to indirect discrimination such as country of origin (proxy for race). Unjustified treatment included intrusion of private life, suspension of benefits, consequential forced evictions, mental burnout.

Danish welfare authority, Udbetaling Danmark (UDK) - fraud detection. The algorithm evaluates and classifies residents over time, with models regularly updated and re-run monthly. It leads to detrimental, unfavorable treatment of people with disabilities, older people, low income groups and migrants, who are flagged for fraud control or investigation and subjected to further monitoring and surveillance, infringing on their right to privacy and risking their right to social security. Irrelevant data including but not limited to travel history, "foreign affiliation" and "unusual" living arrangement are used to flag recipients for fraud investigations, as well as data non-recipients' data such as household or family members' data.

Question 11: Do you have or know concrete examples of AI systems where you need further clarification regarding certain elements of this prohibition to determine whether the AI system is in the scope of the prohibition or not?

- (●) Yes
- ( ) No

Please specify the concrete AI system, how it is used in practice as well as the specific elements you would need further clarification in this regard

1500 character(s) maximum

The difficulty of determining whether a system is a social scoring system stems from the lack of transparency obligations on the deployers of such systems and from the vague formulation of this prohibition. For example, recent Lighthouse Reports and Svenska Dagbladet investigation revealing the discriminatory nature of Swedish welfare automation, also exposed Swedish authorities lack of transparency and refusal to disclose information requested through FOIAs. (https://x.com/gabriels_geiger/status /1861749142369825258). This is not an isolated case. The guidelines therefore need to clarify the burden of proof on authorities to provide sufficient evidence as to how their systems do not qualify as falling under Article 5. In addition, to support a meaningful application of this ban, the Guidelines should reflect the state of play within the European context and refer to existing practices of social scoring, especially in the welfare and migration procedures and in line with civil society calls: (https://www.hrw.org/news/2023/10/09/eu-artificial-intelligence-regulation-should-ban-social-scoring#:~:text=(Brussels%2C%20October%209%2C%202023,regulation's%20prohibition%20on%20social%20scoring.%5D%22%20said%20HMS). We refer to you to the recommendations regarding element clarifications in Q9 and examples from Q10 for this purpose.

D. Questions in relation to individual crime risk assessment and prediction

The prohibition under Article 5(1)(d) AI Act targets AI systems assessing or predicting the risk of a natural person committing a criminal offence solely based on profiling or assessing personality traits and characteristics, without objective

and verifiable facts directly linked to criminal activity and a human assessment thereof. The underlying rationale for the ban is to prevent unacceptable law enforcement practices where AI is used to make an individual a suspect solely based on profiling or their personality traits and characteristics rather than as support of human assessment, which is already based on objective and verifiable facts directly linked to a criminal activity. Such predictive crime and policing AI systems pose an 'unacceptable risk' since they infringe fundamental rights and freedoms in a democracy that is based on rule of law and requires a fair, equal and just criminal legal system. They also endanger individual's liberty without the necessary procedural and judicial safeguards and violate the right to be presumed innocent. Other fundamental rights at risk that the ban aims to safeguard are the right to human dignity, non-discrimination, the right to fair trial, the right to defence, effective remedy, privacy and data protection and the rights of the child if these practices affect children.

Proposed structure of the guidelines

It is proposed that the Commission guidelines would cover the following aspects regarding Article 5(1)(d) AI Act:

- *Rationale and objectives of the prohibition*
- *Main elements of the prohibition*
    - *Individual crime prediction of a natural person committing a criminal offence*
    - *solely based on profiling or the assessment of personality traits and characteristics*
    - *without verifiable facts directly linked to criminal activity and human assessment thereof*
- *Interplay with other Union law (e.g. data protection)*
- *AI systems that are out of the scope of the prohibition (e.g. support of the human assessment)*

Main elements of the prohibition

*Several **cumulative elements must be in place** at the same time for the prohibition in Article 5(1)(d) AI Act to apply:*

*1) The activity must constitute **'placing on the market'** (Article 3(9) AI Act), **'putting into service for this specific purpose'** (Article 3(11) AI Act), or **'use'** of an AI system (Article 3(1) AI Act). The prohibition applies to both providers and deployers of AI systems, each within their own responsibilities.*

*2) The AI system must be intended or used for the specific purpose **of making a risk assessment or prediction of a natural person or persons committing a criminal offence**. The individual crime predictions can be made at any stage of the law enforcement activities such as prevention and detection of crimes, but also investigation, prosecution and execution of criminal penalties. Excluded from the scope are therefore location- and event-based predictions and individual predictions of administrative offences since these are not assessing the risk of individuals **committing a criminal offence**.*

*3) The assessment or the prediction must be **solely** based on either or both of the following:*
*(i)**profiling** of a natural person (defined in Article 4(4) of the General Data Protection Regulation as any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person), or*
*(ii)**assessing a person's personality traits and characteristics** (such as nationality, place of birth, place of residence, number of children, level of debt or type of car)*

*4) Excluded are **AI systems used to support human assessment based on objective and verifiable facts directly linked to a criminal activity**. This means that predictive AI tools could be used for supporting the human assessment of the involvement of a person in the criminal activity if there are objective and verifiable facts linked to a criminal activity on the basis of which a person can be reasonably suspected of being involved in a criminal activity.*

Question 12: Taking into account the provisions of the AI Act, what elements of the prohibition of harmful manipulation and deception do you think require further clarification in the Commission guidelines?

Please select all relevant options from the list

- ☐ *placement on the market, putting into service or use of an AI system*

- [x] *for making risk assessment or prediction of a natural person or persons committing a criminal offence*
- [x] *solely based on the profiling of a natural person or their traits and characteristics*
- [x] *excluded are AI systems used to support human assessment based on objective and verifiable facts directly linked to a criminal activity*
- [ ] *none of the above*

Please explain why the elements selected above require further clarification and what needs to be further clarified in the Commission guidelines?

1500 character(s) maximum

Clarify that the risk of committing a criminal offence is not limited to the sole likelihood of criminality but it includes other proxy elements such as the likelihood of being registered in a police system. Specify that criminal offence includes behaviours qualified as such in Member States and EU legal frameworks. In the migration area, if being irregular or being classified as presenting a risk to public security qualifies as criminal activity, it should be covered by the ban. Clarify if 'based solely' refers to both 'the profiling of a natural person' and 'on assessing their personality traits and characteristics', or it refers only to profiling. Mandate safeguards to prevent exploitation of the exception. Clearly define the meaning of "support" and ensure its narrow definition. If the output of a predictive policing system plays a "determining role" it should be in scope. Define what "objective and verifiable facts" mean and ensure a level of protection through robust independent oversight to avoid bias assessment. In the case of the Amsterdam Top400, non-criminal justice factors were used as indicators of criminality ('you have changed primary school at least 3 times'), as well as mere suspicion of involvement with crime, without actual evidence. Clarify that any use has to be authorised following a 'reasoned request' and subject to approval by a supervisory authority. Authorities that apply this exemption must bear the burden of proving the use will not lead to rights

Question 13: Do you have or know concrete examples of AI systems that in your opinion fulfil all elements of the prohibition described above?

- (•) Yes
- ( ) No

Please specify the concrete AI system, how it is used in practice and how all the necessary elements described above are fulfilled

1500 character(s) maximum

Netherlands, Top 600 – Amsterdam Municipality, police & social services, created in 2012. The Risk modelling and profiling system attempted to profile the 'top 600' young people, over the age of 16, who are most at risk of committing 'High Impact Crime' in the future. The assessment was based on the profiling of individuals using criteria which include having been arrested as a suspect for a high-impact crime, having been presented to a bankruptcy judge. The consequences of being included in the Top600 list were disastrous and included various forms of punishments such as police raids and arrests. These criminal justice consequences occur without any formal trial or assessment of the relevant evidence by a judge or judicial process

RisCanvi, implemented in Catalonia's criminal justice system since 2009, uses predictive algorithms to assess recidivism risk with minimal human oversight. External adversarial audits have revealed significant flaws, including arbitrary correlations in risk factors, insufficient reliability, and a lack of transparency. These issues raise serious concerns about judicial fairness, as decisions may be influenced by biased or inaccurate predictions. Furthermore, the system's deficiencies risk disproportionately affecting marginalized communities, perpetuating systemic discrimination, and undermining trust in the justice system.

Question 14: Do you have or know concrete examples of AI systems where you need further clarification regarding certain elements of this prohibition to determine whether the AI system is in the scope of the prohibition or not?

- ○ Yes
- ◉ No

Please specify the concrete AI system, how it is used in practice as well as the specific elements you would need further clarification in this regard

1500 character(s) maximum

Question 15: Do you have or know concrete examples of AI systems that fulfil all necessary criteria for the prohibition to apply, but fall under the exception of systems that support the human assessment of the involvement of a person in a criminal activity, based on objective and verifiable facts linked to a criminal activity?

- ◉ Yes
- ○ No

Please specify the concrete AI system, how it is used in practice and which exception would apply and why

1500 character(s) maximum

The Guidelines should clarify that over-reliance and automaton bias must be addressed, given that automated risk scoring could determine rather than "support" human assessment. In the case of the recently investigated Swedish welfare risk scoring algorithms by Lighthouse Reports and partners, people flagged as high risk by algorithms were automatically subject to investigations by fraud controllers within the welfare agency, under an assumption of "criminal intent" right from the start. https://www.lighthousereports.com /investigation/swedens-suspicion-machine/
The guidelines should also specify that suspicion of a crime should clearly not qualify as part of "objective and verifiable" human assessment. There are several instances where LEAs used uncorroborated data and mere suspicion of crime to add individuals to crime list, as in the case of Amsterdam Top400, the National Data Analytics Solution created by the West Midlands Police in England, the Durham's Harm Assessment Risk Tool, the Italian Delia crime prediction system. https://www.fairtrials.org/app/uploads/2021/11 /Automating_Injustice.pdf . Because of the biased nature of the concepts of 'support' and 'objective and verifiable', the application of the exception should be allowed only following a request to an independent supervisory authority.

E. Questions in relation to untargeted scraping of facial images

Article 5(1)(e) AI Act prohibits AI systems with the specific purpose of creating or expanding facial recognition databases through untargeted scraping of the internet or CCTV footage.

As to the rationale of the prohibition, untargeted scraping of a large number of facial images from the Internet or CCTV material, along with associated metadata and information, without consent of the data subject(s), to create large-scale facial databases, violates individuals' rights and individuals lose the possibility to be anonymous. Recital 43 of the AI Act justifies the prohibition of Article 5(1)(e) AI Act based on the 'feeling of mass surveillance' and the risks of 'gross violations of fundamental rights, including the right to privacy'.

Proposed structure of the guidelines

It is proposed that the Commission guidelines would cover the following aspects regarding Article 5(1)(e) AI Act:

- *Rationale and objectives of the prohibition*
- *Main elements of the prohibition*
    - *Facial recognition databases*
    - *through untargeted scraping of facial images*
    - *from the internet or CCTV footage*
- *AI systems out of scope of the prohibition*
- *Interplay with other Union law (e.g. data protection)*

Main elements of the prohibition

*Several **cumulative elements must be in place** at the same time for the prohibition in Article 5(1)(e) AI Act to apply:*

*1) The activity must constitute **'placing on the market'** (Article 3(9) AI Act), **'putting into service for this specific purpose'** (Article 3(11) AI Act), or **'use'** of an AI system (Article 3(1) AI Act). The prohibition applies to both providers and*

deployers of AI systems, each within their own responsibilities.

*2) The AI system must be intended or used for the specific purpose of untargeted scraping. The prohibition applies to **scraping AI systems** that are placed on the market or being put into service 'for this specific purpose' of **untargeted scraping of the internet/CCTV** material. This implies that the prohibition does not apply to all scraping tools with which one can build up a database, but only to tools for untargeted scraping.*

*3) The prohibition covers AI system used to **create or expand facial recognition databases**. Database in this context refers to any collection of data, or information, that is specially organized for rapid search and retrieval by a computer. A facial recognition database is a technology that matches a human face from a digital image or video frame against a database of faces, compares it to the database and determines whether there is a match in the database.*

*4) The sources of the images are either the **Internet or CCTV footage**.*

Question 16: Taking into account the provisions of the AI Act, what elements of the prohibition of untargeted scraping of facial images do you think require further clarification in the guidelines?

Please select all relevant options from the list

- [ ] *placement on the market, putting into service or use of an AI system*
- [x] *for creating or expanding facial recognition databases*
- [x] *through untargeted scraping of facial images*
- [ ] *from the internet or CCTV footage*
- [ ] *none of the above*

Please explain why the elements selected above require further clarification and what needs to be further clarified in the guidelines?

1500 character(s) maximum

The guidelines must specify that in order to be considered targeted (and therefore not subject to this prohibition), faces scraped from the internet or a CCTV footage must be likely to have a link to the commission of a specific crime. This is in line with case law of the Court of Justice of the EU. (C-511/18)

Otherwise, the facial images of innocent people/passers-by could be scraped because they appear in the same CCTV footage. Images of all people from a particular country, or with a particular attribute (e.g. protesters), could be scraped, with the false claim that this is a form of targeting. This sort of scraping must be expressly considered as within the scope of the ban, to ensure consistency with the fundamental rights to

privacy, non-discrimination, data protection, and freedoms of expression and assembly.

We further urge the Commission to prevent loopholes by removing the proposed definition of a facial image database. Systems which are intended for and used for untargeted scraping of people's faces, such as Clearview AI and PimEyes, directly fit in the prohibition in Article 5.1.e, in light of recital 43. The Commission's guidelines therefore must not create a loophole for the use of these systems, given that the political intention of the AI Act is clearly to prohibit them.

Question 17: Do you have or know concrete examples of AI systems that in your opinion fulfil all elements of the prohibition described above?

- (●) Yes
- ( ) No

Please specify the concrete AI system, how it is used in practice and how all the necessary elements described above are fulfilled

1500 character(s) maximum

Clearview AI and PimEyes

Question 18: Do you have or know concrete examples of AI systems where you need further clarification regarding certain elements of this prohibition to determine whether the AI system is in the scope of the prohibition or not?

- ( ) Yes
- (●) No

Please specify the concrete AI system, how it is used in practice as well as the specific elements you would need further clarification in this regard

1500 character(s) maximum

F. Questions in relation to emotion recognition

Article 5(1)(f) AI Act prohibits AI systems to infer emotions in the areas of workplace and education institutions except for medical or safety reasons.

As to the rationale of the prohibition, emotion recognition technology is quickly evolving and comprises different technologies and processing operations to detect, collect, analyse, categorise, re- and interact and learn emotions from persons. Emotion recognition can be used in multiple areas and domains for a

wide range of applications, such as for analysing customer behaviour, targeted advertising, in the entertainment industry, in medicine and healthcare, in education, employment, wellbeing, or for law enforcement and public safety.

Emotion recognition can lead to 'discriminatory outcomes and can be intrusive to the rights and freedoms of the concerned persons', in particular the right to privacy. It is therefore in principle prohibited in asymmetric relationships in the context of workplace and education institutions, where both workers and students are in particularly vulnerable positions. The AI Act states in Recital 44 that there are 'serious concerns about the scientific basis of AI systems aiming to identify or infer emotions, particularly as expression of emotions vary considerably across cultures and situations, and even within a single individual. Among the key shortcomings of such systems are the limited reliability, the lack of specificity and the limited generalisability.' At the same time, emotion recognition in specific use contexts, such as for safety and medical care (e.g. health treatment and diagnosis) has benefits and is therefore not prohibited. In such cases, emotion recognition is classified as a high-risk AI system and subjected to requirements aimed to ensure accuracy, reliability and safety.

Proposed structure of the guidelines

It is proposed that the Commission guidelines would cover the following aspects regarding Article 5(1)(f) AI Act:

- *Rationale and objectives of the prohibition*
- *Main elements of the prohibition*
    - *AI systems to infer emotions*
    - *Identification and inference of emotions*
    - *Emotions*
    - *On the basis of their biometric data*
- *Limitation of the prohibition to workplace and educational institutions*
    - *Workplace*
    - *Educational institutions*
- *Exceptions for medical and safety reasons*
- *More favourable Member State law*
- *AI systems out of scope of the prohibition*

- *Interplay with other Union law (e.g. data protection)*

Main elements of the prohibition

*Several **cumulative elements must be in place** at the same time for the prohibition in Article 5(1)(f) AI Act to apply:*

*1) The activity must constitute **'placing on the market'** (Article 3(9) AI Act), **'putting into service for this specific purpose'** (Article 3(11) AI Act), or **'use'** of an AI system (Article 3(1) AI Act). The prohibition applies to both providers and deployers of AI systems, each within their own responsibilities.*

*2) AI systems to infer emotions, as defined in the light of Article 3(39) AI Act, are systems for **identifying or inferring emotions or intentions of natural persons on the basis of their biometric data**. 'Identification' occurs when the processing of the biometric data (for example, of the voice or a facial expression) allows to directly compare and identify with an emotion that has been previously programmed in the emotion recognition system. 'Inferring' is done by deducing information generated by analytical and other processes by the system itself. In this case, the information about the emotion is not solely based on data collected on the natural person, but it is concluded from other data, including machine learning approaches that learn from data how to detect emotions. Emotions have to be defined in a broad sense, but do not include physical states such as pain or fatigue and readily apparent expressions such as smiles.*

*3) The prohibition in Article 5(1)(f) AI Act is limited to emotion recognition systems in the '**areas of workplace and educational institutions**', because there is a power imbalance, an asymmetric relation and a risk of continuous surveillance.*

*4) The prohibition contains an explicit exception for emotion recognition systems used in the areas of the workplace and educational institutions **for medical or safety reasons**, such as systems for therapeutical use.*

Question 19: Taking into account the provisions of the AI Act, what elements of the prohibition of emotion recognition in the areas of workplace and education do you think require further clarification in the Commission guidelines?

Please select all relevant options from the list

- [ ] *placement on the market, putting into service or use of an AI system*
- [x] *for identifying or inferring emotions of natural persons*
- [x] *in the area of workplace and educational institutions*
- [x] *except for medical and safety reasons*
- [ ] *none of the above*

Please explain why the elements selected above require further clarification and what needs to be further clarified in the Commission guidelines?

1500 character(s) maximum

The definition must keep in scope systems that identify/infer emotions, but could exclude, as suggested, "physical states such as pain or fatigue". The definition should not, however, exclude "smiles", which are subject to interpretation. As a general point, this ban should be interpreted to prohibit the attribution of a subjective and judgmental quality about a person's inner state or intentions to physical movements or behavioural signals.

The definition must expressly include proxy inferences/judgments, such as "suspicious" or "untrustworthy". Otherwise such inferences could be used as a proxy for emotion, creating a loophole to the prohibition.

We strongly agree with the Commission's interpretation that this prohibition applies in situations of "power imbalance, an asymmetric relation and a risk of continuous surveillance", which should include policing and migration.

Lastly, we are concerned that the exception for "safety or medical" reasons could be misused. The reference to "therapeutic" uses should be deleted. Some providers have claimed that their systems have a therapeutic effect for people with disabilities (although some disability justice advocates have criticised this claim). The intention of the exception, however, is to ensure that medical equipment (e.g. heart monitors) are not ruled out; it is not to allow companies to experiment with untested, unproven 'mind-reading' technology based on claims that it is "therapeutic".

Question 20: Do you have or know concrete examples of AI systems that in your opinion fulfil all elements of the prohibition described above?

- (•) Yes
- ( ) No

Please specify the concrete AI system, how it is used in practice and how all the necessary elements described above are fulfilled

1500 character(s) maximum

iBorderCTRL was a pilot project designed to perform emotion recognition of people travelling to the EU and predict if they are being truthful in their immigration interviews. The purpose of the system was to assist

border guards in their job to assess immigration applications. It clearly falls within the definition of an emotion recognition system, and it is in a workplace context (the system is being used for the work of the border guard) where there is a profound power imbalance;

Rosalyn (Rosalyn's StableSight Model) (partnering with Synap) was an AI system used for proctoring /supervising exams, used mainly in UK/US. The system uses machine learning, facial recognition, and advanced analytics to detect irregularities and ensure that students adhere to exam protocols. It works by continuously monitoring exam sessions through computer webcams and microphones, analyzing data points such as eye movement, voice, and even keystrokes to identify patterns that may indicate dishonest behavior. Even if the system would meet highest safeguards and a human would be involved in the decision-making process, this system would be prohibited in the EU for the above reasons and due to the power imbalance, an asymmetric relation and a risk of continuous surveillance its use would lead to – exactly what the prohibition aims to protect against.
https://synap.ac/online-exam-platform/proctoring/#rosalyn

Question 21: Do you have or know concrete examples of AI systems where you need further clarification regarding certain elements of this prohibition to determine whether the AI system is in the scope of the prohibition or not?

- (●) Yes
- ( ) No

Please specify the concrete AI system, how it is used in practice as well as the specific elements you would need further clarification in this regard

1500 character(s) maximum

The guidelines should clarify that legitimate health and safety systems such as voice monitors that analyze emergency calls to detect if a person is having a heart attack; safety systems to detect if personnel are wearing protective headgear; systems detecting driver fatigue are *not* emotion recognition systems.

At the same time, we urge the Commission to clarify that systems which attribute an emotion to the person presented as medical or safety tools, should not be categorized as medical or safety devices, given that they suffer from serious, fundamental flaws in their scientific underpinnings and therefore could lead to serious life-threatening consequences for persons subjected to these tools. We urge the Commission to make this distinction in the guidelines.

Question 22: Do you have or know concrete examples of AI systems that fulfil all necessary criteria for the prohibition to apply, but fall under the exception of medical and safety reasons?

- (●) Yes
- ( ) No

Please specify the concrete AI system, how it is used in practice and which exception would apply and why

1500 character(s) maximum

Some stakeholders did in the past and will mistakenly argue that health and safety systems are excluded, such as: medical systems like voice monitors that analyse emergency calls to detect if a person is having a heart attack; safety systems to detect if personnel are wearing protective headgear; or even if drivers are falling asleep. These systems do not need to benefit from the exception because they are not emotion recognition systems. Medical and health systems must be based in scientific evidence, whereas emotion recognition systems are pseudoscience. We urge the Commission to make this distinction in their guidelines, between genuine medical systems with the objective of capturing of physical or physiological signs (e.g. a heartbeat), in contrast to emotion recognition systems that try to establish a causality with the person's inner state or intentions.

Emotion recognition systems are systems that specifically ascribe an emotion, intention or proxy for the emotion like 'untrustworthy'', to the input. 'Tired', 'not wearing headgear' or 'having a heart attack' are not emotions. Such systems would only be prohibited if they then attributed an emotion to the person

---

G. Questions in relation to biometric categorisation

Article 5(1)(g) AI Act prohibits biometric categorisation systems (as defined in Article 3(40) AI Act) that categorise individually natural persons based on their biometric data to deduce or infer their race, political opinions, trade union membership, religious or philosophical beliefs, sex life or sexual orientation. This prohibition does not cover the lawful labelling, filtering or categorisation of biometric data sets acquired in line with Union or national law according to biometric data, which can for example be used in the area of law enforcement (Recital 30 AI Act).

As to the rationale of the prohibition, AI-based biometric categorisation systems for the purpose of assigning natural persons to specific groups or categories relating to aspects such as sexual or political orientation or race violate human dignity and pose significant risks to other fundamental rights such as privacy and discrimination.
A wide variety of information, including 'sensitive' information can be extracted, deduced or inferred from biometric information, even without the individuals knowing it, to categorise them. This can lead to unfair and discriminatory treatment, for example when a service is denied because somebody is considered to be of a certain race.

Proposed structure of the guidelines

It is proposed that the Commission guidelines would cover the following aspects regarding Article 5(1)(g) AI Act:

- *Rationale and objectives of the prohibition*
- *Main elements of the prohibition:*
  - *Biometric categorisation system*
  - *Persons are individually categorised based on their biometric data*
  - *To deduce or infer their race, political opinions, trade union membership, religious or philosophical beliefs, sex life or sexual orientation*
  - *On the basis of their biometric data*
- *AI systems out of scope of the prohibition*
  - *Labelling and filtering based on biometric data*
- *Interplay with other Union law (e.g. data protection)*

Main elements of the prohibition

*Several **cumulative elements must be in place** at the same time for the prohibition in Article 5(1)(g) AI Act to apply:*

*1) The activity must constitute **'placing on the market'** (Article 3(9) AI Act), **'putting into service for this specific purpose'** (Article 3(11) AI Act), or **'use'** of an AI system (Article 3(1) AI Act). The prohibition applies to both providers and deployers of AI systems, each within their own responsibilities.*

*2) The AI system must be a **biometric categorisation system** for the purpose of assigning natural persons to specific categories on the basis of their biometric data, unless it is ancillary to another commercial service and strictly necessary for objective technical reasons (Article 3(40) AI Act).*

*3) **Individual persons** are categorised,*

*4) Based on their **biometric data** (Article 3(34) AI Act),*

*5) Article 5(1)(g) AI Act prohibits only biometric categorisation systems which have as objective **to deduce or infer a limited number of sensitive characteristics: race, political opinions, trade union membership, religious***

***or philosophical beliefs, sex life or sexual orientation**.*

*The prohibition does not **cover labelling or filtering of lawfully acquired biometric datasets**, including in the field of law enforcement.*

Question 23: Taking into account the provisions of the AI Act, what elements of the prohibition of biometric categorisation to infer certain sensitive characteristics do you think require further clarification in the Commission guidelines?

Please select all relevant options from the list

- ☐ *placement on the market, putting into service or use of an AI system*
- ☑ *that is a biometric categorisation system individually categorising natural persons based on their biometric data*
- ☑ *to deduce or infer their race, political opinions, trade union membership, religious or philosophical beliefs, sex life or sexual orientation*
- ☑ *excluded are labelling or filtering of lawfully acquired biometric datasets, including in the field of law enforcement*
- ☐ *none of the above*

Please explain why the elements selected above require further clarification and what needs to be further clarified in the Commission guidelines?

1500 character(s) maximum

"Individually categorising" should not be able to be used as a loophole to prevent the same harmful practices being applied to a group;
Deductions/inferences of "race" should also be interpreted to include inferences about "ethnicity", and those about "sex life or sexual orientation" should also be considered to include gender identity, in accordance with UN conventions on sexual orientation and gender identity;
The consultation document wrongly suggests that labeling or filtering can be permissible in the context of law enforcement among others, whereas the AI Act text is clear that this exception applies only in the law enforcement context. This should be corrected;
The labeling or filtering of lawfully-acquired biometric datasets should be clarified to specifically apply only in forensic contexts, which by definition occur ex post;

Question 24: Do you have or know concrete examples of AI systems that in your opinion fulfil all elements of the prohibition described above?

- ◉ Yes
- ○ No

Please specify the concrete AI system, how it is used in practice and how all the necessary elements described above are fulfilled

1500 character(s) maximum

1. In the migration context, biometric categorisation systems can be used throughout various migration procedures, with the purpose of assisting migration authorities in assessing the credibility of the applicant's claim. Dialect recognition systems used throughout asylum procedures fall under the scope of this prohibition. The system used by the the German Federal Office for Migration and Refugee for the examination of asylum applications. In full violation of the presumption of innocence, the dialect recognition systems is used to verify that asylum applicants are from where they claim to be. The systems process voice data, which qualifies as biometric data, and assign the person to a country of origin, hence inferring ethnicity. Deductions/inferrences of "race" should be interpreted to include inferrences about "ethnicity", hence dialect recognition systems are prohibited under Article 5(1)(g)
2. Viso AI, Deepface is a face recognition and facial attribute analysis library for Python. One of the tasks is the facial attribute analysis (ie. describing the visual properties of face images). Accordingly, facial attributes analysis is used to extract attributes such as age, gender classification, emotion analysis, or race/ethnicity prediction. Given the system categorises on the basis of assumed race it should be prohibited. (https://viso.ai/computer-vision/deepface/

Question 25: Do you have or know concrete examples of AI systems where you need further clarification regarding certain elements of this prohibition to determine whether the AI system is in the scope of the prohibition or not?

- Yes
- No

Please specify the concrete AI system, how it is used in practice as well as the specific elements you would need further clarification in this regard

1500 character(s) maximum

Question 26: Do you have or know concrete examples of AI systems that fulfil all necessary criteria for the prohibition to apply, but fall under the exception of labelling or filtering of lawfully acquired biometric datasets?

- Yes
- No

Please specify the concrete AI system, how it is used in practice and which exception would apply and why

1500 character(s) maximum

H. Questions in relation to real-time remote biometric identification

Article 5(1)(h) AI Act contains a prohibition on real-time use of remote biometric identification systems (Article 3(41) and (42) AI Act) in publicly accessible spaces for law enforcement purposes subject to limited exceptions exhaustively and narrowly defined in the AI Act.

Recital 32 AI Act acknowledges 'the intrusive nature of remote biometric identification systems (RBIS) to the rights and freedoms of the concerned persons, to the extent that it may affect the private life of a large part of the population, evoke a feeling of constant surveillance and indirectly dissuade the exercise of the freedom of assembly and other fundamental rights. Technical inaccuracies of AI systems intended for the remote biometric identification of natural persons can lead to biased results and entail discriminatory effects. Such possible biased results and discriminatory effects are particularly relevant with regard to age, ethnicity, race, sex or disabilities. In addition, the immediacy of the impact and the limited opportunities for further checks or corrections in relation to the use of such systems operating in real-time carry heightened risks for the rights and freedoms of the persons concerned in the context of, or impacted by, law enforcement activities.'

At European level, RBIS are already regulated by EU data protection rules, as they process personal and biometric data for their functioning.

Due to the serious interferences that real-time RBI use for the purpose of law enforcement poses to fundamental rights, its deployment is, in principle, prohibited under the AI Act. However, as most of these fundamental rights are not absolute, objectives of general interest, such as public security, can justify restrictions on exercising these rights as provided by Article 52(1) of the Charter. Any limitation must comply with the requirements of legality, necessity, proportionality and respect for the essence of fundamental rights. Therefore, when the use is strictly necessary to achieve a substantial public interest and when the exceptions are exhaustively listed and narrowly defined, their use outweighs the risks to fundamental rights (Recital 33 AI Act). To ensure that these systems are used in a 'responsible and proportionate manner', their use can only be made if they fall under one of the explicit exceptions defined in Article 5(1)(i) to (iii) AI Act and subject to safeguards and specific obligations

and requirements, which are detailed in Article 5(2)-(7) AI Act. When the use falls under one or more of the exceptions, the remote biometric identification system is classified as a high-risk AI system and subject to requirements aimed to ensure accuracy, reliability and safety.

Proposed structure of the guidelines

It is proposed that the Commission guidelines would cover the following aspects regarding Article 5(1)(h) AI Act:

- *Rationale and objectives of the prohibition*
- *Definition of*
  - *remote biometric identification*
  - *'real-time'*
  - *publicly accessible spaces*
  - *law enforcement purposes*
- *AI systems out of scope of the prohibition*
- *Interplay with other Union law*
- *Conditions and safeguards for exceptions*

Main elements of the prohibition

*Several **cumulative elements must be in place** at the same time for the prohibition in Article 5(1)(h) AI Act to apply:*

*1) The activity must constitute **the 'use' of an AI system** (Article 3(1) AI Act), so, contrary to the previously mentioned prohibitions, this prohibition applies only to deployers of AI systems.*

*2) The AI system must be a **remote biometric identification system** (Article 3 (41) AI Act), i.e. an AI system for the purpose of identifying natural persons, **with out their active involvement**, typically at a distance through the comparison of a person's biometric data with the biometric data contained in a reference database. This **excludes systems for verification or authentication of persons**.*

*3) The system is used in '**real-time**' (Article 3(42) AI Act), i.e. the biometric*

systems capture and further process biometric data 'instantaneously, near-instantaneously or in any event without any significant delay.

*4) The AI system is used in **publicly accessible spaces**, i.e. 'any publicly or privately owned physical space accessible to an undetermined number of natural persons, regardless of whether certain conditions for access may apply, and regardless of the potential capacity restrictions'. This excludes online spaces, border control points and prisons.*

*5) The prohibition of Article 5(1)(h) AI Act applies to **law enforcement purposes** , irrespective of the entity, authority, or body carrying out the activities. Law enforcement is defined in Article 3(46) AI Act as the 'activities carried out by law enforcement authorities or on their behalf for the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, including safeguarding against and preventing threats to public security.' These activities are also those that constitute the subject matters in Article 1 of the Law Enforcement Directive.*

Question 27: Taking into account the provisions of the AI Act, what elements of the prohibition of real-time remote biometric identification for law enforcement purposes do you think require further clarification in the Commission guidelines?

Please select all relevant options from the list

- [x] *use of an AI system*
- [x] *that is a remote biometric identification system*
- [x] *used 'real-time'*
- [ ] *for law enforcement purposes*
- [ ] *in publicly accessible spaces*
- [ ] *none of the above*

Please explain why the elements selected above require further clarification and what needs to be further clarified in the Commission guidelines?

1500 character(s) maximum

We call on the Commission to clarify several key points. Even though the prohibition only covers use, this that cannot be used to legitimise or give a carte blanche for the development of real-time RBI systems for export, given that (as the recital notes) these systems entail a significant limitation on fundamental rights.

The should guidelines to clarify that the "without their active involvement" clause does not mean that law enforcement actors can place posters or flyers in the surveilled space and claim that people are actively

involved and therefore the definition does not apply.

We also caution against the misuse of the term “authentication” and call on the Commission to clarify this in the Guidelines. It is only through technical “verification” that a person can be “authenticated”. Conversely “authentication” is an outcome, not a process. A system which matches people against a pre-enrolled database cannot be considered authentication, but rather closed-set identification. The guidelines must not allow users of any closed-set identification systems to claim that they are doing “authentication” and are therefore not subject to this prohibition.

Lastly, to prevent circumvention of the ban, we recommend that the “significant delay” entailed to make a system not be considered real-time should be a minimum of 24 hours after capture, and must only relate to the processing of inputs from legally-seized material.

Question 28: Do you have or know concrete examples of AI systems where you need further clarification regarding certain elements of this prohibition to determine whether the AI system is in the scope of the prohibition or not?

- Yes
- No

Please specify the concrete AI system, how it is used in practice as well as the specific elements you would need further clarification in this regard

1500 character(s) maximum

Article 5(1)(h)(i) to (iii) AI Act provides for three exceptions to the prohibition for:

*(1) The **targeted search** of victims of abduction, trafficking in human beings or sexual exploitation of human beings, as well as the search for missing persons, i. e. persons whose existence has become uncertain, because he or she has disappeared.*

*(2) The prevention of a **specific, substantial and imminent threat** to the life or physical safety of natural persons or a genuine and present or genuine and foreseeable threat of a terrorist attack. A terrorist attack can include a threat to life, whereas a threat to life does not necessarily qualify as a terrorist attack.*

*(3) The **localisation and identification of a person suspected of having committed a criminal offence**, for the purpose of conducting a **criminal***

investigation or prosecution or executing a criminal penalty for offences referred to in Annex II *and punishable in the Member States concerned by a custodial sentence or a detention order for a maximum period of at least four years. Annex II of the AI Act provides an exhaustive list of serious crimes for which the real-time use of RBI can be authorised.*

The exceptions have to be authorised by national legislation and comply with certain conditions and safeguards (Article 5(2) to (7) AI Act). These include – among others – temporal, geographic and personal limitations, a duty to perform a fundamental rights impact assessment and to register the system in the EU database (Article 49 AI Act), a need for prior authorisation by a judicial or independent administrative authority, and a notification to the relevant market surveillance authorities and data protection authorities.

Question 29: Do you have or know concrete examples of AI systems that fulfil all necessary criteria for the prohibition to apply, but which could fall under one or more of the exceptions of Article 5(1)(h)(i) to (iii) AI Act?

- ○ Yes
- ○ No

Please specify the concrete AI system, how it is used in practice and which exception would apply and why

1500 character(s) maximum

Question 30: Do you need further clarification regarding one or more of the exceptions of Article 5(1)(h)(i) to (iii) AI Act or the conditions or safeguards under Article 5(2) to (7) AI Act?

- ◉ Yes
- ○ No

Please specify the concrete condition or safeguard and the issues for you need further clarification; please provide concrete examples

1500 character(s) maximum

As recognised by the AI Act, the use of real-time RBI entails significant limitations on fundamental rights. Such uses are contrary to the Charter as this limitation is not necessary and proportionate. This

interpretation is supported by the decision of the Italian DPA on the SARI system, which found that it entails mass surveillance, and by the EDPB's 2023 guidelines.
The Commission should therefore make it clear that not being prohibited by the AI Act does not mean that real-time RBI will be lawful and that uses (including those in Annex III) still require a case-by-case assessment.
To mitigate the serious harm entailed by the AI Act's legitimisation of some RBI uses by virtue of the exceptions to the ban, we further urge the Commission to ensure that the exceptions are duly limited in scope, geography, time and person to minimise the risk of harm, as well as to exclude petty crime (in line with CJEU case law). The guidelines must disallow permanent RBI infrastructure, which is by definition designed for repeated/speculative use.
It is vital that the guidelines interpret a "targeted search" strictly and in line with case law of the CJEU, with clear indications that the person being sought is likely to be in the surveilled location. We call on the guidelines to include specific criteria for how this can be achieved, as well as criteria for defining "imminent threats", in order to prevent generalised preventative surveillance based solely on elevated alert levels.

I. Question in relation to interplay with other Union legislation

The prohibitions under the AI Act are without prejudice to prohibitions and specific rules provided for in other Union legislation such as data protection, consumer protection, digital services regulation, etc. As explained above, each section of the Commission guidelines are expected to explain relevant interplay of the prohibitions in relation to other Union law.

Question 31: Do you have or know concrete examples of AI systems where you need further clarification regarding the application of one or more of the prohibitions under the AI Act in relation to other Union legislation?

- (●) Yes
- ( ) No

Please specify the concrete AI system and the prohibition under the AI Act, the relevant provision of a specific Union legislation and where further clarification is needed

1500 character(s) maximum

The guidelines should clarify that international human rights law and the EU charter of fundamental rights are the central guiding basis to define whether a system poses an unacceptable risk to fundamental rights. Further, the guidelines should strongly emphasise that the objective of the prohibitions is to serve a preventative purpose, preventing the use of systems that pose severe harm to fundamental rights- and therefore must be intercepted broadly in the context of harm prevention. It is imperative that the guidelines specify that all AI systems must be viewed within the wider context of discrimination, racism and prejudice. As an example, the Dutch Foreign Ministry of Foreign Affairs (MFA) used a scoring system in its visa procedures that was found to entrench racist assumptions and led to ethnic profiling of visa applicants. At the same time, a report commissioned in 2022 by the Dutch MFA itself concluded that the agency's internal

culture was riddled with structural racism. Similarly, the UK Home office stopped the use of a similar scoring algorithm used in visa procedures, after a legal complaint denouncing the practice to be racist was launched.

Thank you

Thank you for your interest in participating in the consultation. Please do not forget to click on submit.

Contact

Contact Form