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

1 General Remarks

The call for contributions seeks insights into what elements of the definition of an AI system require further clarification beyond the guidance in Recital 12 of the AI Act. Our response focused on the levels of autonomy, emphasizing their importance in defining and understanding AI systems.

Article 5 of the AI Act prohibits certain AI systems that could be misused for manipulative, exploitative, social control, or surveillance practices. While SRIW refrained from commenting on specific AI systems, we shared our perspective on the points where clearer guidelines regarding prohibited use cases under this article are needed.

All input provided in response to the questions adheres to the character limits specified on the platform.

For your convenience, only the questions to which SRIW provided its feedback have been included.

2 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'

2.1 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': **4**
- 'designed to operate with varying levels of autonomy': **10**
- 'may exhibit adaptiveness after deployment': **8**
- 'for explicit or implicit objectives': **6**
- 'infers, from the input it receives, how to generate outputs': **5**
- 'predictions, content, recommendations, or decisions': **9**
- 'can influence physical or virtual environments': **7**

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?

Levels of autonomy:

Autonomy is directly related to independence of actions from human involvement and capabilities to operate without human intervention (Rec. 12). To determine the applicability of Art. 22 of GDPR (automated decision-making having a legal or similarly significant effect), further clarification is needed in terms of the different levels of autonomy that an AI system can have (e.g., limited autonomy, partial autonomy, full autonomy etc.) to understand the level of human oversight and intervention of each AI system.

Additionally, a clarification of the levels of autonomy of AI systems would be of importance to effectively differentiate an AI system from simpler software. In addition to this, autonomous systems can perpetuate bias and discrimination; understanding the level of autonomy would also help mitigate such biases and discrimination as appropriate legal and technical safeguards would be applied based on the relevant risks per level of autonomy. This would be particularly helpful for SMEs producing, deploying or selling AI systems.

Given that the AI Act permits model providers to demonstrate compliance through alternative adequate means, it is recommended to address these challenges via - sector-specific if need be - codes of conduct. Adherence to such codes could serve as a robust safeguard, ensuring that model-level requirements are consistently met and aligned with the relevant risks and autonomy levels of AI systems.

3 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.

3.1 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), '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 'deploy subliminal techniques beyond a person's consciousness (e.g. deploying imperceptible images or audio sounds), purposefully 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? Additional help available

Please select all relevant options from the list

- ✓ placement on the market, putting into service or use of an AI system
- ✓ deploying subliminal, purposefully manipulative or deceptive techniques
- ✓ with the objective or the effect of materially distorting behaviour of a person or groups of persons
- ✓ 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?

The definitions of "placement on the market," "putting into service," and "use" may overlap, creating ambiguity in enforcement responsibilities between providers and deployers of AI systems. Clear guidelines are essential to delineate these activities, particularly in multi-stakeholder contexts like third-party integration or open-source systems. Clarification is needed for transitions from testing to deployment, roles of intermediaries modifying systems pre-deployment, and cross-border application within the EU.

Operational definitions of terms like "subliminal," "purposefully manipulative," and "deceptive" are critical. For example, clarity is required on what constitutes "beyond a person's consciousness" (subliminal) and distinguishing manipulative versus lawful influential practices

in marketing. This includes identifying imperceptible techniques, providing real-world examples, and ensuring a shared understanding of lawful persuasion thresholds.

The subjective nature of "material distortion" requires guidelines for uniform application, such as criteria for assessing autonomy impairment, defining "substantial impact," and distinguishing intentional distortions from unintentional ones like algorithmic bias. Similarly, "significant harm" and "reasonable likelihood" need nuanced interpretation, considering harm type, context, and stakeholder impact. Examples and standards will support consistent application and risk mitigation.

Clarifying the interplay with laws like GDPR and defining causal link criteria will enhance enforcement and reduce complexity. Sector-specific codes of conduct could serve as safeguards to ensure model-level requirements are met.

3.2 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? Additional help available

Please select all relevant options from the list

- ✓ placement on the market, putting into service or use of an AI system
- ✓ exploiting vulnerabilities due to age, disability or specific socio-economic situation
- ✓ with the objective or the effect of materially distorting behaviour of a person or groups of persons
- ✓ 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?

Regarding option 1 please refer to our answer under Question 5. Clarity is required on the scope of "vulnerability" including criteria for determining when an individual is considered vulnerable due to socio-economic status. For instance, it should be specified whether temporary financial instability qualifies as vulnerability or if the focus is limited to extreme cases. Greater precision is also needed regarding what constitutes "material distortion of behaviour" and how the intent versus the effect of distortion should be evaluated. Additionally, the guidelines could elaborate on the definition of "significant harm" by including examples. Finally, given the overlap in concepts between Articles 5(1)(a) and 5(1)(b), more guidance is needed on how the lex specialis principle applies, especially where manipulation and exploitation intersect. This could prevent misinterpretation and ensure consistency. Finally, addressing these issues within a traditional code of conduct would be highly effective. As a flexible coregulatory tool, codes of conduct are dynamic and adaptable, offering a practical means to continuously refine and maintain necessary guidelines in a way that hard law may not achieve.

3.3 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?

- ✓ placement on the market, putting into service or use of an AI system
- ✓ 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
- ✓ with the social score leading to the detrimental or unfavourable treatment of the person or groups of persons
- ✓ in social contexts unrelated to those in which the data was originally generated or collected
- ✓ 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?

The scope of "evaluation or classification of natural persons or groups over a certain period of time" needs further detail to distinguish prohibited practices from legitimate profiling activities permitted under existing laws, such as data protection regulations. Further, the criterion of "social score leading to detrimental or unfavourable treatment" must outline how causality and significant influence by the AI system are assessed, especially when combined with human decision-making. More precise guidelines are needed to interpret "social contexts unrelated to those in which the data was originally generated or collected" to prevent overreach and ensure legal certainty. Lastly, "unjustified or disproportionate treatment" must be defined with examples, focusing on proportionality standards and the thresholds for determining unjust outcomes. Complementarily, codes of conduct could serve as a robust safeguard, providing for clarification and ensuring that model-level requirements are consistently met.

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
- ✓ for making risk assessment or prediction of a natural person or persons committing a criminal offence
- ☐ solely based on the profiling of a natural person or their traits and characteristics
- ✓ 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?

This exclusion appears to be quite broad, and it would benefit from greater specificity. Firstly, the notion of support requires further explanation, especially the extent of such support. In consideration of the Schufa case, guidelines should take into account if such "support" could play a "determining role" in the decision taken by law enforcement authorities.

Secondly, considering that automated decision-making in relation to individual crime risk assessment and prediction could result in a legal or similar significant effect to the individual, the scope of the exclusion should be defined. Providing guidance on the extent of human involvement in such assessment is highly recommended.

Thirdly, precision is necessary on whether "objective and verifiable facts" are linked to a criminal activity or to the specific criminal activity being investigated. Should any criminal activity be included in the scope, bias and discrimination would likely to be the result.

3.4 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? Additional help available

☐ Please select all relevant options from the list
☐ placement on the market, putting into service or use of an AI system
☐ for creating or expanding facial recognition databases
✓ 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?

Additional guidelines to address what would constitute untargeted scraping of facial images are strongly recommended to be provided to establish clearly what is prohibited under the AI Act.

Further clarification would be helpful with regards to permissible targeted scraping of facial images under both the AI Act and the GDPR. For instance, clarification on whether scraping facial data from publicly available sources, such as social media, is fully prohibited under Article 5(1)(e) or permissible under specific legal bases, could address ambiguities. Such guidance also should include the scope for such targeted scraping of facial images, as well as applicable safeguards to protect the rights and freedoms of data subjects.

It would also be helpful to clarify whether the prohibition applies solely to those who directly create or deploy scraping AI systems or extends to users leveraging such systems for database expansion. If the latter is included, providing clear and concrete criteria to differentiate between these roles would be valuable. It is recommended to consider introducing such criteria into traditional sector specific coregulatory tools to facilitate compliance with the AI Act.

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
- ☐ for identifying or inferring emotions of natural persons
- ✓ in the area of workplace and educational institutions
- ✓ 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?

The precise scope of the term "in the area of workplace" remains unclear. This could be interpreted to include the recruitment process, prior to an individual officially becoming an employee, as well as business meetings involving (non-employee) stakeholders. Additional clarification to confirm this interpretation would be beneficial.

Educational institutions could be understood to include all schools and universities. However, it would be helpful to clarify whether institutions providing professional certifications, workshops, courses and training would also fall within the scope of such prohibition.

It is understood that all other types of emotion recognition AI systems would be considered high risk under the AI Act. Additional clarity would be valuable regarding the permissibility of their use for medical and safety purposes to ensure there are no gaps in the application of the AI Act – e.g., where a prohibited emotion recognition AI might inadvertently be classified as high risk due to insufficient guidance.

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?

Further clarification is needed in several key areas. It remains unclear whether inferred characteristics, such as emotional, social or behavioral traits are included and if emerging characteristics (genetic traits or health-related data) fall under the restriction. Guidelines providing clear definitions to distinguish biometric categorization from identification would be therefore welcomed.

Regarding what qualifies as indirect inference it is understood that when biometric data is used to infer political affiliation or religious beliefs, it must be clear when such inferences cross into prohibited territory. Clear guidelines outlining the scope of the prohibition on the inference of sensitive characteristics would ensure consistency and clarity.

Another key area requiring clarification is the assessment of discriminatory risk and bias mitigation. The guidelines should set clear limits for evaluating discriminatory risks in biometric categorization systems. Concrete practical strategies for bias mitigation are necessary. Additionally, it should be provided guidance on how transparent providers and deployers must be regarding their use of AI with biometric relevance and avoidance of bias.

Finally, clarification is needed on how the AI Act interacts with GDPR, especially concerning the explicit consent required for processing biometric (sensitive) data when AI use is not prohibited. How would both frameworks work together to protect data subjects? An idea might be through sector-specific Codes of Conduct.

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,* ***without 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

- ☐ *use of an AI system*
- ✓ ***that is a remote biometric identification system***
- ✓ ***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?

Guidelines providing a precise definition of remote biometric identification systems, distinguishing them from systems used for verification or authentication would be welcomed in order to avoid ambiguities on the scope of the prohibition. Codes of Conduct could be a valuable tool supporting in this line.

Next, the concept of "real-time use" needs a precise definition, particularly in relation to processing delays. The guidelines should clarify whether slight delays or near-instantaneous processing fall under the real-time category, especially in borderline cases where systems operate with minimal delay. The key technical challenge lies in systems with near-instantaneous or minimal delays potentially attempting to avoid the prohibition by claiming they are not operating in real-time. The guidelines should address the potential for circumventing the prohibition by artificially introducing minor delays or reclassifying systems as non-real-time, even when their operation is effectively like "real-time".

Finally, further clarification is necessary on what constitutes "law enforcement purposes", particularly regarding public security measures. The guidelines should specify whether preventive actions or general public surveillance by law enforcement are covered, in addition to criminal investigations. Additionally, there needs to be guidance on how Member States should ensure transparency regarding these uses, particularly in relation to their citizens, to maintain trust and accountability.

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 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

First, the concept of an "imminent threat" requires clearer definition. It is important to specify the criteria that should apply to determine the immediacy and severity of the threat, and how Member States must ensure that this use is in line with these criteria.

Next, more precision is needed in distinguishing between genuine and present threats, for example in the context of terrorist attacks. The guidelines should clarify whether this distinction applies only to threats already detected or if it also covers proactive intelligence, to prevent ambiguity in applying the exceptions.

Specifically, the process for judicial or independent authorization needs to be clearly outlined, including time frames and specific criteria that must be met. Further clarification on the temporal restrictions for RBI use would ensure that these systems are deployed in a manner consistent with fundamental rights and proportionality.