EUROPEAN
COMMISSION

Brussels, 20.7.2026
C(2026) 5054 final

ANNEX

ANNEX

to the

COMMUNICATION TO THE COMMISSION

Approval of the content of the draft Communication from the Commission - Guidelines on the implementation of the transparency obligations for certain AI systems under Article 50 of Regulation (EU) 2024/1689 (the 'AI Act')

Table of Contents

1. Background and Objectives 3

2. Overview of the Transparency Obligations and Horizontal Topics 3
2.1. Overview of the transparency obligations in Article 50 AI Act 3
2.2. Rationale and objectives 4
2.3. Responsible actors under Article 50 AI Act 5
2.4. Exclusions from the scope of the AI Act 7
2.4.1. Deployers acting in the context of purely personal non-professional activity 8
2.4.2. Research & Development 8
2.4.3. AI released under free and open-source licences 9
2.5. Interplay with the prohibited practices, the requirements for high-risk AI systems and AI literacy 9
2.6. Interplay with rules applicable to general-purpose AI models and/or systems 10

3. Article 50(1) AI Act: Transparency for Interactive AI systems 10
3.1. Main components, concepts and related transparency obligation(s) 11
3.1.1. AI systems intended to interact directly with natural persons 11
3.1.2. Information obligation under Article 50(1) AI Act 13
3.2. Exceptions to the information obligation under Article 50(1) AI Act 16
3.2.1. Exception for obvious interaction with an AI system 16
3.2.2. Exception for AI systems authorised by law for law enforcement purposes 18
3.3. Interplay with other Union legal acts 19

4. Article 50(2) AI Act: Marking and Detection of AI-generated or manipulated content 20
4.1. Main components and concepts of Article 50(2) AI Act 21
4.1.1. AI systems generating or manipulating synthetic content 21
4.1.2. Modalities of synthetic content in scope 21
4.1.3. Content falling outside the scope of Article 50(2) AI Act 22
4.2. The marking and detection obligation of Article 50(2) AI Act 24
4.2.1. The marking obligation 24
4.2.2. The detection obligation 25
4.2.3. Compliance with the requirements for technical solutions: effective, interoperable, robust and reliable 25
4.3. Exceptions to the obligations under Article 50(2) AI Act 28
4.4. Interplay with other Union legal acts 29

5. Article 50(3) AI Act: Emotion Recognition Systems and Biometric Categorisation Systems 30
5.1. Main components, concepts and related transparency obligation(s) under Article 50(3) AI Act 31
5.1.1. The notion of an emotion recognition system 31
5.1.2. The notion of a biometric categorisation system 31
5.1.3. The information obligation under Article 50(3) AI Act 31
5.2. Exception for law enforcement purposes 32
5.3. Interplay with other Union legal acts 32

6. Article 50(4): Labelling of Deep Fakes and Certain Text Publications 33

6.1. Main components, concepts and related transparency obligation(s) for deep fakes under Article 50(4) AI Act 33
6.1.1. The notion of 'deep fake' 33
6.1.2. The disclosure obligation under Article 50(4), first subparagraph, AI Act 37
6.1.3. Transparency of artistic, creative, satirical, fictional or analogous deep fake content 37
6.1.4. Exception for law enforcement 39
6.1.5. Interplay with other Union legal acts 40
6.2. Main components, concepts and related transparency obligation(s) for AI-generated or manipulated text under Article 50(4) AI Act 41
6.2.1. Text published with the purpose of informing the public on matters of public interest 41
6.2.2. The disclosure obligation under Article 50(4), second subparagraph, AI Act 43
6.2.3. Exception from the transparency obligation for text under human review or editorial control and editorial responsibility 43
6.2.4. Exception for law enforcement 45
6.2.5. Interplay with other Union legal acts 45

7. Horizontal Requirements Applicable to the Information Provided under Article 50(5) AI Act 45
7.1. Information provision in a clear and distinguishable manner 46
7.2. First interaction or exposure 46
7.3. Compliance with applicable accessibility requirements 47

8. Enforcement of Article 50 AI Act 47
8.1. Effects of adhering to a code of practice assessed as adequate 47
8.2. Market Surveillance Authorities 49
8.3. Penalties 49
8.4. Entry into application 49

9. Review and Update of the Guidelines 50

1. BACKGROUND AND OBJECTIVES

(1) Regulation (EU) 2024/1689 of the European Parliament and the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending certain regulations ('the AI Act') entered into force on 1 August 2024 ([1]). The AI Act lays down harmonised rules for the placing on the market, putting into service, and use of artificial intelligence ('AI') in the Union ([2]). Its aim is to promote innovation in and the uptake of trustworthy AI, while ensuring a high level of protection of health, safety and fundamental rights in the Union, including democracy and the rule of law.

(2) The AI Act follows a risk-based approach, classifying AI systems into four different risk categories, one of which is AI systems posing transparency risks that are subject to the obligations laid down in Article 50 AI Act. These transparency obligations apply two years after the entry into force of the AI Act, i.e. as from 2 August 2026.

(3) Pursuant to Article 96(1)(d) AI Act, these Commission Guidelines are issued with the aim to serve as practical guidance to assist competent authorities, as well as providers and deployers of AI systems, in ensuring compliance with the transparency obligations under Article 50 AI Act in a consistent, effective, proportionate and uniform manner.

(4) The drafting of these Guidelines was informed by input from a variety of stakeholders collected during a broad consultation organised by the Commission and input from the Member States in the AI Board as well as stakeholder input on the draft guidelines that were published for consultation.

(5) These Guidelines are non-binding. Any authoritative interpretation of the AI Act may ultimately only be given by the Court of Justice of the European Union ('CJEU').

2. OVERVIEW OF THE TRANSPARENCY OBLIGATIONS AND HORIZONTAL TOPICS

2.1. Overview of the transparency obligations in Article 50 AI Act

(6) Article 50 AI Act includes four transparency obligations, each applying to different types of AI systems or their outputs.

Provision	Type of AI system/output	Transparency obligation	Exceptions or special regimes
Art. 50(1)	AI systems directly interacting with natural persons	Providers must develop and design the AI system in such a way that the natural persons concerned are informed they are interacting with an AI system.	Exceptions 1) if the artificial origin of the interaction is obvious, or 2) the system is authorised by law to detect, prevent, investigate or prosecute criminal offences, unless the system is available to the public to report a criminal offence.
Art. 50(2)	AI systems generating or	Providers must ensure the AI system's outputs are	Exceptions if 1) the AI system performs an assistive function for

([1]) Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) (OJ L, 2024/1689, 12.7.2024).

([2]) Article 1 AI Act.

	manipulating synthetic image, video, audio or text content	marked in a machine-readable format and detectable as artificially generated or manipulated with technical solutions that are effective, interoperable, robust and reliable.	standard editing or does not substantially alter the input data or the semantics thereof, or 2) the AI system is authorised by law to detect, prevent, investigate or prosecute criminal offences.
Art. 50(3)	Emotion recognition or biometric categorisation AI systems	Deployers must inform the natural persons exposed to the system of AI system's operation.	Exception if the AI system is permitted by law to detect, prevent or investigate criminal offences.
Art. 50(4)	AI systems generating or manipulating deep fake or text published to inform the public on matters of public interest	Deployers must disclose that the content has been artificially generated or manipulated.	Exceptions if 1) the AI system is authorised by law to detect, prevent, investigate or prosecute criminal offence, or 2) if the text publication has undergone human review or editorial control and is subject to editorial responsibility. Special disclosure regime applies to deep fakes part of artistic, creative, fictional, satirical or analogous works or programmes.

(7) Each of these four obligations is examined in Sections 3 to 6 of these Guidelines. In addition, Article 50(5) AI Act mandates that the information referred to in its paragraphs 1 to 4 shall be provided to the natural persons concerned in a clear and distinguishable manner at the latest at the time of the first interaction or exposure and conform to the applicable accessibility requirements. Section 7 clarifies these requirements.

(8) The various transparency obligations laid down in Article 50 AI Act may apply cumulatively to (the output of) a single AI system, engaging possibly the responsibility of different actors (providers or deployers). For example, if an AI system generates images as part of a direct interaction with a natural person, then the provider of the system may have to comply with both Article 50(1) and (2) AI Act. Where that system is also able to generate images constituting deep fakes or text published on matters of public interest, the deployer may additionally have to comply with Article 50(4) AI Act.

2.2. Rationale and objectives

(9) As explained in Recitals 132-136 AI Act, the purpose of the transparency obligations is to reduce the risks of impersonation, deception, misinformation, manipulation at scale, and fraud and to mitigate potential adverse impacts on democratic processes and societal trust caused by AI-generated or manipulated content or interaction. Informing individuals about the AI-origin of interaction and content will help them take informed decisions, while also contributing to safeguarding trust and the integrity of the information ecosystem. Transparency is not only a technical compliance requirement but a prerequisite for enabling natural persons to recognise AI-mediated interactions and

content or the use of emotion recognition or biometric categorisation systems. Ensuring transparent and trustworthy use of interactive and generative AI systems will also allow to uphold personal autonomy and European values, human dignity and other fundamental rights (e.g. freedom of thought, privacy, equality and non-discrimination, consumer protection, intellectual property etc.), while also preserving the freedom of expression and credibility of media content.

2.3. Responsible actors under Article 50 AI Act

(10) According to Article 3(3) AI Act, **providers** are natural or legal persons, public authorities, agencies or other bodies that develop AI systems, or have them developed, and place them on the Union market or put them into service under their own name or trademark, whether for payment or free of charge. Providers must comply with the relevant transparency obligations in Article 50(1) and (2) AI Act applicable to them, irrespective of whether they are established or located within the Union or in a third country ([3]). Providers established or located outside the Union are also subject to the AI Act if the output of their AI system is used in the Union ([4]). However, incidental, unforeseeable or unauthorised downstream use should not alone trigger the application of the obligations to such third country providers of interactive or generative AI systems that are not placed on the market or put into service in the Union, without prejudice to the powers for competent market surveillance authorities to consider their outputs as non-compliant with the AI Act.

(11) Providers must ensure their interactive or generative AI systems have transparency embedded into the system's design and operation and ensure all relevant transparency obligations in Articles 50(1), (2) and (5) AI Act are met at the latest when those systems are placed on the market or put into service. All natural persons interacting with the AI system or exposed to its outputs in scope must be effectively and consistently informed throughout the lifecycle of the AI system, as placed on the market or put into service by that provider (irrespective of who the deployers or other users of the system are).

For example, a company provides a generative or interactive AI application (e.g. a chatbot, image generator, AI agent) on the Union market under its own name or trademark to consumers, professional users or other legal entities who may directly use the AI system for interaction or content generation (including by integrating the system into their existing processes and using it accordingly without the need for modifications or additional components). The company is a provider responsible for compliance with the transparency obligations in Article 50(1) and/or (2) AI Act, regardless of whether the AI system is provided for free or for payment and regardless of whether the provider is established or located in the Union or in a third country.

Another example of a provider is a company or another organisation (e.g. a public authority) that has developed an interactive AI system (e.g. chatbot) in-house and puts it into service in the Union for its own use and under its name or trademark.

([3]) Articles 2(1)(a), 3(3), (9) and (11) AI Act. For further guidance on the concepts of placing on the market, putting into service and use, see the Commission Guidelines on the prohibited artificial intelligence practices, C(2025) 5052 and Commission Notice – The 'Blue Guide' on the implementation of EU product rules 2022, 2022/C 247/01, Section 2.

([4]) Article 2(1)(c) and recital 22 AI Act.

If a company takes an already existing generative AI system placed on the market by another provider and modifies that system (e.g. with new training data), which it afterwards puts into service under its own name or trade market, then that company becomes a provider of the new system without prejudice to the responsibility of the provider of the initial AI system for the latter.

(12) Deployers are natural or legal persons, public authorities, agencies or other bodies using AI systems under their authority, unless the use is for a personal non-professional activity ([5]). The 'authority' over an AI system should be understood as assuming responsibility over the decision to deploy the system and over the manner of the actual use of the system (including its outputs). It does not necessarily require technical control over the operation of the AI system, so long as the deployer takes the decision for what purposes and how to use the AI system (including in decentralised workflows and group corporate structures). Deployers involved in complex content production and distribution value chains should take proportionate measures to ensure that the labelling of the content they have implemented pursuant to Article 50(4) AI Act is displayed in a clear and distinguishable manner for the targeted and foreseeable audience at the point of first exposure in accordance with Article 50(5) AI Act (e.g., via contractual conditions with distributing partners, user experience (UX) settings and interfaces to be displayed).

(13) Deployers fall within the scope of the transparency obligations in Article 50(3) and (4) AI Act if their place of establishment or location is within the Union, or if they are established or located in a third country where the output of the AI system is used in the Union ([6]). In the latter case, the transparency obligations apply to entities located or established outside the Union where the deployer itself foresees dissemination and use of the AI outputs in the Union (i.e. by directing or authorising distribution within the Union, including by posting deep fakes on the globally accessible internet). However, third country deployers are not bound by the transparency obligations where the content of the AI system reaches audiences in the Union through channels that are unforeseeable and outside their control.

(14) Where the deployer of an AI system is a legal person under whose authority the system is used (e.g. an advertising company), the individual employees that act under the instructions and under the control of that legal person (e.g. digital animators, web designers, content creators, journalists) should not be considered as separate deployers for that system. A legal person remains a deployer even if it involves third parties (e.g. contractors, freelancers) in the operation of the system on its behalf and under its responsibility and control.

For example, a media outlet established in the Union, whose editors use an AI system to support their written coverage of current events with AI-generated text publications on matters of public interests posted on a globally accessible website, is a deployer falling within the scope of the AI Act. Similarly, a company established in a third country that uses an AI system to generate a deep fake of a celebrity featured in an advertisement displayed in the Union is also a deployer falling within the scope of the AI Act.

([5]) Article 3(4) AI Act.
([6]) Article 2(1)(b) and (c) AI Act.

By contrast, a company that merely commissions an advertising agency to produce an advertisement, without taking decisions and exercising control over whether and how the advertising agency uses AI in the production process, is not a deployer.

(15) Operators (e.g. providers and deployers) may fulfil more than one role concurrently in relation to an AI system. For example, a company might develop an in-house generative AI system and use it for the generation of deep fakes. In such a case, it will qualify as both a provider and a deployer that must ensure compliance with the machine-readable marking obligations under Article 50(2) and the requirements for a clear and perceivable labelling of the deep fakes under Article 50(4) AI Act.

(16) By contrast, other actors (such as providers of hosting services, including online platforms, or broadcasters) whose role is limited to disseminating or transmitting AI-generated or manipulated content created by third parties, or who receive or are exposed to AI-generated or manipulated content without having authority over the use of the AI system for the purpose of the AI content generation or manipulation, are not deployers within the meaning of the AI Act. Without being responsible and liable under the AI Act, those actors may nevertheless play an important role in the value chain to preserve the trust and integrity of the information ecosystem. They are, therefore, strongly encouraged to preserve the marking and labelling implemented pursuant to Article 50 AI Act and implement proportionate and technically feasible measures to enable natural persons to detect the AI-generated or manipulated content, without prejudice to the responsibility of providers and deployers of generative AI systems or any obligations or liability under Regulation (EU) 2022/2065 (the Digital Services Act or 'DSA') ([7]) or other applicable Union or national law.

(17) Other actors acting in their professional capacity and disseminating content (even if not directly engaged under Article 50 AI Act) are also encouraged to take proportionate and appropriate measures, so that the natural persons ultimately exposed to the content are effectively informed and able to detect its artificially generated or manipulated origin. In addition, these parties may be required to label the respective deep fakes or text publications on matters of public interest under other applicable legal or deontological rules (e.g. a broadcaster labelling and showing as part of a TV-programme a deep fake created by a private person). However, the mere fact that an actor labels content or uses labelled content does not mean that the latter qualifies as a 'deployer' within the meaning of the AI Act. The AI Act does not extend or affect the legal responsibility of those actors under the applicable legal frameworks, if they do not qualify at the same time as providers or deployers of AI systems.

2.4. Exclusions from the scope of the AI Act

(18) Article 2 AI Act provides a number of general exclusions from the scope of the AI Act which have already been clarified in the Commission Guidelines on prohibited artificial intelligence practices ([8]). This section will accordingly address only those aspects that are relevant to the transparency obligations in Article 50 AI Act.

([7]) Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services; OJ L 277, 27.10.2022, pp. 1.

([8]) See Section 2.5 of Commission Guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689 (AI Act), Brussels, 29.7.2025, C(2025) 5052 final.

2.4.1. Deployers acting in the context of purely personal non-professional activity

(19) According to Article 2(10) AI Act, the regulation "does not apply to obligations of deployers who are natural persons using AI systems in the course of a purely personal, non-professional activity" ([9]). Any activity through which natural persons gain an economic benefit on a regular basis or are otherwise involved in a professional, business, trade, occupational or freelance activity should be considered as a 'professional' activity. The specification of 'purely personal' is a qualifier of non-professional, meaning that the person should act in both a personal and a non-professional capacity to fall within the scope of the exception. Any use by natural persons acting on behalf or under the authority of a deployer acting in a professional capacity will fall within the scope of the professional deployer's transparency obligations laid down in Article 50(3) and (4) AI Act.

Examples of personal non-professional use of AI systems excluded from the scope:

- individuals using an AI system to generate deep fakes of the members of their household;

- individuals using an AI system in a personal capacity to generate text on various topics (e.g. students using AI to write their homework);

- individuals using an AI system in their personal capacity to generate deep fakes disseminated on social media;

(20) The exclusion in Article 2(10) AI Act applies only as regards the obligations of deployers when using the system for purely personal, non-professional activities. The system as such remains within the scope of the AI Act as regards the obligations of providers placing the system on the market or putting it into service and of deployers using the system for professional use. The exception is also without prejudice to the application of other relevant Union and national law.

For example, the provider of the AI system used to generate deep fakes to feature on a personal Christmas card should mark those deep fakes in a machine-readable manner in accordance with the transparency obligation laid down in Article 50(2) AI Act. At the same time, the use of the system by natural persons for purely personal, non-professional activities will not be covered by the deployer obligation laid down Article 50(4) AI Act.

2.4.2. Research & Development

(21) Article 2(6) AI Act provides an exclusion for AI systems or AI models, including their outputs, specifically developed and put into service for the sole purpose of scientific research and development. That exclusion is without prejudice to applicable research ethical standards and EU data protection and other legislation. In the context of Article 50 AI Act, this exclusion covers not only interactive AI systems, but also the outputs of generative AI systems (including deep fake content) used in the context of scientific research. However, if those systems are put into service or their outputs are being used for other purposes than for the sole purpose of scientific research and development, the relevant transparency obligations in Article 50 AI Act would still need to be complied with.

([9]) The definition of deployer also excludes users engaged in such activities (see Section 2.3 above).

For example, if researchers develop and place on the market or put into service an interactive AI system and wish to investigate whether natural persons are able to distinguish between AI systems and humans when engaging in spoken interaction with that AI system, they do not need to implement the respective transparency obligations under Article 50(1) and (2) AI Act. However, if the same interactive AI system is also put into service for purposes other than research and development, the relevant transparency obligations in Article 50 AI Act would need to be complied with.

(22) Article 2(8) AI Act furthermore excludes research, testing or development activities regarding AI systems or AI models prior to their placement on the market or putting into service. Testing in real world conditions (inside or outside of AI regulatory sandboxes) is not covered by that exclusion. Furthermore, once an AI system is placed on the market or put into service as a result of such research and testing activities, it should nonetheless comply with the relevant transparency obligations under Article 50 AI Act ([10]).

2.4.3. AI released under free and open-source licences

(23) According to Article 2(12) AI Act, AI systems released under free and open-source licences are outside the scope of the AI Act provided they do not fall under the prohibitions of Article 5 AI Act or the transparency obligations of Article 50 AI Act, or they are not classified as high-risk ([11]). This means that providers and deployers of open-source AI systems within the scope of Article 50 AI Act still need to ensure compliance with their respective transparency obligations.

(24) By contrast, providers of free and open-source AI components that cover software, data, models, tools, services or processes for customisation and integration into an AI system and that do not constitute in themselves (in part or in whole) an AI system, are not subject to the transparency obligations in Article 50 AI Act. Providers of such systems are nevertheless encouraged to design those models, components, and services in a manner described in paragraph 27 below, insofar as this is proportionate and feasible.

2.5. Interplay with the prohibited practices, the requirements for high-risk AI systems and AI literacy

(25) Recital 137 AI Act clarifies that compliance with the transparency obligations under Article 50 AI Act for a particular AI system cannot be interpreted as indicating that the use of that AI system or its output is lawful under the AI Act. In particular, AI systems subject to the obligations in Article 50 AI Act may fall under Article 5 AI Act and thus be prohibited (e.g., such as the use of emotion recognition systems in the area of workplace or education institutions) ([12]). Furthermore, AI systems falling within the scope of Article 50 AI Act may also be classified as high-risk ([13]) and be subject to the relevant requirements and obligations for such systems. Compliance with Article 50

([10]) Recital 25 AI Act.
([11]) Recital 102 AI Act explains that a release of software and data under free and open-source licence "allows them to be openly shared and where users can freely access, use, modify and redistribute them or modified versions thereto".
([12]) For further information on Article 5 AI Act, see Commission Guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689 (AI Act), C(2025) 5052 final.
([13]) See Article 6 AI Act.

AI Act therefore does not replace, limit or affect obligations arising under other provisions of the AI Act ([14]). At the same time, AI systems can also fall within the scope of Article 50 AI Act without being classified as high-risk pursuant to Article 6 AI Act. Literacy requirements under Article 4 AI Act also apply to both providers and deployers of AI systems within the scope of Article 50 AI Act.

2.6. Interplay with rules applicable to general-purpose AI models and/or systems

(26) Article 50 AI Act applies to certain AI systems, which includes general-purpose AI systems ('GPAI systems') ([15]), such as chatbots supporting direct interaction with natural persons or AI systems capable of synthetic content generation or manipulation (including deep fakes) ([16]).

(27) Article 50 AI Act does not explicitly apply to GPAI models. However, if an interactive or generative AI system that falls within scope of Article 50(1) or (2) AI Act is built upon a GPAI model provided by the same provider, the transparency measures set out in these Guidelines could also be implemented at the model level (see also Section 4 below). Furthermore, while they fall outside the scope of Article 50 AI Act, providers of other GPAI models and of other generative AI models are also encouraged to implement appropriate transparency measures at the model level for the identification of interactive AI systems and the marking of AI-generated and manipulated content. This is particularly beneficial to facilitate compliance by downstream AI system providers with their obligations under Article 50(1) and (2) AI Act ([17]) and to contribute to preserving the trust and integrity of the information ecosystem. For GPAI models presenting systemic risk, model-level safety measures for the identification of AI interactions and marking of AI-generated or manipulated content may, where appropriate, constitute part of a broader set of mitigating measures that providers may take as part of their obligation to mitigate systemic risks pursuant to Article 55(1)(b) AI Act.

3. ARTICLE 50(1) AI ACT: TRANSPARENCY FOR INTERACTIVE AI SYSTEMS

(28) Article 50(1) AI Act is addressed to providers of AI systems directly interacting with natural persons, who must design and develop their systems in such a way that the natural persons concerned are informed that they are interacting with an AI system. As explained in Recital 132 AI Act, the purpose of this information obligation is to enable those natural persons to take informed decisions regarding the system's outputs, to avoid overreliance on such systems, and to allow those natural persons to calibrate their trust in the content and the interactions accordingly.

([14]) In particular, Article 26(11) AI Act imposes a transparency obligation on deployers of high-risk AI systems that make decisions or assist in making decisions related to inform natural persons that they are subject to the use of the high-risk AI system. These transparency obligations are complementary to those in Article 50 AI Act, since they concern different information that must be provided to the natural persons, but they can be combined in single notifications.

([15]) Article 3(66) AI Act defines a general-purpose AI ('GPAI') system as an AI system which is based on a general-purpose AI model and which has the capability to serve a variety of purposes, both for direct use as well as for integration in other AI systems.

([16]) Article 50(2) AI Act contains an explicit reference to GPAI systems.

([17]) Recital 133 AI Act.

3.1. Main components, concepts and related transparency obligation(s)

(29) The transparency obligation in Article 50(1) AI Act applies to providers of AI systems where the system is intended to interact directly with natural persons and that interaction does not fall under any of the following exceptions: (1) it is obvious from the point of view of a natural person who is reasonably well-informed, observant and circumspect, or (2) the system is authorised by law to detect, prevent, investigate or prosecute criminal offences, unless the system is available for the public to report a criminal offence.

3.1.1. AI systems intended to interact directly with natural persons

(30) Four elements must be present for a system to fall within the scope of Article 50(1) AI Act: (i) the system must be an AI system, (ii) intended to interact, (iii) directly, (iv) with natural persons.

i. **An AI system**: The system must fulfil the elements of the AI system definition ([18]), which excludes simple non-AI automated response mechanisms (e.g., traditional out-of-office emails, rule-based quick message answers).

ii. **Intended to interact:** The AI system must be intended to interact with natural persons. Interaction entails the capability of a bidirectional exchange of information or actions between natural persons and the AI system with a genuine conversational or responsive character. It can occur in various forms (e.g. auditory, visual and physical). Natural persons should be able to provide input (e.g., written text or other forms of content, voice, or physical actions) to the AI system (e.g. via technical communication means including online interfaces, phone or e-mail, or directly), while the AI system should be capable of responding with contextual output (including any type of content or actions) and vice-versa. The interaction may cover any type of output or format that can be perceived and understood by humans. It does not have to be initiated by a human; it may also cover AI content or actions eliciting a human response. The interaction may be a one-time exchange (e.g. single prompt followed with a single reply) or take place over a certain period time (multi-turn). Systems that only passively collect data and are not capable of engaging in an exchange with natural persons (e.g. automated facial-recognition access controls) or just collect one-time feedback (e.g. spam filters), are not considered to interact with natural persons.

iii. **Direct interaction:** The interaction between the AI system and the natural persons must be direct. This typically involves real-time or near real-time interaction (including through capabilities of the system to write and send messages to natural persons or otherwise interact with the physical or virtual environment that can be perceived by natural persons). Direct interaction excludes indirect or mediated human interaction where a natural person is exposed to AI outputs without directly interacting with the system (for example, if customer service representatives use AI assistance tools to help them communicate with natural persons or if the AI output is not made available to the person by the AI system itself, but by another person disseminating the content). However, the mere possibility for humans to intervene or review the AI system's outputs should not be used to circumvent the application

([18]) See Article 3(1) AI Act. For further guidance on the concept of AI system, see Commission Guidelines on the definition of an AI system, 29.7.2025, C(2025) 5053 final.

of the disclosure requirement under Article 50(1) AI Act. Cases of products and services blending AI-generated responses with human-curated content fall within the scope of Article 50(1) AI Act and require disclosure for the specific outputs in the interaction that are AI-generated, unless those AI outputs have been properly reviewed and sent by humans as the main interlocutors with the natural persons.

iv. **With natural persons:** Article 50(1) AI Act requires that the AI system interacts with natural persons. Those persons may be professional deployers, other users (also using the system for purely personal activities) or other persons using the system on their behalf. AI systems that operate in a closed physical environment (such as part of an industrial machinery set-up), backend machine-to-machine calls between AI systems (including AI agents) whose outputs are not intended to reach natural person, or virtual environments without any direct contact with natural persons are excluded from the scope of the obligation.

(31) AI agents are covered by Article 50(1) AI Act if they are capable of interacting with the persons instructing them or with other natural persons in the execution of the tasks (e.g. making bookings, managing correspondence, negotiating or concluding contracts, executing purchases, etc). AI agents must be designed and developed in such a way that they disclose both their artificial nature and the person on whose behalf they are acting, considering the need for transparency of the origin and the delegation of authority and accountability for the consequences of their actions. This includes cases where AI agents are capable of interacting in complex multi-agent architectures with other AI agents that directly interact with the natural persons in the execution of the task. Nevertheless, it may not always be possible for the provider of the AI agent to identify in advance individual instances of direct interaction of an AI agent with other natural persons once the AI agent is deployed in different contexts. Where the provider cannot reliably determine whether the AI agent will directly interact with a natural person before placing the AI agent on the market or putting it into service, the agent should be designed at the architecture level, and instructed, to disclose itself as such in every situation where it is reasonably likely that the agent may interact with a natural person (including where a natural person may represent or act for a legal entity). Furthermore, AI agents should also disclose themselves, to the persons instructing them at key steps (e.g. at the point of authorisation, reporting, validation etc., including when the deployed AI agent receives, processes, or relies upon outputs generated by other AI systems rather than directly by a natural person) and at every new interaction.

Examples of AI systems intended to interact directly with natural persons falling within scope of Article 50(1) AI Act:

- AI-enabled voice assistants, chatbots/conversational agents in various contexts (e.g. public service, customer support, complaints management, e-commerce, finance, healthcare, education etc.),

- AI hotline for incidents, fraud or other reporting, (humanoid) robots/cobots, AI companions; robotic companion pets; AI avatars (e.g. in virtual reality environments),

- AI bots on social networks and media, coding agents and other AI agents capable of directly interacting with natural persons (including in complex multi-agent architectures).

3.2.1Examples of AI systems not directly interacting with natural persons falling outside the scope of Article 50(1) AI Act:

- AI-enabled traditional industrial robots that operate in a closed industrial setting and not intended to interact with humans (e.g. electronic assembly robots),

- AI-enabled recommender systems, spam filters, automated translation or transcription tools,

- AI-enabled search and retrieval of information without generating or modifying content,

- AI-enabled text and code auto-completion techniques, authentication/biometric recognition systems, backend decision-support systems where the user only sees the AI output without possibility for direct interaction with the system;

- AI-enabled predictive maintenance and optimisation systems in factories and industrial applications, AI-assisted product visualisation (e.g., virtual try-on, room staging, product rendering), AI-enabled single-turn route-planning tools, transcription of customer calls processed for behavioural analysis etc.

3.1.2. Information obligation under Article 50(1) AI Act

(32) Article 50(1) AI Act obliges providers of AI systems directly interacting with natural persons to ensure that their AI systems (including AI agents and GPAI systems) are designed and developed in such a way that the natural persons concerned are informed that they are interacting with an AI system.

(33) While the mechanism to inform should be embedded in the design and development of the system and in the way the system interacts with the natural persons, notification should be provided during the operation of the AI system and at the latest at the time of the first interaction with the natural person (as required in Article 50(5) AI Act, see Section 7 below). Providers are responsible for ensuring that, after the AI system's placement on the market or putting into service, natural persons interacting with the AI system are effectively and consistently informed throughout the lifecycle of the AI system, as placed on the market or put into service by that provider.2.2

(34) **Format of disclosure:** Article 50(1) AI Act does not specify any particular disclosure technique. Providers can therefore choose any appropriate technique, so long as two requirements are met: (i) the requirements set out in Article 50(5) AI Act are fulfilled regarding the provision of clear and distinguishable information at the latest at the time of the first interaction (see Section 7 below); and (ii) the characteristics of natural persons belonging to vulnerable groups due to their age (e.g. children or the elderly) or disability are taken into account to the extent that the AI system is intended or reasonably likely to interact with natural persons belonging to those groups (e.g. including when the AI system is reasonably likely to interact with the general public) ([19]). The information must also be provided in an accessible format for natural persons with disabilities, as specified in Article 50(5) and Recital 132 AI Act (see Section 7). Where the system may interact with children, those notifications must be child-friendly, age-appropriate, easy-to-understand and easily accessible to all children, including those with disabilities or additional accessibility needs. Where available and appropriate, providers are

([19]) See Recital 132 AI Act. This does not require that the system is targeted at those vulnerable groups. If a system is intended to interact with any member of the public, then members of the vulnerable groups are included as well and should be considered by providers when implementing the notification measures.

recommended to use standardised disclosure practices and labels to ensure effectiveness and consistency of the notifications and improved user awareness.

(35) **Substance of the notification:** Article 50(1) AI Act requires providers of interactive AI systems to design and develop them in a manner that all natural persons interacting with such AI systems are explicitly informed about the artificial, non-human nature of the interacting counterpart. If it is reasonably foreseeable to the provider that the AI system may interact with natural persons belonging to the above-mentioned vulnerable groups, then the information should be adapted accordingly and provided in an accessible format, as required under applicable Union law to also ensure the full understanding of those natural persons ([20]).

(36) The notifications should provide information on the artificial origin of the interaction in ways that are appropriate to the context in which the interaction takes place. This should be done in a manner that ensures that the persons concerned are properly and effectively informed, avoiding any risks of deception or manipulation before and throughout the interaction, while ensuring feasible and proportionate disclosures. Examples of providing the information include notifications in writing (e.g., a chatbot that starts a conversation by mentioning that it is based on AI technology), visual means (e.g. an email generated by an AI agent sent to a natural person that features an AI label at the top), auditory means (e.g. a voice assistant that says at the beginning of a session that it is powered by AI), disclosure of AI identifiers, and credentials (e.g. AI agents that disclose their AI identity to the extent feasible in a verifiable manner ([21]).

(37) **Format of the notifications**: The design of the AI system with multimodal and accessible disclosure approaches is recommended as a best practice. Such notifications may combine different techniques to reinforce user understanding, where appropriate and depending on the context and the target and foreseeable audience groups and the available resources of the providers, in particular Small and Medium Sized Enterprises (SMEs). Examples of appropriate techniques may include:

- **Textual disclosure** (User Interface (UI)-based): Prominent, plain-language labels or banners (e.g. “You are interacting with an AI system”) and first-turn greetings in chatbots that may be combined, as appropriate, with persistent badges or labels visible throughout the interaction as proportionate to the specific context and risks. Furthermore, it is recommended to position accompanying disclosures close to the interaction interface (e.g. near the input/output field) and use simplified wording, particularly for users with lower digital literacy and children.
- **Auditory disclosure:** In voice-based or telephony contexts, explicit spoken statements at the beginning of the interaction (e.g. “This is an AI-powered assistant”) combined, as appropriate, with periodic reminders in longer interactions, in particular in case of interruptions or a change of the role of the AI system during a user journey. Distinct audio cues (e.g. tones or earcons) may support recognition, particularly for visually impaired users, but are not considered sufficient by themselves.

([20]) See Section 7.3 for applicable accessibility requirements.

([21]) For example, electronic attestations of attributes as established under Regulation (EU) No 910/2014 and as made available in the EU Digital Identity Wallets established under the same Regulation, as well as the proposed European Business Wallets, can provide secure and efficient means of identifying AI agents. The Wallets can store and manage electronic attestations that verify the AI agent's identity, attributes, and authorisations, thereby enabling seamless and trustworthy disclosure to natural persons.

- **Visual/graphical cues:** Design of the AI system with persistent icons, watermarks, coloured frames or recognisable "AI" symbols to complement textual disclosures. The use of standardised visual indicators, including across services of a provider, is recommended, since such indicators can significantly reduce cognitive burden and increase recognisability.
- **Multi-modal combinations:** Combining text, audio and visual cues (e.g. a chatbot displaying a label, while also providing a first-turn textual disclosure) to ensure accessibility and reinforce clarity across different user groups.

(38) Certain techniques are not necessary and, when used alone, insufficient for effectively fulfilling the transparency obligation in Article 50(1) and (5) AI Act, notably:

- Disclosures contained only in terms and conditions, URLs, or documentation (such disclosures may complement, though not replace, in-context disclosure);
- Machine-readable markings (e.g. metadata or watermarks) that are not perceivable by users at the point of interaction. This does not affect the appropriateness of such marks to comply with Article 50(2) AI Act (see Section 4);
- Unclear or ambiguous signals (e.g. generic references to "assistant") or human-like representations that may mislead users;
- Generalised disclosures that are not sufficiently specific to the AI system's outputs and interactions (e.g. on a platform providing a variety of services or individual inputs, a general disclosure like "Services on this website use AI" is insufficient);
- Technical or capability-based descriptions: statements solely referring to underlying technologies (e.g. "this system uses LLMs") without explaining the function or implications of the system for the user and its artificial origin.

(39) Disclosures should be proportionate and calibrated to the different contexts and target audiences of the AI system to remain effective, user-friendly and avoid, to the extent, possible disruptions of the user experience or habituation effects (so-called "banner blindness").

(40) A single, prominent notification before the first interaction of the AI system with a particular natural person is likely to suffice in most instances to meet the obligation in Article 50(1) AI Act. However, in certain riskier contexts, periodic reminders and context-aware disclosures are likely to be necessary to ensure continued user awareness. This is in particular the case in the context of AI systems interacting with vulnerable persons (e.g. children, the elderly, persons with disabilities and those with lower literacy), sustained and evolving interactions in sensitive or immersive contexts (e.g. where users may express or experience emotional distress or vulnerability or be at risk of addiction-like behaviours), where there is an increased risk of users being misled (e.g. financial advice, insurance, legal assistance, health advice, and complaints handling), or of forming emotional attachments or dependencies (e.g. AI companions), as well as when there is a need to ensure awareness and control over the actions of AI agents. Providers must also design and develop the AI system to ensure disclosure in all situations where the AI system is being asked questions relating to its nature or to the origin of the interaction, or where it can be reasonably assumed from the exchanges with the natural person that that person is likely to be misled or confused about the AI origin of the interaction.

3.2. Exceptions to the information obligation under Article 50(1) AI Act

(41) Article 50(1) AI Act provides two exceptions to the information obligation regarding interactive AI systems examined below.

3.2.1. Exception for obvious interaction with an AI system

(42) To be able to rely on the first exception, providers will need to assess and demonstrate: i) the obvious artificial nature of the interaction to ii) a natural person who must be reasonably well-informed, observant and circumspect, taking into account the circumstances and the context of use.

(43) This standard draws on established EU consumer protection law regarding the notion of 'average consumer' and should be interpreted consistently, while taking into account the specific and varied contexts and uses of AI systems directly interacting with natural persons and the objectives pursued by Article 50(1) AI Act.

(44) Conceiving such 'an average' natural person comprises two steps. First, providers should consider the audience with whom the AI system is intended and reasonably likely to interact. Second, the provider should examine how reasonably well-informed, observant and circumspect an average member of that audience is, taking into account the circumstances and the context of use. Several factors should be considered when performing this assessment:

- Providers should consider the target audience of the AI system, as well as a broader potential audience that is reasonably foreseeable. This will especially be the case if the general public can easily access the AI system or if the system is operating in virtual or physical public spaces and capable of interacting with natural persons present in those spaces. The audience that may be exposed to an AI system does not always equals the target audience.
- Providers should consider the potential diverse composition of the target and reasonably foreseeable audience in line with Recital 132 AI Act. Where persons with disabilities, elderly people or minors are likely to be part of that audience, the expected levels of information, observance and circumspection of an average member of such an audience will be lower as compared to an average member of an audience that does not include those categories of persons. By contrast, if the interactive AI system is only available to a professional or specialised target audience, then this could positively impact the expected levels of knowledge, observance and circumspection and an average member of that audience may be considered to be better informed and more observant and circumspect.
- As part of the broader circumstances and context of use, providers should consider the level of digital and AI literacy of the target and reasonably foreseeable audience and relevant social, cultural, and linguistic factors.

(45) **Obviousness**: Upon establishing the hypothetical reasonably well-informed, observant and circumspect natural person, providers should evaluate whether it will be obvious to such a person that they are directly interacting with an AI system. The exception should be interpreted restrictively given that it deprives natural persons from the protection and the right to be informed with clear and distinguishable disclosures that they are interacting with AI systems. The general awareness of consumers and other natural

persons that AI systems (including chatbots and agents) exist does not imply that they recognise them in interactions. Several factors may be relevant and considered as part of the assessment whether the interaction is obvious:

- The anticipated nature of the interaction: In instances where the AI system interacts physically with natural persons, pertinent elements may include the visibility of mechanical components (which would increase obviousness) or the degree to which the AI system authentically replicates its non-artificial equivalent (which would decrease obviousness). For interactions that are visual or auditory in nature, elements such as writing or speech patterns, vocal tone (robot voice vs. genuine human-sounding voice), user interface design (e.g. a profile picture related to a chatbot that displays a human), and the capability for advanced personalised interaction may impact this assessment.
- The composition of the target and foreseeable audience: If the interactive AI system is only available to a professional or specialised audience, without potential exposure to vulnerable groups, this could increase the possible obviousness of the artificial nature of the interaction to an average member of the respective audience. By contrast, if the AI system may be accessed by the general public or consumers that include also vulnerable persons (e.g. such as children, the elderly, persons with cognitive, physical or mental disabilities, or persons with lower AI literacy) for whom the interaction is less obvious, the exception cannot be relied upon in order to protect those vulnerable users efficiently.
- Consideration should also be given to the fact that it is becoming increasingly hard for natural persons to know whether an interaction with an AI system or an actual human being is taking place. Due to the use of avatars and agents, the general anonymity of interactions online, and the capabilities of AI systems to write and interact in a human-like manner, the "obviousness" exception should be limited to cases where there is almost no doubt left about the nature of the interaction for an average person from the targeted and reasonably foreseeable audience.

Examples of obviousness for which the transparency obligation under Article 50(1) AI Act does not apply:

- AI-powered code assistance and code review chatbots available only to professional developers who by virtue of their expertise and the context of use have no reasonable expectation of interacting with a human and can easily recognise that suggestions are generated by an AI system.

- An internal employee-facing assistant for properly trained staff who is AI literate and aware that they are using AI systems for internal organisational purposes (e.g. HR, legal, procurement, compliance or IT support).

- Interactive AI systems intended only to be used by properly trained health professionals to support medical diagnosis and suggest related treatments.

- Ambient-AI embedded in home appliances with functionalities constrained to the home appliance assistance and operation.

- Interactions with AI-enabled Non-Playable Characters (NPCs) in a single-player video game where the nature of the video game makes it clear that no other natural person can participate

and interact with the user and the AI nature of the interaction is also obvious for all users (including vulnerable persons).

Examples that do not fulfil the obviousness-exception and for which the transparency obligation under Article 50(1) AI Act applies:

- An AI-powered robotic companion pet, looking highly similar to its natural equivalent and designed to mimic typical human-pet interaction, creating ambiguity as to whether the interaction is with an AI system or not.

- AI systems embedded in immersive environments (e.g. virtual or augmented reality) using realistic avatars or voices resembling humans, where users and particularly children, the elderly, or persons with disabilities may not be able to distinguish easily between human and AI interaction.

- AI chatbots embedded in online platforms or assistance support tools (helpdesks) whereby users directly interact and receive AI outputs (e.g. replies to queries or other AI-generated content) they may perceive as human-generated.

3.2.2. Exception for AI systems authorised by law for law enforcement purposes

(46) Providers of interactive AI systems are exempted from the transparency obligation under Article 50(1) AI Act if they are authorised by law to detect, prevent, investigate, or prosecute criminal offences, subject to appropriate safeguards for the rights and freedoms of third parties.

(47) 'Authorised by law' is to be understood to comprise Union law, as well as Member States' national law adopted in compliance with Union law. The law authorising the use of the AI system must clearly specify the purposes and circumstances of permitted use and provide for appropriate safeguards to protect the rights and freedoms of third parties. Considering the technology neutral approach of the AI Act, it is not necessary to identify or authorise specific AI systems by law where the use of the AI system is generally envisaged within the scope of the legally authorised activity and complies with the applicable legal framework.

(48) To fall within this exception, the purpose of the AI system must be to detect, prevent, investigate, or prosecute criminal offences (e.g. AI-undercover agent, AI systems used under human control of investigators for online investigations or communication with persons who may be involved in criminal activities). The exception is not restricted to the use of such AI systems only by law enforcement authorities as defined in Article 3(48) AI Act, but may also cover interactive (or generative) AI systems used by other Union or national public authorities, so long as their use is authorised by law to detect, prevent, investigate, or prosecute criminal offences, and subject to appropriate safeguards to protect the rights and freedoms of third parties. The exemption does not require the AI system to be capable of conducting the investigations independently. If a respective system can both be used for law enforcement purposes, as well as for other regular purposes, the applicability of the exemption should be assessed in light of the particular deployment. Transparency obligations continue to apply where the same AI system is deployed for other purposes than the legally authorised detection, prevention, investigation or prosecution of criminal offences. Transparency still needs to be provided

to natural persons when they are interacting with the system in the context of its regular purposes.

(49) The exception does not apply if (i) the AI system is available to the public, and (ii) the system offers a functionality to individuals to report criminal offences. Such systems may be used, but they remain subject to the transparency obligation under Article 50(1) AI Act, so natural persons must be informed they are interacting with such systems.

For greater clarity, this exception concerns solely the obligation under Article 50(1) to inform natural persons that they are interacting with an AI system ([22]). It is without prejudice to applicable Union and national law, including procedural rules concerning the documentation, disclosure, admissibility and assessment of evidence.

Examples of AI systems made available to the public to report criminal offences that are not exempted and fall within the scope of Article 50(1) AI Act:

- AI-enabled police chatbots deployed on official police websites or mobile applications that allow individuals to report criminal offences or complaints. These systems guide users through structured questions, collect relevant information, and triage reports before forwarding them to human officers for assessment.

· AI-powered telephone hotlines operated by the police or other law enforcement authority for general public inquiries.

- AI-assisted fraud reporting hotlines or digital portals operated by financial institutions or public authorities, where users can report suspected financial crimes. The AI system interacts with the user to gather details, categorise the report, and prioritise cases, while human investigators validate the information before any action is taken.

- Virtual assistants used for witness statement collection, where victims or witnesses can submit information through an AI-driven interface. Such AI systems may support multilingual input, accessibility features, and structured evidence submission, but the final assessment and investigative steps remain with human authorities.

3.3. Interplay with other Union legal acts

(50) The transparency obligation under Article 50(1) AI Act applies without prejudice to existing Union consumer protection, data protection, digital or non-discrimination legislation that may require other information disclosures not specific to the AI origin of the interaction. In particular, under Directive 2005/29/EC (the Unfair Commercial Practices Directive or 'UCPD') ([23]), misleading actions or omissions with regards to main characteristics of a product or service are prohibited. Furthermore, Directive 2011/83/EC (the 'Consumer Rights Directive') ([24]) requires traders to provide clear and comprehensible pre-contractual information about the main characteristics of the goods or services. Where a service is AI-driven (such as a subscription-based chatbot, an AI

([22]) This also applies for the similar exceptions to the benefit of law enforcement activities provided for in Article 50 (2)- (4) AI Act.

([23]) Directive 2005/29/EC of the European Parliament and of the Council of 11 May 2005 concerning unfair business-to-consumer commercial practices in the internal market; OJ L 149, 11.6.2005, pp. 22.

([24]) Directive 2011/83/EU. of the European Parliament and of the Council of 25 October 2011 on consumer rights; OJ L 304, 22.11.2011, pp. 64.

companion application, or a virtual coaching service), the AI functionality may qualify as an essential characteristic that must be disclosed before the consumer is bound by a contract. The information obligations under Union consumer protection law apply irrespective of whether the interaction is considered "obvious" under Article 50(1) AI Act. Those obligations are also complementary as general consumer protection information obligations, while the notification obligation under the AI Act is more contextual and aims to ensure ongoing awareness of all natural persons about a direct interaction with an AI system at the least before its start and throughout the interaction.

(51) The DSA also includes transparency requirements for online platforms' recommender systems, including a requirement to explain the main ranking parameters. These obligations are complementary to the transparency obligations laid down in Article 50(1) AI Act, since they have a different scope and purpose and concern systems that, in principle, do not interact directly with users within the meaning of Article 50(1) AI Act.

(52) As regards the information obligations towards data subjects under Union data protection law, the obligation in Article 50(1) AI Act fulfils a different objective than those obligations and, as such, does not affect those information obligations.

(53) Similarly, the transparency obligation under Article 50(1) AI Act applies without prejudice to Union and national law concerning the provision of information to persons in the context of criminal proceedings (e.g. Directive 2012/13/EU or Directive 2012/29/EU, as amended by Directive (EU) 2026/1472) ([25]).

4. ARTICLE 50(2) AI ACT: MARKING AND DETECTION OF AI-GENERATED OR MANIPULATED CONTENT

(54) Article 50(2) AI Act requires providers of AI systems generating synthetic content to implement technical solutions that meet certain quality requirements for machine-readable marking and detection of their AI systems' outputs. The objective is to enable natural persons to distinguish AI-generated or manipulated content from other content (for example, content created by humans) and to verify its origin, thus also contributing to increased integrity and trust in the information ecosystem.

(55) The beneficiaries of this transparency obligation are therefore all natural persons likely to be exposed to the AI-generated or manipulated content, as well as key actors who play an important role for the trust and integrity of the information ecosystem, such as competent market surveillance authorities and other competent authorities, independent researchers, civil society organisations, media, trusted flaggers, fundamental rights defenders, providers of online platforms and online search engines, rightsholders and other content creators and distributors etc.

([25]) Directive 2012/13/EU of the European Parliament and of the Council of 22 May 2012 on the right to information in criminal proceedings; OJ L 142, 1.6.2012, pp. 1-10; Directive 2012/29/EU of the European Parliament and of the Council of 25 October 2012 establishing minimum standards on the rights, support and protection of victims of crime; OJ L 315, 14.11.2012, pp. 57.

4.1. Main components and concepts of Article 50(2) AI Act

(56) Article 50(2) AI Act applies if several cumulative conditions are fulfilled:

i. The system must qualify as an AI system;

ii. The AI system must be capable of generating or manipulating synthetic content;

iii. The content must be in one or more of the following modalities: audio, image, video or text;

iv. The AI system must not fall within any of the following exceptions: (1) the AI system performs an assistive function for standard editing; (2) the AI system does not substantially alter the input data provided by the deployer or the semantics thereof; or (3) the AI system is authorised by law to detect, prevent, investigate or prosecute criminal offences.

4.1.1. AI systems generating or manipulating synthetic content

(57) Article 50(2) AI Act applies to AI systems ([26]), including GPAI systems, generating synthetic audio, image, video or text content. Article 50(2) AI Act covers a subset of systems that fulfil the AI system definition in Article 3(1) AI Act and that generate such outputs. Article 50(2) AI Act also covers AI systems used to manipulate content.

(58) Content generation refers to the generation of synthetic material by the AI system (e.g. based on a human prompt), such as generating a synthetic image or a song. In practice, this means that Article 50(2) AI Act may apply to generative AI systems designed with a narrow intended purpose to produce specific outputs. At the same time, Article 50(2) AI Act also applies to AI systems that may serve multiple purposes, including GPAI systems, and can produce a variety of types of content, as well as agentic AI systems, so long as they generate synthetic audio, image, video or text content.

(59) By contrast, content manipulation refers to already existing content (whether synthetic or not) that is altered by an AI system (e.g. an existing image or a voice recording that is manipulated by an AI system in accordance with human instructions). For content to fall under Article 50(2) AI Act, it is sufficient that it is AI-generated or manipulated beyond standard editing under the exceptions set out in Article 50(2) AI Act (see Section 4.3). Additionally, Article 50(2) AI Act does not require that the content is solely AI-generated or manipulated. Content that is mixed with human-created material also qualifies as synthetic content, if manipulated or generated in one of the modalities referred to in Article 50(2) AI Act.

4.1.2. Modalities of synthetic content in scope

(60) Article 50(2) AI Act provides an exhaustive list of the modalities of the synthetic content covered by the transparency obligation: audio, image, video or text. This also includes multimodal content, that is, content made up of a mix of these modalities. For the purpose of Article 50(2) AI Act, the different modalities are to be interpreted in line with practical and technological developments that may evolve over time. For the time being, they should be understood as follows:

([26]) As defined in Article 3(1) AI Act.

- Text refers to discrete symbolic content composed of characters, or numbers arranged in a particular sequence, that are capable of being read and interpreted semantically by humans.
- Images refer to static spatial representations encoding visual information that are capable of being seen by humans.
- Audio refers to a time-varying signal encoding sound that is capable of being perceived through hearing by humans. This may cover speech, instrumental music or other audio signals.
- Video refers to a time-based sequence of images, which may be synchronised with audio, capable of being seen by humans. In that case, both video and audio content must be marked and detectability must be ensured, as set out below.

(61) AI systems generating or manipulating 3-D images, 3-D videos and 3-D audio, as well as virtual, augmented or mixed reality, fall within the scope of Article 50(2) AI Act. 'Virtual reality' covers AI technologies that enable users to experience and interact with a computer-generated environment simulating physical presence and that allow real-time user interaction with that environment. As virtual and augmented reality is made up of a time-based sequence of images, optionally synchronised with audio, that represents visual change over time, it can be regarded as 'video' content. Relevant 'output' to be marked is the final form of the individual AI-generated asset with special exceptions for the real-time ephemeral interactive sessions under point 88 below. Digital artefacts used to generate 3-D images, 3-D videos and 3-D audio content such as a points cloud a 3D mesh are not required to be marked.

(62) Digital twins that provide a virtual replica of persons, physical objects or systems, integrating multiple data types, also fall within the scope of Article 50(2) AI Act, if they are AI generated or manipulated and take the form of audio, image, video or text, unless they fall within the exceptions and qualify as 'industrial applications' or 'business-to-business applications' under point 87 below or real-time ephemeral content under point 88 below.

(63) Article 50(2) AI Act may also apply to AI agents if the AI agent takes an action, the output of which is AI-generated or manipulated content perceptible by natural persons in the form of audio, image, video or text. Those outputs must therefore be marked and detectable as described below. In contrast, any AI agent´s action that does not include AI-generated or manipulated content (e.g., intermediate processing steps such as reasoning and chain of thought), or AI-generated action that is not intended to be directly perceived by natural persons (e.g. a web request or browser action), does not qualify as synthetic content within the scope of Article 50(2) AI Act and does not need to comply with the marking and detection obligations.

4.1.3. Content falling outside the scope of Article 50(2) AI Act

(64) Content generated by an AI system based on simple data processing, which is not specifically AI-generated or manipulated, for example a rendered frame, falls outside the scope of Article 50(2) AI Act.

(65) Furthermore, an AI system's output that merely reproduces existing content or enables the presentation or arrangement of already existing content is not covered by

Article 50(2) AI Act. This may be the case, e.g., for music playlists or other recommender systems that only select or rank existing content based on user preferences or activity, without creating anything new or manipulating the content itself. The same applies to AI systems used in internal analytical processes that extract and structure data, but do not summarise it.

(66) The same is true for an AI system's input and output that constitutes mere observations and recordings of data from physical or virtual environments, that is not AI-generated or manipulated (e.g. by robots or other AI-enabled sensors) or data transmissions by AI systems without any alteration. This covers a variety of AI systems' inputs and outputs, in particular in industrial settings such as manufacturing data for robots observing data; consumption recorded by AI-enabled smart meters, grid frequency and voltage measurements; and recording of GPS location data from vehicles, etc.

(67) Furthermore, since the objective of the transparency obligation in Article 50(2) AI Act is to enable humans to distinguish AI-generated or manipulated content, so as to address the risks of deception, manipulation and ensure integrity and trust in the information ecosystem, any content that is not related to that objective, and that is not perceptible or not intended to be interpreted, verified or acted upon by natural persons, is not targeted by this provision.

(68) In particular, the following content should be considered to fall outside the scope of Article 50(2) AI Act:

- Outputs generated in the form of a short sequence of numbers, symbols or letters (e.g. single words, image captions, alt-text, UI labels and icon-scale graphics, image captions and other data labels);
- Source code understood as content written in a programming, scripting, markup, query or configuration language that is intended to be interpreted, compiled or executed by a computing system, regardless of whether it is composed of text characters. This may also include natural-language comments and contextual information that form an integral part of the source code. Similarly excluded are Software Development Kits (SDKs), Structured Query Language (SQL), infrastructure-as-Code and 'YAML Ain't Markup Language' (YAML), JavaScript Object Notation (JSON) configuration. schemas, scripts, machine-readable specifications, Application Programming Interfaces (APIs & software libraries);
- Outputs of an AI system intended to be exclusively communicated from machine to machine and processed automatically without any exposure to humans, for example agent-to-agent communication, anti-spam signals, as long it is not perceived by natural persons;
- Outputs that are only used in closed loop environments in industrial and product development workflows (for example for film, animation, games or advertising production), unless they are the final output of the AI system in the form of AI-generated or manipulated text, audio, image or video content. In those cases, only the final output from workflow stages has to be marked and detectable.

4.2. The marking and detection obligation of Article 50(2) AI Act

(69) Article 50(2) AI Act sets out a transparency obligation for providers of generative AI systems falling within the scope of that provision comprising two distinct, but inherently interlinked elements. First, providers must ensure that the outputs of the AI system are marked in a machine-readable format. Second, providers must ensure that the outputs are detectable as artificially generated or manipulated content. For every marking solution deployed, providers should ensure that corresponding means for detection are available (as outlined below) to enable natural persons exposed to the content and other relevant actors to identify and distinguish the AI system's artificially generated content from other content. A technical solution is to be understood as a combination of techniques for marking and means for detection that the provider has implemented to fulfil the transparency obligation under Article 50(2) AI Act.

(70) Each of the two above-mentioned elements must be fulfilled to achieve effectively the objectives of the transparency obligations of Article 50(2) AI Act. Fulfilling only one element (e.g. for machine-readable marking of outputs without the means for their detection being available) will not suffice to comply with that provision ([27]).

4.2.1. The marking obligation

(71) The scope of the marking obligation is limited to implementing marks in a machine-readable format. A machine-readable format means that marks are structured in a way that allows software applications to easily identify, recognise and extract them without human intervention. Perceptible marks and labels are not excluded as a complementary measure, where appropriate, with a view to facilitating the compliance of deployers with their obligation to label deep fakes pursuant to Article 50(4) AI Act.

(72) To comply with the marking obligation, providers may rely on a single marking technique or a combination of techniques, so long as their overall technical solution is machine-readable and meets the requirements for effectiveness, interoperability, robustness and reliability to the legally required degree, as provided by the second sentence of Article 50(2) AI Act (see Section 4.2.3. below).

(73) Recital 133 AI Act provides examples of such techniques to include watermarks, metadata identifications, cryptographic methods for proving provenance and authenticity of content, logging methods, fingerprints or other techniques, and a combination of such techniques. While methods for proving provenance and authenticity are mentioned in Recital 133 AI Act, providers are not required to record or keep a full provenance chain containing information on content origin and modifications or any other relevant assertion concerning the history of the content. However, such provenance methods may also be used for compliance with Article 50(2) AI Act and be conducive in enabling natural persons to distinguish AI-generated or manipulated content from authentic content.

(74) Providers may implement the marking solution for the outputs at different stages of the value chain (e.g. post hoc after content generation at the level of the AI system, at the

([27]) Recital 133 AI Act clarifies that it is appropriate to require providers of those systems to embed technical solutions that enable marking in a machine-readable format and detection that the output has been generated or manipulated by an AI system and not a human. Recital 135 AI Act refers to the obligation regarding detection.

level of the underlying AI model or integrated in the AI system's inference process). They may rely on the marking solution implemented by an upstream model provider or a third party providing the solution (e.g. as an open standard or a specialised service) to the extent that the marking solution is compliant with Article 50(2) AI Act. Such reliance is without prejudice to the responsibility of the provider of the AI system to demonstrate compliance with Article 50(2) AI Act.

4.2.2. The detection obligation

(75) Providers of AI systems falling within scope of Article 50(2) AI Act must ensure that the output of their systems are detectable as AI-generated or manipulated content. This means that the provider is obliged to ensure that the means of detection are available to the persons potentially exposed to the content. Furthermore, pursuant to Article 50(5) AI Act, such detection solutions should be able to produce human-readable results indicating whether the content has been AI-generated or manipulated.

(76) A detection tool is a mechanism that detects whether content has been AI-generated or manipulated, typically identifying technical markers or signatures that verify its origin. To ensure full interoperability, providers must rely on publicly-available industry standard detection solutions that allow any third party to implement detection that is ideally locally executable on the digital device and allow seamless integration into self-contained, isolated networks. Where such standards are not available, in particular at the initial stage of the implementation of Article 50(2) AI Act for watermarking technologies, the provider may rely on its own detection solution or on a third party or shared detection solution, so long as those solutions ensure interoperability with the detection solutions used by other providers (see Section 4.2.3.). That possibility should be limited in time until harmonised standards and a standardised provider-agnostic interoperable detection solution emerge that is secure, privacy-preserving and locally executable.

(77) Article 50(5) AI Act further specifies that the information referred to under Article 50(2) AI Act should be provided to the natural persons exposed to the content in a clear and distinguishable manner at the latest at the time of the first interaction or exposure. This information is notably the result of the detection, which indicates whether content is AI-generated or manipulated and not the AI-generated or manipulated content itself. 'At the time of first exposure' should be understood as the moment at which natural persons are willing to verify the origin of a specific piece of content and access the detection solution to obtain the detection results.

(78) If the provider of the AI system depends on another actor for the provision or access to a detection solution, it must ensure that the detection solution complies with Article 50(2) and (5) AI Act and that those actors display the detection results in compliance with Article 50(5) AI Act (see Section 7).

4.2.3. Compliance with the requirements for technical solutions: effective, interoperable, robust and reliable

(79) The second sentence of Article 50(2) AI Act provides that the technical solutions for marking and detection must be effective, interoperable, robust and reliable. These requirements should be understood as follows:

- **Effectiveness** refers to the capability of the technical solution implemented by providers to detect their marks and enable natural persons to distinguish artificially generated or manipulated content produced by their AI system, and thus contribute to the trust and integrity of the information ecosystem.
- **Reliability** refers to the capability of the technical solution to accurately identify and distinguish AI-generated or manipulated content from other content in nominal conditions across the variety of content generated or manipulated by the provider´s AI system.
- **Robustness** refers to the capability of the technical solution to accurately identify and distinguish AI-generated or manipulated content produced by the provider's AI system from other content under varying conditions, covering both common alterations and adversarial attacks.
- **Interoperability** refers to the capability of the different technical solutions for marking and detection to operate seamlessly across multiple systems, actors, contexts and technical implementations to enable detection of AI-generated or manipulated content, regardless of the marking technique(s) deployed by providers of different AI systems.

(80) Providers must ensure that the combination of the technical solutions for marking and detection holistically meet all aforementioned requirements to the legally required degree, taking into account the limitations and the complementarities of the various solutions deployed. The technical solutions must comply with all requirements insofar as this is technically feasible, taking into account the specificities and limitations of various types of content, the costs of implementation, and the generally acknowledged state of the art, as may be reflected in relevant technical standards.

(81) 'Technically feasible' solutions within the meaning of Article 50(2) AI Act are solutions that are capable of being implemented for the modalities covered in the scope of the provision, using currently available technology, methods, and engineering practices, within the specific technical architecture and operational environment concerned. The provider is not obliged to make use of a technical solution that is not yet developed or available on the market, or that is technically unfeasible for implementation. Technical feasibility is an objective notion that is not dependent on the specific resources and capabilities of individual providers.

(82) Article 50(2) AI Act requires providers to implement technically feasible solution(s) that align with the generally acknowledged state of the art, in order to meet the four quality requirements of effectiveness, robustness, reliability and interoperability.

(83) The 'state of the art' is to be understood as a developed stage of technical capability at a given time as regards products, processes and services, based on the relevant consolidated findings of science, technology and experience and which is accepted as good practice in technology ([28]). The state of the art does not necessarily imply the latest scientific research still in an experimental stage or with insufficient technological

([28]) See Annex II of Commission Implementing Decision C(2025) 3871 final on a standardisation request to the European Committee for Standardisation and the European Committee for Electrotechnical Standardisation as regards high-risk AI system in support of Regulation (EU) 2024/1689 of the European Parliament and of the Council and repealing Implementing Decision C(2023)3215 final.

maturity. Providers must continuously adapt their marking and detection solutions in a timely and proportionate manner as the technology and state of the art evolves.

(84) In light of Article 50(7) AI Act, providers may rely on a code of practice, assessed as adequate pursuant to that provision to demonstrate compliance with the marking and detection obligations in Article 50(2) and (5) AI Act, which is aligned with the state of the art. For further information regarding the effects of adherence to such a code and the supervision of other providers who do not adhere to a code, see Section 8.1 below.

(85) The costs of implementation of certain technical solutions for marking and detection can also be taken into account. In certain cases, as described in the following points, the costs may be disproportionate to marginal gains with more limited value for enabling humans to distinguish AI-generated or manipulated content, and ensuring the integrity and trust in the information ecosystem.

(86) There may therefore be narrowly defined cases where a technical solution based on less robust metadata markings is sufficient for the purpose of complying with Article 50(2) AI Act, taking into account the specificities of the AI system's outputs if they pose inherently lower risks of deception, manipulation or negative effects on the information ecosystem. This is specifically the case where a generative AI system is embedded in physical products generating outputs in a technically controlled and closed environment that is mainly instructive in nature (e.g., an AI system embedded in navigation systems in vehicles) to the extent effective technical measures are embedded in the product to prevent the output from leaving the product environment, for example by being captured or exported and disseminated externally.

(87) There may also be limited cases of "industrial AI applications" or "business to business applications" (excluding public and consumer-facing AI systems) where the proportionality principle and the limited benefits for the transparency objectives means that no marking and detection of AI-generated or manipulated outputs is required. This applies only if the following requirements are met cumulatively:

i. The AI system's generated output is strictly technical in nature, for example engineering designs, industrial production workflows, technical instructions, output generated as a result of predictive system maintenance processes in industrial settings, internal documentation processes or production steps and workflows before the output is finalised and made available to other external persons or the public.

ii. The AI system's generated output is only intended to be perceived and processed by a limited pre-defined number of natural persons acting in a professional capacity within the organisation of the provider and the deployer.

iii. The AI system's generated output must not be intended to be shared outside the company or to be usable by external persons, with appropriate safeguards in place to avoid reasonably foreseeable misuse (e.g. cloud isolation, role-based controls).

(88) Real-time content generation that is ephemeral and consumed immediately, without being recorded, stored or disseminated further (e.g. in video games, virtual reality applications), may also be exempted when marking is not technically feasible and the persons exposed to the content are made aware that the content is AI-generated or manipulated (e.g. in-experience disclosure, session-level notifications).

4.3. Exceptions to the obligations under Article 50(2) AI Act

(89) Article 50(2) AI Act provides three explicit exceptions from the transparency obligations laid down in that provision

(90) The first exception concerns AI systems to the extent that they perform an assistive function for standard editing. Standard editing should be understood as the process of preparing existing content for publication or distribution (e.g., small edits to improve readability and grammar, quality and format) and does not involve generating new content. Standard editing aims at ensuring that content is, among others, free from obvious technical or grammatical linguistic errors, is aligned with applicable lay-out, presentation, formatting or accessibility requirements, or in conformity with sectoral practices. Editing goes beyond standard editing if the content is changed in a material way (substantive modifications, structural changes etc.) that affect its meaning, style or intent.

(91) The second exception applies where an AI system does not substantially alter the input data provided by the deployer or the semantics thereof. An alteration should be considered substantial if the input data or its semantics have been manipulated significantly by the AI system during its output generation process. This must be based on an assessment of relevant factors, such as the format, media content type, style and changes in the content that affect its meaning, style or intent. Whether that is the case requires a case-specific assessment.

(92) If an AI system can be used both for content generation or manipulation and for non-substantial minor alterations of input data, then the transparency obligation under Article 50(2) AI Act will not apply to the content altered in a minor non-substantial manner.

Examples of standard editing and minor alterations benefitting from the exception under Article 50(2) AI Act:

- Grammar correction and spellchecking, linguistic and minor stylistic polishing that do not change the substance, meaning, style or messaging of text, AI-generated translations of text;

- Formatting, format conversions, technical compression, noise reduction or removal for enhanced clarity without changing the meaning or the substance of the content;

- Minor cropping, minor colour adjustments or corrections, lightening or darkening, sharpening for enhanced clarity or other standard editing and technical corrections;

- Removal of dust spots caused by a dirty lens or sensor, removal of red-eye caused by flash photography; deleting and obscuring backgrounds that are visible in the original file, pixelation or blurring of faces;

- Rescaling of a video clip, dynamic range compression and equalisation; limited video stabilisation;

- Standard adjustments to colour and contrast, minor adjustments to playback speed, minor corrections to level the horizon of an image, applying pixel filters to amplify certain parts of an image, or applying colour maps to grayscale images, edge image completion, converting a black & white to colour image or video and vice versus;

- Pixel filling for aspect ratio adaptation or enhanced clarity, automatic transition clips, other non-substantive edits for cosmetic and technical purposes.

- Medical image technical processing or reconstruction within medical devices (e.g. annotations, boxes, contours, circles, heatmaps);

- Transcriptions of conversations; AI-generated content that only transforms authentic human input through assistive technologies allowing persons with disabilities to communicate (e.g., augmentative and alternative communication (AAC) or customized neural voices (CNV)) since they do not alter semantically the meaning of the content.

Examples of semantic changes that require marking under Article 50(2) AI Act:

- AI-generated summaries of text; paraphrasing or rewriting text that changes style, structure and meaning beyond mere grammatical and minor stylistic correction;

- Removal, replacement or insertion of objects or persons in existing images and videos that changes meaning and substance of the content; face replacement or substantial facial modification,

- Synthesis of realistic speech in a specific person's voice, or generation of realistic video depicting events that did not occur; altering the body shape or the skin colour of a person; extreme lightening, darkening, colour and contrast changes that change the meaning, intent and messaging of the content;

- Creation of composite images or video clips that modifies the representation of persons, objects, events or facts, any other substantial alteration of the content.

(93) The third exception applies to generative AI system authorised by law to generate or manipulate synthetic content to detect, prevent, investigate or prosecute criminal offence The clarifications provided in points 46-48 above for this exception are also relevant in this case.

4.4. Interplay with other Union legal acts

(94) The marking and detection obligations under Article 50(2) AI Act focuses on how the content has been created and its artificial origin, not on who created the content. Any marking and detection solutions employed by providers of AI systems must therefore be compliant with applicable Union data protection law, including data protection principles and obligations, such as the obligation of data protection by design and by default, the data minimisation and the security and confidentiality principles. For example, for the purpose of marking and detection the information related to the creator of the content should not be processed and any other personal data processed should be deleted as soon as the purpose of detection is fulfilled, security and confidentiality must be ensured, etc.

(95) Recital 136 AI Act further highlights the particular relevance of the transparency obligations under Article 50(1), (2) and (4) AI Act to facilitate the effective implementation of the DSA. This applies in particular as regards the obligations of providers of very large online platforms ('VLOPs') or very large online search engines ('VLOSEs') to diligently identify, analyse, assess and mitigate systemic risks that may arise from the dissemination of content that has been artificially generated or manipulated, in particular the risk of the actual or foreseeable negative effects on

democratic processes, civic discourse and electoral processes, including through disinformation (Articles 34 and 35 DSA). Machine-readable marks may enable those providers to detect content that is AI-generated or manipulated and to provide appropriate labelling and disclosures, thus diminishing the risk of AI-generated disinformation across their services.

(96) The obligations to mark and label content generated or manipulated by AI systems under Article 50(2) and (4) AI Act are without prejudice to the obligation in Article 16(6) DSA for providers of hosting services to process notices on illegal content received pursuant to Article 16(1) DSA ([29]). In particular, marking or labelling applied to AI-generated or manipulated content should not influence the assessment and the decision on the illegality of the specific content under other regulatory frameworks. That assessment should be performed solely with reference to the rules governing the legality of content in accordance with those frameworks ([30]). For example, if a labelled deep fake image is flagged as potentially child sexual abuse material (CSAM) or infringing on trademark or copyright law, then the application of a label does not affect the potential illegality of the content under those regulatory frameworks. That assessment should be conducted solely on the basis of the respective, applicable laws (e.g. criminal law, trademark or copyright rules).

(97) Conversely, the machine-readable marking embedded in the content pursuant to Article 50(2) AI Act can help providers of VLOPs and VLOSEs to fulfil their obligations under Article 34 and 35 DSA to identify, analyse and assess systemic risks that may arise from the dissemination of content that has been artificially generated or manipulated, in particular the risk of the actual or foreseeable negative effects on democratic processes, civic discourse and electoral processes, including through disinformation, ([31]) and to put in place reasonable, proportionate and effective measures to mitigate those risks, tailored to the specific systemic risks identified. Providers of VLOPs and VLOSEs may also qualify as providers of AI systems if they meet the relevant conditions under the AI Act where they integrate AI systems into their services (see also Section 6.1.5).

(98) Providers of online platforms, online search engines and other relevant actors distributing content along the value chain are in particular encouraged to preserve and enable identification of the marks implemented by providers pursuant to Article 50(2) AI Act, so that the natural persons exposed to the content can be informed about its origin.

5. ARTICLE 50(3) AI ACT: EMOTION RECOGNITION SYSTEMS AND BIOMETRIC CATEGORISATION SYSTEMS

(99) Article 50(3) AI Act imposes an obligation on deployers of emotion recognition systems and biometric categorisation systems to inform natural persons who are exposed to those systems of the operation of the system.

(100) Recital 132 AI Act explains that the purpose of that obligation is to ensure that natural persons (including persons with disabilities) are aware that they are exposed to emotion recognition and biometric categorisation systems that can be intrusive for their privacy.

([29]) See also recital 11 AI Act.
([30]) See also recital 137 AI Act.
([31]) Recital 136 AI Act.

The obligation therefore applies regardless of whether the persons are exposed to such systems in real-time or those systems are operated ex post.

5.1. Main components, concepts and related transparency obligation(s) under Article 50(3) AI Act

5.1.1. The notion of an emotion recognition system

(101) Article 3(39) AI Act defines an 'emotion recognition system' as 'an AI system for the purpose of identifying or inferring emotions or intentions of natural persons on the basis of their biometric data'. The detailed explanation of this notion provided in the Commission guidelines on the classification of high-risk AI systems, including the practical examples set out therein, are equally relevant in the context of Article 50(3) AI Act.

(102) Since all emotion recognition systems are also classified as high-risk, unless prohibited under Article 5(1)(f) AI Act in the areas of workplace and education, the transparency obligation in Article 50(3) AI Act should apply in conjunction with the other safeguards and requirements applicable to high-risk AI systems.

5.1.2. The notion of a biometric categorisation system

(103) Article 3(40) AI Act defines a 'biometric categorisation system' as 'an AI 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'. The detailed explanation of this notion provided in the Commission guidelines on the classification of high-risk AI systems, including its specific elements as well as the practical examples set out therein, are equally relevant in the context of Article 50(3) AI Act.

(104) Importantly, unless prohibited under Article 5(1)(g) AI Act, Article 50(3) AI Act applies to any biometric categorisation systems (e.g. age or gender classification based on biometric data), regardless of whether they fall or not in the scope of high-risk AI systems under the AI Act.

5.1.3. The information obligation under Article 50(3) AI Act

(105) **The scope of the information to be provided**: Deployers are required to inform natural persons concerned that they are exposed to an emotion recognition system or to a biometric categorisation system, which is being operated. The AI Act does not require including information about the reasons for the system's operation, or if the system also processes information for purposes other than emotion recognition or biometric categorisation, which are required to be disclosed under Union data protection law (see Section 5.3.).

(106) **The addressees of the information:** Article 50(3) AI Act requires that deployers inform all natural persons exposed to the operation of the AI system, including children.

(107) **The means of providing information:** The AI Act does not prescribe a specific means to provide the information, except that Article 50(5) AI Act requires that the information needs to be provided in a clear and distinguishable manner and in accordance with applicable accessibility requirements (responding to the needs of natural persons with disabilities). How natural persons exposed to the system are to be informed may depend on the place of deployment (e.g., virtual gaming platform, bricks and mortar store, train station), the possible addressees (e.g., children, elderly, persons with disabilities, customers), the length and type of the exposure (e.g., long-term, one-off, continuous, intermittent), and the possible existence of a relationship between the deployer and the persons exposed (e.g., if there is an existing communication channel with the individual). Depending on the use case, the information may be provided in writing, by standardised icons (also when presented electronically), orally, or by using combinations of such ways.

Examples of informing natural persons:

- A centrally placed pop-up with an onboarding message before a computer game is launched indicating that the player's face is recorded, capturing their emotions.

- A visible notice at each possible entrance to an exhibition room with information that visitors' facial images are captured when entering the room to assign them to a specific age group.

(108) **The timeline for providing the information**: The information must be provided at the latest at the time of the first interaction with or exposure to the AI system in accordance with Article 50(5) AI Act. Where appropriate, providing the required information in advance is not excluded.

5.2. Exception for law enforcement purposes

(109) The information obligation does not apply to emotion recognition systems and biometric categorisation systems that are permitted by law to detect, prevent or investigate criminal offences subject to appropriate safeguards for the rights and freedoms of third parties and in accordance with Union law. The clarifications provided for this exception in points 46-48 above are also relevant in this case with one difference. While the other exceptions in Article 50 AI Act require an explicit law authorising the non-transparent use of the AI systems, Article 50(3) AI Act makes such use possible when permitted under the legal rules governing the powers of law enforcement authorities to use emotion recognition and biometric categorisation systems to detect, prevent, or investigate criminal offences without informing the persons concerned, subject to appropriate safeguards.

5.3. Interplay with other Union legal acts

(110) The information obligation under Article 50(3) AI Act does not, in itself render the use of an emotion recognition and biometric categorisation system lawful. Nor does it legitimise intrusive or discriminatory uses that might be unlawful and prohibited under Article 5 AI Act or other Union law. Deployers need to comply with the information obligation in Article 50(3) AI Act, in addition to any applicable information requirements under Union data protection law. In certain cases, deployers can consider adding notifications pursuant to Article 50(3) AI Act to the information provided to data subjects under Union data protection law.

6. ARTICLE 50(4): LABELLING OF DEEP FAKES AND CERTAIN TEXT PUBLICATIONS

(111) Article 50(4) AI Act lays down two separate obligations addressed to deployers of generative AI systems requiring clear and distinguishable disclosures of (i) deep fakes and (ii) AI-generated or manipulated text published with the purpose of informing the public on matters of public interest except in defined cases. These labelling obligations apply in addition and without prejudice to the obligations for machine-readable marking and detection under Article 50(2) AI Act applicable to providers of AI systems generating or manipulating synthetic images, video, audio or text content.

6.1. Main components, concepts and related transparency obligation(s) for deep fakes under Article 50(4) AI Act

(112) Article 50(4), first subparagraph, AI Act applies if several conditions are fulfilled:

i. The system must be an AI system;

ii. Used by deployers for professional purposes (see Sections 2.3. and 2.4.1.);

iii. To generate or manipulate image, audio or video content that constitutes a deep fake;

iv. The AI system must not fall under the exception for use authorised by law to detect, prevent, investigate or prosecute criminal offence.

6.1.1. The notion of 'deep fake'

(113) Article 3(60) AI Act defines 'deep fakes' as 'AI-generated or manipulated image, audio or video content that resembles existing persons, objects, places, entities or events and would falsely appear to a person to be authentic or truthful'. The difference between manipulation and generation has been clarified in Section 4.1 above. For the purpose of clarifying the notion of deep fake, four cumulative criteria from that definition are examined below: (i) resemblance of (ii) existing (iii) persons, objects, places, entities or events (iv) that would falsely appear to a person to be authentic or truthful.

i. **Resemblance:** Article 3(60) AI Act specifies that AI-generated or manipulated image, audio or video content must 'resemble' existing subjects to constitute a deep fake. Recital 134 AI Act adds that such resemblance should be 'appreciable'. AI-generated or manipulated image, audio or video content appreciably resembles a subject if there is a high level of similarity between the deep fake content and the subject (including any of its recognisable elements) being simulated by the deep fake. The content does not need to be identical to the subject. Whether the level of resemblance is appreciable is a case-by-case assessment to be done by the deployer based on an objective comparison between the simulated subject and the deep fake content, based on, among others, the extent to which characteristic or distinctive features are represented by the deep fake.

ii. **Existing:** Article 3(60) AI Act specifies that AI-generated or manipulated image, audio or video content must resemble 'existing' subject(s) to constitute a deep fake. To effectively fulfil the purpose of Article 50(4) AI Act to reduce the risks of impersonation, deception, misinformation, manipulation and fraud, the characteristic

of existence implies that the AI-generated or manipulated image, audio or video content should resemble realistic subjects (e.g., persons, objects, places). Therefore, it is sufficient for simulated persons, objects, places, entities or events to resemble someone or something that exists, can plausibly exist or could have plausibly existed in reality to be considered a deep fake, since such content can be a major source of deception, manipulation and disinformation. By contrast, simulated persons, objects, places, entities or events that, for example, defy the laws of nature or physics or depict lifeforms that are not commonly accepted in biology (such as e.g. humans flying without mechanical aids, dragons, or elephants driving cars) and have no potential to mislead are considered unrealistic and therefore fall outside the scope of the transparency obligation.

iii. **Persons, objects, places, entities or events:** Article 3(60) AI Act specifies that AI-generated or manipulated image, audio or video content should resemble existing 'persons, objects, places, entities or events' to constitute a deep fake. 'Persons' is to be understood as realistic, human beings (including digital replicas of real persons, realistic AI-generated human avatars or personas, and personal characteristics or expressions, such as image, voice, behaviour, performances etc.). 'Objects' is to be understood as realistic, inanimate material items, including buildings, artworks, machinery, consumer goods etc. 'Places' is to be understood as realistic locations. 'Entities' is to be understood as realistic, non-human but animate beings including animals or other biological lifeforms. 'Events' is to be understood as realistic scenes or situations that can involve persons, objects, places and entities (e.g. evoking historical events or the depiction of professional or consumer services).

iv. **False appearance to a person to be authentic or truthful:** Fourth, Article 3(60) AI Act specifies that the AI-generated or manipulated image, audio or video content resembling existing persons, objects, places, entities or events 'would falsely appear to a person to be authentic or truthful'. This criterion relates to the essential characteristic of deep fake content and its capacity to potentially deceive or mislead a person regarding the content's authenticity or truthfulness ([32]).Content authenticity refers to whether the content is genuinely what it purports to be in terms of its source or creation process (including e.g. the involvement of real human beings or animals and their actual behaviour or actions, the actual appearance or use of objects or the delivery of services, the accurate unfolding or course of an event). Truthfulness pertains to the veracity of the content (e.g. factual correctness of the representations in the deep fake).

(114) Whether content 'falsely appears to a person to be authentic or truthful' should be assessed as a whole, taking into account the level of resemblance, the potential substantive message of the content, the intended and foreseeable deployment contexts, the environment in which the content is presented, and the intended and reasonably foreseeable audience composition and their expectations. However, this assessment is objective and does not require the intention of the deployer to deceive or mislead the natural persons exposed to the content for it to constitute a deepfake. As regards the deployment context and audience expectations, this means that if the content is used in specific deployment contexts where the intended and reasonably foreseeable audience does not expect content to be authentic or truthful, the AI-generated or manipulated

([32]) Therefore, this criterion must be understood independently and distinctly from the concept of the deception as mentioned by Directive 2005/29/EC concerning unfair business-to-consumer commercial practices in the internal market (Unfair Commercial Practices Directive), https://eur-lex.europa.eu/eli/dir/2005/29/oj/eng

content may not *falsely* appear to be authentic or truthful even if it is, in fact, non-authentic or untruthful. In such cases, the fourth criterion of the deep fake definition will not be fulfilled. For example, the AI-generation or manipulation of background scenes, special effects, or technical pre- and post-processing as part of standard movie production processes are not likely to make content *falsely* appear to the audience to be authentic or truthful. Conversely, the use of AI for generating or manipulating essential elements that impact audience perception is likely to do so (such as fully AI-generated actors, digital replicas of real or deceased actors, de-aging, simulated performances including parts thereof, or non-authentic or untruthful representation of persons, objects, places or events in documentaries). For image and video content, it is acknowledged that a high degree of photorealism renders it more likely that such content should be considered a deep fake (as it likely resembles existing subjects), but photorealism alone is not determinative for the assessment. What should still be assessed, is whether such content is capable of deceiving or misleading a person regarding the content's authenticity or truthfulness while taking into account, among others, the intended and foreseeable deployment context.

(115) As opposed to Article 50(1) AI Act, the assessment of false appearance to a person to be authentic or truthful should not be based on a hypothetical "average" person expected to be exposed to the content (see Section 3.2.1), but should duly take into account the possible diverse composition of the reasonably foreseeable audience that may be exposed to the deep fake content. Due consideration of audience composition is especially important if it is reasonably foreseeable that the deep fake content may, at some point, be perceived by children, the elderly, or other groups of persons with lower digital and AI literacy or general knowledge levels, since they may be more easily deceived or misled regarding the content's authenticity or truthfulness. In such cases, false appearance of the content to be authentic or truthful to that part of the foreseeable audience may suffice to qualify content as a deep fake. However, deployers are not expected to consider potential further dissemination of the deep fake content by third parties beyond the reasonably foreseeable audience by the deployer, taking into account the intended distribution channels (e.g. if deep fake content is shown solely on a subscriber-only part of a website or as part of a corporate newsletter, this does not imply that the deployer should consider broad public accessibility by default).

(116) In a similar vein, AI-supported manipulation of insignificant substantive or technical aspects of pre-existing content may be of minor relevance for a person's assessment of the authenticity or truthfulness of the content, not rendering the resulting content to become a deep fake. This could include, for example, editing background details (e.g. removing passerby), lighting adjustments, adapting audio parameters, colour correction, noise reduction or removal, improving accessibility or file compression, cosmetic adjustments and enhancements. Whether the manipulation will fall into this category depends on the context and the impact on the persons' perception of the content's authenticity or truthfulness in the specific case. For example, AI-powered colour correction, background extensions of existing content, adjustments or replacements of backgrounds for clearly aesthetic purposes, compositions and arrangements of existing products, or re-scaling of images applied in product advertisements or packaging is likely to have only a minor impact on a person's perception of the authenticity and truthfulness of the advertisement and the product. Conversely, substantial AI-powered editing of background details of journalistic images beyond standard technical, editorial practices may negatively affect a person's perception of the content's authenticity and truthfulness.

Examples of deep fakes under Article 3(60) AI Act

- AI-manipulated image of a scene involving two real professional footballers in front of a building resembling a football stadium.

- AI-generated audio involving voice cloning of a newspaper podcast's regular presenters and a guest interlocutor discussing some news events.

-AI-generated video of an individual resembling a politician, holding a speech in front of an audience.

- AI-generated video featuring an AI-generated depiction of a celebrity influencer in an advertising or promotional context.

- AI-generated video featuring a realistic synthetic avatar of a company CEO congratulating employees with their work and the corporate results of the past year.

- An AI-generated image of a product in advertisement or packaging that can affect the audience's perception and mislead as to the actual product appearance, characteristics or use (e.g. making the product appear not identical to the real product, more appealing or with improved quality than in real life).

Examples that do not constitute deep fakes under Article 3(60) AI Act

- AI-generated image of a scene involving a sphinx flying over the Eiffel Tower.

- AI-generated video of mice arguing in human language over the best type of cheese as part of an advertisement campaign for a cheese manufacturer.

- AI voice replication for fictional characters (e.g., for audiobooks, games or animation) when there is no deception as to the identity of the narrators.

- AI-manipulated radio broadcast adjusting technical audio parameters (e.g. normalising volume levels, noise reduction, audio compression) without altering the actual words spoken by speakers or their way of speaking.

- AI-generated cartoon of a pre-existing image depicting a historical event.

- Video of a presenter of a scientific TV-programme who walks across a mountainous region while addressing the retreat of glaciers which is illustrated by an AI-manipulated background animation that shows how the glaciers retreated from their historical levels to their current levels.

- AI-generated fictitious environments (e.g. fictitious forests and castles) in video games.

- Movies where real existing actors (not manipulated by AI) are playing against an AI-generated background (e.g. depicting a fictitious stretch of prairie, historical buildings of ancient cities etc.).

- A real product (e.g., a car) shown in an advertisement against an AI-generated background and surrounding environment as long as the ad is not likely to mislead the audience about the product's actual representation and its characteristics and use.

6.1.2. The disclosure obligation under Article 50(4), first subparagraph, AI Act

(117) Article 50(4), first subparagraph, AI Act obliges deployers of AI systems that generate or manipulate image, audio or video content constituting a deep fake to 'disclose that the deep fake content has been artificially generated or manipulated'. Recital 134 AI Act clarifies that deployers using AI systems to generate or manipulate deep fake content should clearly and distinguishably disclose that such content has been artificially created or manipulated by labelling the AI output accordingly and disclosing its artificial origin. Labelling or disclosure methods applied in accordance with Article 50(4), first subparagraph, AI Act should be understandable and perceivable by natural persons (e.g. with visible or audible labels), without them needing to rely on any specific technical tools or performing dedicated actions. Therefore, deployers cannot rely on the machine-readable marking embedded in the content by the provider under Article 50(2) AI Act, since those markings are not immediately clear and distinguishable for the natural persons exposed to the deep fake content.

(118) In light of Article 50(7) AI Act, deployers may adhere to a code of practice, assessed as adequate pursuant to Article 50(7) AI Act to demonstrate compliance with this disclosure obligation. For further information regarding the effects of adherence to such a code and the supervision of other deployers not adhering to a code, see Section 8.1 below.

6.1.3. Transparency of artistic, creative, satirical, fictional or analogous deep fake content

(119) Article 50(4), first subparagraph, AI Act foresees an attenuated transparency obligation for deep fakes forming part of evidently artistic, creative, satirical, fictional or analogous works or programmes. For such content, the transparency obligation is limited to the disclosure of the deep fake in an appropriate manner that does not hamper the display or enjoyment of the work.

(120) **Relevant content categories**: To benefit from the attenuated transparency disclosure, the deep fake (image, audio or video) should form part of, at least, one of the following content categories: artistic, creative, satirical, fictional, or analogous works or programmes. Whether a deep fake falls under one or more of these categories requires a case-by-case assessment by the deployer. For the purpose of Article 50(4) AI Act, those categories are understood as follows:

- Artistic works are to be understood as works that have been created for the purpose of art, including music, cinematographic works, and visual arts;
- Creative works are to be understood as works that involve creative choices. Works that are mainly motivated by functional or technical considerations cannot be regarded as creative;
- Satirical works are to be understood as works that are intended to criticise society, politics, business or public figures through the use of humoristic techniques (including irony, sarcasm, mockery, pastiche etc.);
- Fictional works are to be understood as cultural works that involve persons, objects, places, entities or events, in an imaginary, but verisimilitude setting;

- Analogous works are to be understood as works sharing core traits or similar expressive or functional characteristics with the above categories, but not fitting neatly into one.

(121) The above content categories may also apply to programme(s) that should be understood as an individual item within a schedule or catalogue, established by a media service provider, that is comparable in form and content to television broadcasting ([33]). The reference to 'works' in the context of Article 50(4) AI Act has no bearing on whether the content qualifies for protection under Union law on copyright and related rights.

(122) **Evidently:** The fact that a deep fake falls within one (or more) of the above content categories should be evident to the natural persons exposed to it. This implies that the categories should be interpreted strictly, justified by the fact that a lighter disclosure regime applies due to the potential impact on the freedom of expression and the freedom of the arts and sciences, while risks of misinformation, manipulation and deception still need to be addressed adequately. Therefore, content whose nature is potentially unclear or ambiguous to the audience is excluded from the scope of this lighter transparency regime. Relevant factors for assessing the evident nature of artistic, creative, satirical or fictional content include: (i) whether the content displays formats or styles that are characteristic of the content categories (e.g. irony or exaggeration for satirical works, certain art styles, displaying distinctive themes, techniques or aesthetics, etc.); (ii) the context in which the content is presented (e.g. if the platform, medium or place where the content is presented, made available, or distributed is associated with artistic, creative, satirical or fictional use); and (iii) audience expectations (e.g. a movie, gaming environment, or virtual reality scene). In addition, the condition that content should 'evidently' fall in one of the above content categories excludes content from the scope of those content categories if its nature is exclusively informative or commercial and is recognisable as such (e.g. news reporting). In this respect, some kinds of content (e.g. advertisements or documentaries) containing deep fakes might be regarded as evidently creative or fictional works in certain, specific situations, but not in others, since the assessment is case-specific. When the deep fake combines multiple characters (e.g. informative and creative), the informative character should always prevail and require compliance with the standard labelling requirements.

(123) **Appropriate disclosure not hampering the display or enjoyment of the work**: Deep fakes that form part of evidently artistic, creative, satirical or fictional works or programmes are not excluded from the transparency obligation of Article 50(4), first subparagraph, AI Act. Deployers still need to disclose the AI-origin of the content or its manipulation, but they can do so in an appropriate manner that does not hamper the display or enjoyment of the work. Recital 134 AI Act clarifies that such appropriate disclosure should not hamper the normal exploitation and use, while allowing to maintain the utility and quality of the work. Determining which disclosure measures are to be considered appropriate is a case-by-case assessment, whereby deployers may consider all relevant factors (including nature of the work, audience, context, etc.). In any case, deployers need to comply with Article 50(5) AI Act (see Section 7). In light of Article 50(7) AI Act, deployers may adhere to a code of practice, assessed as adequate pursuant to that provision, to demonstrate compliance with this disclosure obligation. For further information regarding the effects of adherence to such a code and supervision of other deployers not adhering to a code, see Section 8.1 below.

([33]) See in a similar vein Article 1(b) AVMSD.

(124) **Appropriate safeguards for the rights and freedoms of third parties:** Recital 134 AI Act clarifies that compliance with the attenuated transparency obligation and the use of artistic, creative, satirical or fictional deep fakes must be 'subject to appropriate safeguards for the rights and freedoms of third parties'. Therefore, deployers need to ensure that the rights and freedoms of third parties (including e.g. their right to the protection of personal data or intellectual property rights) are adequately safeguarded and respected, when creating or publishing deep fakes, even if the deep fake is to be considered artistic, creative, satirical or fictional. Reliance on the attenuated transparency obligation cannot be a justification for failing to respect the fundamental rights of individuals or rightsholders under Union law on intellectual property or Union data protection law ([34]).

Examples of deep fakes that are creative, satirical, fictional or analogous works

- Artistic/fictional work: movies (or movie trailers) featuring AI-generated or manipulated deep fakes of de-aged, existing actors or digital replicas of dead actors or other persons, available in a cinema or on a streaming platform.

- Artistic/creative work: AI-generated music in any kind of genre resembling the individual style of existing artists.

- Satirical/fictional work: AI-manipulated image of an existing politician placing him in a scene clearly meant to criticise in a humorous manner certain policy decisions taken or supported by that person.

- Analogous creative/fictional work: AI-generated gaming imagery involving deep fake simulations of real, existing persons.

Examples of deep fakes that do not constitute artistic, creative, satirical, fictional or analogous work:

- AI-manipulated video in the style of a teleshopping channel involving deep fake simulation of humans advertising a product in an AI-generated scene depicting the functional use of the product by the simulated consumers with the aim of persuading viewers to buy the product.

- AI-generated image of celebrities implying their involvement in activities that never happened, lacking any fictional, satirical or analogous purpose.

-AI-manipulated video featuring a realistic synthetic influencer testing out a sponsored real product, focused solely on displaying the product functionalities.

- AI-generated video depicting realistic, holocaust scenes being shared on publicly available social media platforms.

6.1.4. Exception for law enforcement

(125) If the use of a deep fake (e.g. an AI-generated video simulating events that never happened) is authorised by law to detect, prevent, investigate or prosecute criminal offences, deployers are exempted from the transparency obligation under Article 50(4),

([34]) This equally applies to published texts informing the public on matters of public interest.

first subparagraph, AI Act. The clarifications provided in points 46-48 above for a similar exception in Article 50(1) AI Act are also relevant for this provision.

6.1.5. Interplay with other Union legal acts

(126) The transparency obligation under Article 50(4), first subparagraph, AI Act will interact with the obligations of providers of VLOPs and VLOSEs to identify and mitigate systemic risks pursuant to Articles 34 and 35 DSA, in particular, Article 35(1)(k) DSA ([35]). The latter provision lists, among the possible risk-mitigation measures that providers of VLOPs and VLOSEs may adopt, measures ensuring that 'an item of information, whether it constitutes a generated or manipulated image, audio or video that appreciably resembles existing persons, objects, places or other entities or events and falsely appears to a person to be authentic or truthful is distinguishable through prominent markings when presented on their online interfaces'. In addition, providers of VLOPs and VLOSEs may also provide an easy-to-use functionality which enables recipients of the service to indicate such information. While this DSA provision seems similar to Article 50(4) AI Act, it is complementary to it and contains two important distinctions:

- **Different material scope of application**: Compared to Article 50 (4) AI Act, Article 35(1)(k) DSA covers a broader range of content that may falsely appear authentic or truthful, regardless of the technologies used to create the content (i.e. it covers content that has been generated or manipulated by using AI systems or by any other tool).

- **Different personal scope of application**: Article 50(4) AI Act obliges deployers of AI systems generating or manipulating deep fakes to label such content, while Article 35(1)(k) DSA applies to providers of VLOPs and VLOSEs disseminating the content or making such content discoverable through search. For example, deployers of AI systems integrated into a VLOP under the AI Act may be considered 'recipients of the service' under Article 3(b) DSA ([36]) if they use the VLOP to disseminate content generated with the respective AI system, while the VLOP may be considered the provider of the AI system if the conditions of Article 3(3) AI Act are fulfilled. Where providers of VLOPs or VLOSEs make labelling tools available to such deployers enabling them to label their deep fake content in compliance with Article 50(4) AI Act (i.e. a clear and distinguishable disclosure of the AI-origin), those deployers can rely on such tools to fulfil their transparency obligation under that provision within the context of the VLOP or VLOSE used. Providing such a functionality is without prejudice to the responsibility of the deployers under the AI Act to fulfil their labelling obligations under Article 50(4) AI Act. Furthermore, providers of VLOPs and VLOSEs will also be considered deployers of an AI system under the AI Act if they use an AI-system under their authority for their own professional purposes (e.g. creating visuals to be used as part of marketing material).

(127) If the deep fake includes information relating to a directly or indirectly identifiable natural person (who is alive), then deployers as controllers under Union data protection law will also need to comply with the respective obligations under Union data protection

([35]) See also Section 4.4 for further information regarding the interplay between Article 50 AI Act and the DSA.
([36]) Recital 2 DSA.

law ([37]). This includes, for example, reliance on an appropriate legal basis for the processing of the personal data, information regarding the processing of data relating to the depicted data subjects, etc.

(128) Deployers of AI systems generating deep fakes will furthermore need to ensure that they comply with relevant Union intellectual property laws (such as Union trademark law and Union law on copyright or related rights) if their deep fake incorporates subject matter protected under those laws. The application of a label or icon to a deep fake does not, as such, have any impact on its eligibility for copyright protection, which should be assessed in accordance with applicable Union copyright law.

(129) Where existing persons are depicted in deep fakes, deployers will also need to comply with relevant legal regimes protecting personality rights over personal images or voices. The transparency obligation of Article 50(4), first subparagraph, AI Act does not imply that AI-generated or manipulated deep fakes that are harmful and unlawful under the applicable Union or national law (e.g. misleading advertising or criminal law such as child sexual abuse material and non-consensual intimate images) may be generated and disseminated (see also Section 4.4).

6.2. Main components, concepts and related transparency obligation(s) for AI-generated or manipulated text under Article 50(4) AI Act

(130) Article 50(4), second subparagraph, AI Act applies if several conditions are fulfilled:

i. The system must be an AI system;

ii. Used to generate or manipulate text (see Section 4.2.1.) published with the purpose of informing the public on matters of public interest;

iii. The use of the AI system does not fall under any of the exceptions where (1) the AI-generated content has undergone a process of human review or editorial control and where a natural or legal person holds editorial responsibility for the publication of the content or (2) the use is authorised by law to detect, prevent, investigate or prosecute criminal offences.

6.2.1. Text published with the purpose of informing the public on matters of public interest

(131) For the purpose of clarifying the scope of Article 50(4), second subparagraph, AI Act, three elements are examined below: (i) published text (ii) with the purpose of informing the public (iii) on matters of public interest.

i. **Published text:** The AI-generated or manipulated text should be published. This means that the text should be accessible by an indeterminate, fairly large number of unrelated, potential readers simultaneously and/or successively, whether or not against payment (e.g. subscriptions) ([38]). By contrast, text is not considered published if access is

([37]) For more detailed clarification, see the joint Commission and EDPB guidelines on the interplay between the AI Act and EU data protection law (under preparation).

([38]) See, by way of analogy, C-89/04, *Mediakabel,* EU:C:2005:348, para. 33; C-192/04, *Lagardère*, EU:C:2005:475, para. 31; C-306/05, *SGAE,* EU:C:2006:764, para. 37-40; C-527/15, *Stichting Brein*,

restricted to specific individuals belonging to a closed, private group (e.g. a small, closed group on an instant messaging app, or if the group is too small or insignificant). Examples of text not considered published include e.g. private, interpersonal correspondence (for professional purposes), or organisation-internal texts or communications (e.g. publications on internal corporate networks).

ii. **Informing the public:** The AI-generated or manipulated published text should aim to inform the public. This entails that the text should intend to communicate knowledge, opinions or facts. By contrast, short texts which do not materially communicate knowledge, opinions or facts, cannot be deemed to inform the public.

iii. **On matters of public interest:** The text should inform the public 'on matters of public interest'. In general, such matters should be understood to cover those relevant to society at large, whether at a local, national, Union or international level, and meriting public debate or scrutiny. In that regard, texts should be considered to address public interest matters if they cover topics on politics and democratic processes, public administration and services, the administration of justice and law enforcement, the protection of fundamental rights, public security, public health, environmental protection, consumer safety, and any economic, financial, political, scientific, or cultural development that may be relevant subject of public debate. Matters that may be considered to be of public interest can evolve over time and across contexts.

Examples of text published with the purpose of informing the public on matters of public interest under Article 50(4) AI Act Act which must be labelled as artificially generated or manipulated unless subject to human review/editorial control (see Section 6.2.3.):

- AI-generated summary of a human-authored article on a newspaper's website discussing a recent decision by a town council.

- AI-manipulated parts of a lifestyle-website article comparing the effects of various diets on a particular disease in middle-aged women.

- AI-manipulated corporate reports published on a listed company's website containing investor information.

- AI-generated message on a meteorological institute's social media profile warning citizens about stormy weather and related precautionary measures.

Examples of text which is not published with the purpose of informing the public on matters of public interest falling outside the scope of Article 50(4) AI Act:

- AI-generated fantasy novels.

- AI-manipulated text that is part of a company's advertisement or product descriptions (not including any claims related to e.g. health, consumer safety or sustainability).

- News summary generated by a chatbot that is only available to the user that prompted the chatbot.

- AI-manipulated text by a consultant for a client advice regarding measures to be taken for regulatory compliance with applicable legislation.

EU:C:2017:300, para. 43-44 and case-law cited; Case C-135/23, *GEMA*, EU:C:2024:526, para. 38 and case-law cited).

6.2.2. The disclosure obligation under Article 50(4), second subparagraph, AI Act

(132) Article 50(4), second subparagraph, AI Act obliges deployers of AI systems that generate or manipulate text which is published with the purpose of informing the public on matters of public interest to disclose that such text has been artificially generated or manipulated. Recital 134 AI Act clarifies that, similarly for deep fakes, deployers of text publications falling within the scope of the provision should clearly and distinguishably disclose that such content has been artificially created or manipulated by labelling the AI output accordingly and disclosing its artificial origin. As required for deep fakes, labelling or disclosure methods (including disclaimers) applied in accordance with Article 50(4), second subparagraph, AI Act should also be clear and perceivable by natural persons (e.g. visible or audible measures) without them needing to rely on any specific technical tools or performing dedicated actions. In light of Article 50(7) AI Act, deployers may adhere to a code of practice, assessed as adequate pursuant to that provision, to demonstrate compliance with this disclosure obligation For further information regarding the effects of adherence to such a Code and supervision of deployers not adhering to such a code, see Section 8.1 below.

6.2.3. Exception from the transparency obligation for text under human review or editorial control and editorial responsibility

(133) Article 50(4), second subparagraph, AI Act foresees an exception to the transparency obligation laid down in that provision where two cumulative conditions are met: (i) the AI generated or manipulated text must have undergone human review or editorial control and (ii) a legal or natural person must hold editorial responsibility for the publication. Where relevant, deployers may rely on relevant applicable professional or deontological standards to demonstrate compliance with those requirements as outlined below.

i. Text under human review or editorial control

(134) The first condition for the exception to Article 50(4), second subparagraph, AI Act to apply requires that the AI-generated or manipulated text has been subject to human review or editorial control. Human review refers to the deliberate examination of the substance of the content by one or more natural persons possessing relevant knowledge and professional judgement pertaining to the subject matter under scrutiny (e.g. academic peer review or professional validation chains). Fact-checking the accuracy of the content is a minimum requirement that should be part of that review. Editorial control refers to the control exercised in practice by a responsible editorial entity (e.g. an editor-in-chief) over the content having the authority to approve, alter or reject the substance of the text based on substantive grounds (incl. fact-checking of information and ensuring the trustworthiness of sources). This is without prejudice to existing review and editorial procedures and professional standards applicable to media service providers.

(135) Superficial, solely formal or procedural checks (e.g. spell-checking or grammatical correction), the mere existence of an editorial policy, automated review processes or cursory editorial approval without substantive engagement by the human reviewer or the editorial entity, cannot fulfil the conditions for human review or editorial control for the purposes of this exception.

(136) Where AI systems are used to modify, supplement, or reformulate content following editorial sign-off, the resulting content must be treated as AI-generated or manipulated

for the purposes of Article 50(4) AI Act. Any substantive AI intervention occurring after the human review or editorial control process has taken place will therefore cause the exception to become void.

(137) In light of Article 50(7) AI Act, deployers may adhere to a code of practice, assessed as adequate pursuant to that provision, to demonstrate compliance with the condition of human review or editorial control (especially by deployers that are not media service providers subject to existing editorial professional or deontological standards) laid down in that provision. For further information regarding the effects of adherence to such a code and supervision of deployers who do not adhere to such a code, see Section 8.1.

ii. Editorial responsibility

(138) The second condition for the exception to Article 50(4), second subparagraph, AI Act to apply requires a legal or natural person to hold editorial responsibility for the publication of the content. This entails that said person must hold the ultimate legal responsibility over the publication of the content, including the human review or editorial control (e.g. an individual, editorial board, or the publishing company). To ensure public accountability and trust, and in line with existing media professional standards, the identity and contact details of the legal person, the natural person or the function with editorial responsibility should be made publicly available on an easily findable location (if not yet otherwise available). This can happen online through e.g. a website's terms and conditions or other user-facing legal information. Offline, such information can be included in a publication's colophon or edition notice.

Examples of text subject to human review or editorial control with a legal or natural person holding editorial responsibility

- An AI-manipulated article in a newspaper or AI-generated summary of an article that has been subject to editorial control of the respective editor-in-chief with editorial responsibility held by the legal person that publishes the newspaper.

- An AI-manipulated academic blog which has undergone internal peer review and where the respective research centre managing the blog holds editorial responsibility.

- AI-generated public safety warnings approved by a public official before being distributed to citizens, under the responsibility of the relevant public agency for civil protection.

- AI-generated sustainability reports published on a listed company's website having undergone human review by professionals in relevant functions (e.g. compliance).

- AI-supported translation of a human-written article whereby the translation has undergone human review.

Examples of text that do not meet the required human review or editorial control to benefit from the exception:

- A website where AI-generated articles on Union policy are posted without any deliberate human review or editorial control.

- AI-generated articles that are reviewed and edited by another AI system and where a human editor performs a mere superficial, grammatical check before publication.

- An AI-generated, self-published book on climate change, made available on an e-commerce platform that has not undergone any review by a competent natural or legal person (nor by the platform).

6.2.4. Exception for law enforcement

(139) If the use of AI systems for generation of text published with the purpose of informing the public on matters of public interest without human review (e.g. an AI-generated police warning regarding a terrorist threat published on social media) is authorised by law to detect, prevent, investigate or prosecute criminal offences, deployers are exempted from the transparency obligation under Article 50(4) AI Act. The clarifications provided in points 46-48 above for a similar exception in Article 50(1) AI Act are also relevant for this provision.

6.2.5. Interplay with other Union legal acts

(140) For Union media law, editorial responsibility is a crucial concept defined in Article 2(8) of Regulation (EU) 2024/1083 (the European Media Freedom Act or EMFA) ([39]). In that context, 'editorial responsibility' means 'the exercise of effective control both over the selection of programmes or press publications and over their organisation, for the purposes of the provision of a media service, regardless of the existence of liability under national law for the service provided'. In certain situations, this definition may overlap with the notion of editorial responsibility used in Article 50(4), second subparagraph, AI Act, in particular where the deployers of the AI system also qualify as media service providers under the EMFA. Editorial responsibility within the meaning of Article 50(4) AI Act should therefore be interpreted in line with the existing media acquis, notwithstanding the fact that it remains a distinct concept that may also apply in broader contexts and to other deployers, so long as they assume editorial responsibility for the text publication. Media service providers may rely on their existing editorial processes and standards to benefit from the exception under Article 50(4), second paragraph, AI Act.

7. HORIZONTAL REQUIREMENTS APPLICABLE TO THE INFORMATION PROVIDED UNDER ARTICLE 50(5) AI ACT

(141) Pursuant to Article 50(5) AI Act, the information to be provided in accordance with Article 50(1) to (4) AI Act should be provided to the natural persons concerned in a clear and distinguishable manner, at the latest at the time of the first interaction or exposure. The information should also conform to the applicable accessibility requirements. Natural persons concerned in this context include the persons directly interacting with the AI system (Article 50(1) AI Act), active or passive users and other persons exposed to the AI-generated or manipulated synthetic content (with or without direct interaction with the AI system (Article 50(2) and (4) AI Act)), and persons subject to the use of emotion recognition or biometric categorisation systems (Article 50(3) AI Act).

([39]) Regulation (EU) 2024/1083 of the European Parliament and of the Council of 11 April 2024 establishing a common framework for media services in the internal market and amending Directive 2010/13/EU (European Media Freedom Act); OJ L, 2024/1083, 17.4.2024.

7.1. Information provision in a clear and distinguishable manner

(142) Information will be considered to be provided in a clear manner where it is noticeable, easy to understand by and accessible for the natural person concerned (including persons with disabilities). Information will be considered to be provided in a distinguishable manner where it is easy to identify as separate from other information and the environment in which the content is presented. It must be also easily understood by the audience, including by specific groups, such as children or persons with disabilities when they are part of the audience foreseeably exposed to the content or interacting with the system ([40]). To provide information in a clear and distinguishable manner, the relevant information may be *inter alia* presented as part of the interaction (under Article 50(1) AI Act) or the relevant content (under Article 50(2) and (4) AI Act). Information will not be considered to be provided in a clear and distinguishable manner where it can be easily overlooked or missed by natural persons under normal exposure or interaction conditions (e.g. only included as part of a manual or hidden under layers of menu options on an online interface, part of terms of use that are often not read by users).

7.2. First interaction or exposure

(143) Article 50(5) AI Act requires the information to be provided to the natural persons concerned at the latest at the time of the first interaction or exposure. The first interaction or exposure refers not only to the first natural person interacting with or exposed to (the output of) an AI system, but also any subsequent, first interaction with or exposure to (the output of) the AI system by any other natural person. As regards the first interaction with an AI-system, the information should, at least, be provided once at the start of an interactive session, while considering the specific contexts mentioned in point above. As regards AI-generated or manipulated content falling within the scope of Article 50(2) and (4) AI Act, the information obligation applies to each output of an AI system with respect to any natural person exposed to the content. It should be understood in the sense of any moment in time at which a natural person is reasonably likely to be exposed to the output of the AI system and perceive the disclosure, taking into account the specificities of disclosure in case of evidently creative, artistic and other works (see Section 6.1.3.). Moreover, deployers are allowed to inform natural persons earlier than the first actual interaction or exposure (e.g. at the beginning of content featuring deep fakes). However, if it is reasonably foreseeable that persons may not perceive content from its beginning, then only disclosure at the beginning of content does not adequately inform those persons and should be complemented with disclosure at later moments, where possible. This may be implemented as appropriate, taking into account the objective of the transparency obligations to ensure an effective provision of the information and meaningful user awareness in increasingly immersive and realistic AI environments where users have

([40]) When children are part of the target audience of the content those notifications must be: (i) child-friendly, age-appropriate, easy-to-understand and easily accessible to all children, including those with disabilities and/or additional accessibility needs; (ii) presented clearly in a way that is easy to understand and is as simple and succinct as possible; (iii) presented in ways that are easy to review and that provide for immediate and intuitive access, at the points at which they become relevant; (iv) presented in the official language(s) of the Member State the service is provided in; (v) engaging for children. This may also include the use of graphics, videos, and/or characters or other techniques; (vi) given to children gradually and overtime to maximise retention by the user.

increasing difficulties to distinguish the origin and the truthfulness of the content and the interaction.

Examples of information provision at the first interaction or exposure under Article 50(5) AI Act

- First interaction: when launching a conversation with a chatbot or starting physical interaction with an AI system.

- First exposure: at the start of a video featuring a deep fake, at the start of an AI-generated or manipulated text publication on matters of public interest, or when encountering AI-manipulated deep fakes when scrolling on social media.

7.3. Compliance with applicable accessibility requirements

(144) Providers and deployers of AI systems falling within scope of the various transparency obligations listed in Article 50(1)-(4) AI Act must be aware of the applicable accessibility requirements (e.g. under Directive 2016/2102 ([41]) and Directive 2019/882 ([42])) and assess whether their product, service or content needs to comply with those rules. In such a case, they must ensure accessibility of the information conveyed in line with Article 50 AI Act in accordance with the applicable accessibility requirements. Article 50 AI Act does not impose distinct or additional accessibility requirements.

8. ENFORCEMENT OF ARTICLE 50 AI ACT

(145) This section aims to clarify: (i) the effects of adhering to a code of practice as regards compliance with the transparency obligations for AI-generated content; (ii) the role of the competent supervisory authorities in the enforcement of Article 50 AI Act; (iii) the applicable penalties for infringements of those obligations; and (iv) the moment at which the obligations in Article 50 AI Act enter into application.

8.1. Effects of adhering to a code of practice assessed as adequate

(146) Providers and deployers of AI systems falling within the scope of Article 50(2) and (4) AI Act may demonstrate compliance with their respective transparency obligations for AI-generated content under those provisions by adhering to a code of practice that is assessed as adequate pursuant to Article 50(7) AI Act ([43]). Such a Code does not replace the AI Act or these Guidelines. It complements them by providing the only Union-wide recognised practical framework for providers and deployers of generative AI systems to

([41]) Directive (EU) 2016/2102 of the European Parliament and of the Council of 26 October 2016 on the accessibility of the websites and mobile applications of public sector bodies (Text with EEA relevance), *OJ L 327, 2.12.2016, pp. 1–15.*

([42]) Directive (EU) 2019/882 of the European Parliament and of the Council of 17 April 2019 on the accessibility requirements for products and services (Text with EEA relevance), PE/81/2018/REV/1**,** *OJ L 151, 7.6.2019, pp. 70–115.*

([43]) See relevant information about the Code of Practice on Transparency of AI-Generated Content | Shaping Europe's digital future published on 10 June 2026.

demonstrate compliance with those obligations, regardless of their place of establishment, operation, and competent market surveillance authorities.

(147) While providers and deployers may also demonstrate compliance with those obligations through adequate alternative means, adherence to a code of practice that is deemed adequate is a straightforward, predictable, and legally certain way of demonstrating compliance. For signatories of a code of practice that is assessed as adequate pursuant to Article 50(7) AI Act ([44]), the Commission and the competent market surveillance authorities will focus their supervisory activities on assessing whether those signatories have adhered to the code of practice and implemented the measures contained therein. As adhering providers and deployers will be transparent about the measures they implement to comply with Article 50(2), (4) and (5) AI Act, they will benefit from increased trust from the Commission, the other competent market surveillance authorities, and other stakeholders (including the general public). Any opt-out from sections by signatories of a code of practice that is deemed adequate will result in those providers and deployers losing the benefit of facilitating the demonstration of compliance in that respect.

(148) Providers and deployers that are not signatories to a code of practice that is deemed adequate pursuant to Article 50(7) AI Act are expected to demonstrate how they have complied with their obligations under Article 50(2), (4) and (5) AI Act through other adequate means. Furthermore, such providers and deployers are expected to explain how the measures they implement ensure compliance with their obligations under the AI Act. For instance, they should carry out a gap analysis that compares the measures they have implemented with the measures set out by a code of practice that is assessed as adequate. Since competent authorities (including the AI Office) will have less understanding of how providers and deployers that are not signatories to a code of practice ensure compliance with their obligations under Article 50(2), (4) and (5) AI Act, they will likely need more detailed information when monitoring for compliance. For example, providers not adhering to a code of practice may also be subject to a larger number of requests for information and requests for access to assess the effectiveness, interoperability, robustness and reliability of the technical solutions implemented to comply with Article 50(2) AI Act. Deployers may also be subject to such requests with regard to their labelling practices under Article 50(4) AI Act.

(149) Competent authorities may take commitments implemented in line with a code of practice that is deemed adequate into account as a mitigating factor when fixing the amount of fines, depending on the specific circumstances ([45]).

(150) If a code of practice is not deemed adequate pursuant to Article 50(7) AI Act, the Commission may adopt an implementing act specifying common rules for the implementation of the obligations of Article 50(2), (4) and (5) AI Act, which would be applicable to all relevant providers and deployers of AI systems falling within the scope of those provisions ([46]).

([44]) See Commission Opinion on the assessment of the Code of Practice on Transparency of AI-generated content | Shaping Europe's digital future

([45]) Article 99(7)(e) AI Act.

([46]) Article 50(7) AI Act.

8.2. Market Surveillance Authorities

(151) Market surveillance authorities designated by the Member States, the AI Office ([47]), and the European Data Protection Supervisor ([48]) are responsible for supervising and enforcing the rules for AI systems falling within their competence, including the transparency obligations laid down in Article 50 AI Act. Such enforcement takes place within the system of market surveillance and compliance of products established by Regulation (EU) 2019/1020 and the AI Act. The enforcement powers of market surveillance authorities in relation to AI systems are laid down in those regulations. Those authorities can take enforcement actions in relation to the obligations listed in Article 50 AI Act on their own initiative or following a complaint, which every affected person or any other natural or legal person having grounds to consider such violations has the right to lodge ([49]). The responsibilities of the market surveillance authorities under the AI Act do not affect the powers and tasks of other supervisory authorities or fundamental rights authorities under their own legal framework in situations where both Article 50 AI Act and that framework apply (e.g. data protection, consumer protection).

8.3. Penalties

(152) Provider and deployers that do not comply with the applicable transparency obligations laid down in Article 50 AI Act may be fined up to EUR 15 000 000 or, if the offender is an undertaking, up to 3% of its total worldwide annual turnover for the preceding financial year, whichever is higher. EU institutions, bodies and agencies that violate the transparency obligations may be subject to administrative fines of up to EUR 750 000. In the case of small and medium-sized enterprises (SMEs), including start-ups, each fine shall be up to the above percentages or amount, whichever is lower. The fines shall be effective, proportionate and dissuasive and take into account the interests of SMEs and small mid-cap enterprises (SMCs), including start-ups, and their economic viability. When deciding whether to impose an administrative fine and when deciding on the amount of the administrative fine in each individual case, the competent authorities should assess all relevant circumstances of the specific situation and, as appropriate, a range of factors that affect the seriousness of the infringement (e.g. nature, gravity and duration of the infringement and of its consequences, negligent or intentional character, other aggravating or mitigating circumstances, the degree of cooperation with market surveillance authorities etc.).

8.4. Entry into application

(153) According to Article 113 AI Act, Article 50 AI Act will apply as from 2 August 2026. This requires all in scope AI systems placed on the market or put into service in the Union to be compliant with that provision on that date, regardless of their date of placement on

([47]) As the market surveillance authority for AI systems built on GPAI models provided by the same provider (Article 75(1) AI Act). This covers the obligation applicable to providers under Article 50(1) and (2) AI Act. regarding Article 50(3) and (4) AI Act, the AI Office will only be competent if the respective GPAI model and system provider, is also a deployer of the system.

([48]) As the market surveillance authority for the EU institutions, agencies and bodies.

([49]) Article 85 AI Act.

the market or putting into service ([50]). Regulation amending the AI Act (the AI Omnibus), which has been recently adopted by the Union legislature, envisages a targeted grandfathering rule only with regard to the marking and detection obligations under Article 50(2) AI Act for generative AI systems placed on the market or put into service before 2 August 2026. It gives providers of those existing systems a transitional period to bring their systems in conformity by 2 December 2026. Systems that are partly interactive and partly generative may benefit from this transitional period only with regard to the marking obligation under Article 50(2) AI Act, while compliance with the disclosure obligation for AI systems directly interacting with natural persons must be ensured as of 2 August 2026.

(154) AI-generated or manipulated outputs falling within the scope of Article 50(2) AI Act and deep fakes within scope of Article 50(4), first subparagraph, AI Act, which have been generated or manipulated before 2 August 2026 do not need to be marked or labelled retroactively. The same applies to texts on matters of public interest falling within scope of Article 50(4), second subparagraph, AI Act that have been AI-generated or manipulated and published before 2 August 2026. Conversely, if texts that have been generated or manipulated before 2 August 2026 are published on or after that date, they need to be labelled. Considering the objective of transparency and increased trust and integrity of the information ecosystem pursued by the AI Act, deployers of AI systems and other actors who are in the possession of or disseminate pre-existing unlabelled deep fakes are, nonetheless, encouraged to do so without being expected to engage in disproportionate efforts to that end (such as auditing of pre-existing content databases or modifying already printed product packaging).

9. REVIEW AND UPDATE OF THE GUIDELINES

(155) These Guidelines constitute a first interpretation with practical examples of the transparency obligations laid down in Article 50 AI Act. The Commission will review these Guidelines as soon as necessary in view of practical experience gained in the implementation of the transparency obligations and the pace of technological, societal, and regulatory developments in this area. This also includes any relevant experience from market surveillance enforcement actions and interpretations on Article 50 AI Act given by the CJEU. During such a review, the Commission may decide to withdraw or amend these Guidelines. The Commission encourages providers and deployers of AI systems, national market surveillance authorities through the AI Board, the AI Advisory forum, the research community, and civil society organisations to contribute to this process by responding to future calls for public consultation.

([50]) The special grandfathering rule in Article 111(2) AI Act applies to high-risk AI systems placed on the market before the date of the application of the high-risk rules only with respect of the compliance of those systems with the requirements and obligations for high-risk AI systems. If the high-risk AI system is also subject to one or more transparency obligations under Article 50 AI Act, this special grandfathering rule does not apply since the application of the different obligations for high-risk and Article 50 transparency are cumulative and the grandfathering is justified and limited only with regard to the compliance with the high-risk provisions.