EUROPEAN COMMISSION
DIRECTORATE-GENERAL FOR COMMUNICATIONS
NETWORKS, CONTENT AND TECHNOLOGY

Artificial Intelligence Office
Artificial Intelligence Regulation and Compliance

Draft 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

Disclaimer: This is a draft working document and does not prejudge the final decision that the Commission may take on the guidelines. The draft is published for stakeholder consultation to provide input to the Commission before a finalised version of the guidelines are adopted by the Commission.

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 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 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 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.

(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)	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.	If artificial interaction is obvious, or 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.

([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.

Art. 50(2)	Generating or manipulating synthetic image, video, audio or text content	Providers must ensure the AI system's outputs are marked in a machine-readable format and detectable as artificially generated or manipulated with technical solutions that are effective, interoperable, robust and reliable.	If the AI system performs an assistive function for standard editing or does not substantially alter the input data or the semantics thereof, or the system is authorised by law to detect, prevent, investigate or prosecute criminal offences.
Art. 50(3)	Emotion recognition or biometric categorisation	Deployers must inform the natural persons exposed to the system of its operation.	If the system is permitted by law to detect, prevent or investigate criminal offences.
Art. 50(4)	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.	If the system is authorised by law to detect, prevent, investigate or prosecute criminal offence, or if the text publication has undergone human review or editorial control and is subject to editorial responsibility. Special disclosure regime applies to 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 can apply cumulatively to (the output of) a single AI system, engaging possibly the responsibility of different actors (providers or deployers).

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.

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, irrespective of whether those providers are established or located within the Union or in a

third country[3]. Providers of AI systems 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]. Providers must ensure that their AI systems meet all relevant transparency obligations laid down in Articles 50(1), (2) and (5) AI Act before placing those systems on the market or putting them into service.

For example, the provider of an interactive AI system falling within scope of Article 50(1) AI Act is the provider that puts the system into service in the Union under its trademark, regardless of its place of establishment. Similarly, a third country provider of a generative AI system may be subject to the obligation laid down in Article 50(2) AI Act, even if the system is only marketed outside the EU, if the system's outputs are intended to be used in the Union.

(11) **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]. '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 output). Deployers fall within the scope of the 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]. Where the deployer of an AI system is a legal person under whose authority the system is used (e.g. a movie company or a newspaper publisher), the individual employees that act within the procedures/instructions and under the control of that legal person (e.g. digital animators or journalists) should not be considered as separate deployers. A legal person remains a deployer even if it involves third parties (e.g. contractors, freelancers or external staff) 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 AI systems to support their written coverage of current events posted on a globally accessible website, is a deployer falling within the scope of the AI Act. Similarly, a third country advertising company 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.

(12) By contrast, actors whose role is limited to disseminating or transmitting AI-generated or manipulated content created by third parties (including online platforms), or who receive or are exposed to AI-generated or manipulated content without directly having authority over the use of the AI system, are not deployers within the meaning of the AI Act. Those actors may nevertheless still play an important role in the value chain and are encouraged to preserve the marking and labelling of the content implemented pursuant to the AI Act.

(13) Operators (e.g. providers and deployers) may fulfil more than one role concurrently in relation to an AI system.

(14) Since safeguarding trust and integrity of the information ecosystem is a shared responsibility, other actors, acting in their professional capacity and disseminating

([3]) Article 2(1) (a) and Article 3(3), (9) and (11) AI Act. For further guidance on the concepts of placing on the market, putting into service and use, please consult 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) (a) AI Act.

([5]) Article 3(4) AI Act.

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

contentacross the value chain (even if not directly engaged under Article 50 AI Act) are also encouraged to take appropriate measures so that the natural persons exposed to the content will be effectively informed about 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).

2.4. Exclusions from the scope of the AI Act

(15) 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[7]. This section will accordingly address only those aspects that are relevant to the transparency obligations in Article 50 AI Act.

2.4.1. Purely personal non-professional activity

(16) According to Article 2(10) AI Act, the regulation 'does not apply to obligations of deployers who are natural persons using systems in the course of a purely personal, non-professional activity'[8]. 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. 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 deployers' transparency obligations laid down in Article 50(3) and (4) AI Act.

(17) 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. The exclusion should therefore not encompass criminal activities since these should not be considered purely personal, even if no economic benefit is sought or attained. Similarly, an AI-generated or manipulated deep fake that is made publicly available by a person, and which may have an impact on public discourse on matters of public interest (e.g. due to its political or economic content), should not be considered a purely personal non-professional activity either[9].

For example, individuals using an AI system to create Christmas cards to be sent to relatives, featuring deep fakes of the members of their household, fall under the exclusion of Article 2(10) AI Act so that those deep fakes would not have to be labelled as such in accordance with Article 50(4) AI Act. By contrast, if an individual generates a deep fake of the mayor of their town to criticise certain local policy decisions and publicly shares that deep fake on social media, then such a deep fake cannot benefit from the exclusion and should be labelled as it is not a purely personal activity.

(18) 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

([7]) 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.

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

([9]) The concept of deep fake is clarified in part 6.1.1. What matters constitute 'matters of public interest' is clarified in part 6.2.1.

system on the market or putting it into service and of deployers using the system for non-personal professional use.

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 for purely personal, non-professional will not be covered by the deployer obligation laid down Article 50(4) AI Act.

2.4.2. Research & Development

(19) 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. In the context of Article 50 AI Act, this exclusion does not only cover 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 also used outside of the scientific research context[10], the relevant transparency obligations in Article 50 AI Act would still need to be complied with.

For example, if researchers develop 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 interactive AI systems used is also put into service outside of the research setting, the relevant transparency obligations in Article 50 AI Act would need to be complied with.

(20) 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. 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[11].

2.4.3. AI systems released under free and open-source licences

(21) According to Article 2(12) AI Act, AI systems released under free and open-source licences are outside the scope only if they do not fall under the prohibitions, high-risk or the transparency obligations of Article 50 AI Act[12]. 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.

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

(22) Recital 137 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

(10) The Commission is preparing guidelines on research exemption that will provide further clarifications.
(11) Recital 25 AI Act.
(12) Recital 102 AI Act describes 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'.

system or its output is lawful under the AI Act. In particular, AI systems mentioned in Article 50 AI Act may fall under Article 5 AI Act and be prohibited (e.g., such as the use of emotion recognition in the area of workplace or education institutions)[13]. Furthermore, AI systems falling within the scope of Article 50 AI Act may also be classified as high-risk[14] and be subject to the relevant requirements and obligations for such systems. 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.

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

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

(24) 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 could also be implemented at the model level (see also Section 4). Furthermore, while outside the scope of Article 50 AI Act, other providers of GPAI 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, to facilitate compliance by downstream AI system providers with their obligations in Article 50(1) and (2) AI Act[17]. For GPAI models presenting systemic risk, model-level safety measures for identification of AI systems 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

(25) 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 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 that those natural persons over rely on such systems, and to allow those natural persons to calibrate their trust in the content and the interactions accordingly.

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

(26) 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

(13) For further information on Article 5 AI Act, please consult the Commission Guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689 (AI Act), C(2025) 5052.
(14) See Article 6 AI Act.
(15) See Article 3(66) AI Act defines general-purpose AI 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.

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

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

i. **An AI system**: The systems must fulfil the elements of the AI system definition[18], excluding 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 possibility of a bidirectional exchange of information or actions between natural persons and AI systems and 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 can 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 or take place over a certain period time. Systems that only passively collect data and are not capable of engaging in an exchange with natural persons are not considered to be intended to interact with natural persons (e.g. automated facial-recognition access controls).

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 the 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).

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) or virtual environment without any direct contact with natural persons are excluded from the scope of the obligation.

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

(28) AI agents are covered by Article 50(1) AI Act if they are designed to interact with the persons instructing them and potentially with other natural persons in the execution of tasks. However, it may not always be possible for the provider to identify individual instances of interaction of an AI agent with other natural persons. Where the provider cannot reliably determine whether the AI agent will interact with a natural person, the agent should be instructed to disclose itself as such in every situation where it is likely that the agent may interact with a natural person.

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, e-commerce, finance, healthcare, education etc.), (humanoid) robots/cobots, AI companions; robotic companion pets; AI avatars (e.g. in virtual reality environments), bots on social networks and media, coding agents and other agentic AI systems.

Examples 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, algorithmic recommender systems, spam filters, automated translation or transcription tools, authentication/biometric recognition systems, backend decision-support systems where the user only sees the AI output without possibility for direct interaction with the system; predictive maintenance and optimisation systems in factories and industrial applications.

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

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

(30) While the mechanism to inform should be embedded in the design and development of the system and in the way it 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 a natural person, as required in Article 50(5) AI Act (see Section 7 below). Providers are responsible for ensuring that natural persons are effectively informed throughout the lifecycle of the AI system, including after its placement on the market and its putting into service.

(31) **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 elderly) or disability are taken into account to the extent that the AI system is intended to interact with natural persons belonging to those groups[19]. The information must also be provided in accessible

([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 notification measures.

formats for persons with disabilities, as specified in Article 50(5) and Recital 132 AI Act (see Section 7). When 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 and/or additional accessibility needs.

(32) **Substance of the notification:** Article 50(1) AI Act requires providers of interactive AI systems to explicitly inform all natural persons interacting with such AI systems 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 to also ensure the full understanding of those persons.

(33) The notifications should provide the information on the artificial origin of the interaction in ways that are appropriate to the context in which the interaction takes place, so that the persons concerned are properly and effectively informed, avoiding any risks of deception before and throughout the interaction. 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), and AI identifiers or credentials (e.g. AI agents that disclose their AI identity, including as appropriate in a verifiable manner[20]).

(34) In terms of the **format of the notifications**, the use of multimodal and accessible disclosure approaches is strongly recommended[21], potentially combining different techniques to reinforce user understanding, where appropriate and tailored to the target and potential audience groups. Examples of appropriate techniques may include:

- **Textual disclosure** (UI-based): Prominent, plain-language labels or banners (e.g. “You are interacting with an AI system”), first-turn greetings in chatbots, and persistent badges visible throughout the interaction. Furthermore, it is recommended to position accompanying disclosures close to the interaction interface (e.g. near the input/output field) and using 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. Distinct audio cues (e.g. tones or earcons) may support recognition, particularly for visually impaired users, but are not considered sufficient by themselves.
- **Visual/graphical cues:** Use of persistent icons, watermarks, coloured frames or recognisable “AI” symbols to complement textual disclosures. The use of standardised visual indicators across services of a provider is recommended, as they can significantly reduce cognitive burden and increase recognisability.

([20]) For example, electronic attestations of attributes as established under Regulation (EU) No 910/2014 and available in the EU Digital Identity Wallets established under the same Regulation and the proposed European Business Wallets, can provide a 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.

([21]) See summary report on Stakeholder Feedback received from the public consultation November 2025.

- **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.

(35) Certain techniques, when used alone, are not considered appropriate 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), which are not perceivable by users at the point of interaction and therefore cannot fulfil the transparency obligation in Article 50(1) AI Act;
- Unclear or ambiguous signals (e.g. generic references to "assistant"), or human-like representations that may mislead users;
- 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 non-human origin.

(36) Caution may be needed regarding overly intrusive disclosure techniques that may undermine the effectiveness of providing the information and disrupt user experience and lead to habituation effects (so-called "banner blindness").

(37) At the same time, certain approaches may prove insufficient in terms of timing and continuity depending on the context and the use. In particular, one-time disclosures at the beginning of an interaction may not be considered adequate in the context of sustained or evolving interactions, especially in sensitive contexts (e.g. where users may express or experience emotional distress or vulnerability) or where there is an increased risk of users being misled or forming emotional attachments (e.g. AI companions). In such cases, periodic reminders and context-aware disclosures may be necessary to ensure continued user awareness. Providers are also encouraged to ensure disclosure in situations where the AI system is asked questions relating to its nature or to the origin of the interaction by the interacting natural person, or where it can be reasonably assumed from the exchanges with the natural person that the person is likely to be misled or confused about the AI origin.

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

(38) 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

(39) 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.

(40) 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.

(41) Conceiving such 'an average' natural person comprises two steps. First, providers should consider the target audience with whom the AI system is intended 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 interacting with natural persons present in those spaces.
- Providers should consider the potential diverse composition of the target audience in line with Recital 132 AI Act. Where persons with disabilities, elderly people or minors are 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 information, 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 audience and relevant social, cultural, and linguistic factors.

(42) **Obviousness**: Upon establishing the hypothetical reasonably well-informed, observant and circumspect natural person, a provider should evaluate whether it would be obvious to such a person if they are directly interacting with an AI system. Several factors may be relevant and considered as part of this assessment:

- 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 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.

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

- AI-powered code assistance 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 readily recognise that suggestions are generated by an AI system.

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

- Interactions with AI-enabled Non-Playable Characters (NPCs) in a videogame

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, elderly individuals, 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. recommendations or other generated content) they may perceive as neutral or human-generated.

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

(43) 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.

(44) Authorized 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 AI system, the purposes and circumstances of permitted use and provide for appropriate safeguards to protect the rights and freedoms of third parties.

(45) 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). 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 EU or national public authorities or even private actors, such as security companies or financial institutions, 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.

(46) 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 remain subject to the transparency obligation under Article 50(1) AI Act, so natural persons must be informed they are interacting with those systems.

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. These systems guide users through structured

questions, collect relevant information, and triage reports before forwarding them to human officers for assessment.

- 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

(47) The transparency obligation under Article 50(1) AI Act applies without prejudice to existing Union consumer protection, data protection and digital 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)[22], traders must ensure that consumers are not misled about the main characteristics of a product or service. Directive 2011/83/EC (the Consumer Rights Directive)[23] 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 companion application, or a virtual coaching service), the AI functionality may qualify as an essential characteristic that must be disclosed prior to purchase. The information obligations under EU consumer protection law apply irrespective of whether the interaction is considered "obvious" under Article 50(1) AI Act.

(48) The DSA also includes transparency obligations relevant to AI-mediated interactions. For example, providers of online platforms must clearly inform users about their content moderation, recommender system, and advertising practices. These obligations are complementary to the transparency obligations laid down in Article 50(1) AI Act, since they concern the provision of a different type of information, they are not AI-specific, and they require disclosure of relevant information for systems that in most cases do not interact directly with the users.

(49) Article 50(1) AI Act fulfils a different objective than the information obligations towards data subjects under EU data protection law and, as such, does not affect those information obligations[24].

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

(50) 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 system outputs. The objective is to enable natural persons to distinguish AI-generated or manipulated content from other content (for example content

([22]) Directive 2005/29/EC of the European Parliament and of the Council as amended and in force.
([23]) Directive 2011/83/EU as amended and in force.
([24]) 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).

created by humans) and to verify its origin, thus also contributing to increased integrity and trust in the information ecosystem.

(51) 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 in ensuring the integrity of the information ecosystem, such as competent market surveillance authorities and other competent authorities, independent researchers, civil society organisations, media organisations, trusted flaggers, fundamental rights defenders, providers of online platforms, etc.

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

(52) 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: (a) they perform an assistive function for standard editing; (b) they do not substantially alter the input data provided by the deployer or the semantics thereof, or (c) they are authorised by law to detect, prevent, investigate or prosecute criminal offences.

4.1.1. AI systems generating or manipulating synthetic content

(53) Article 50(2) AI Act applies to AI systems, including general-purpose AI systems, generating synthetic audio, image, video or text content. This covers all AI systems as defined under Article 3(1) AI Act. The fact that those AI systems generate content and might be referred to as ‘generative AI systems’ does not distinguish them from any other AI system, since all AI systems generate outputs. Article 50(2) AI Act also covers AI systems used to manipulate content.

(54) Content generation refers to the creation 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 applies to generative AI systems designed with a narrow intended purpose to produce specific outputs (such as those generating or manipulating medical imagesor generating evaluations of learning outcomes resulting in text-based decisions or recommendations). 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.

(55) By contrast, content manipulation refers to already existing (not necessarily synthetic) content 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 under 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-generated 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

(56) 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:

- 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 refers to static spatial representations encoding visual information across one moment in time 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.

(57) AI systems generating or manipulating 3-D images, videos and 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 time-based sequence of images, optionally synchronized with audio, that represents visual change over time, it can be regarded as 'video' content.

(58) Digital twins that provide a virtual replica of persons, physical objects or systems and integrate multiple data types also fall within the scope of Article 50(2) AI Act, unless they qualify as 'industrial applications' as explained in point (79) below.

(59) The rules for generative AI systems within Article 50(2) AI Act also apply to agentic AI systems that pursue goals by, for example, planning, making decisions, or taking actions while interacting with physical or virtual environments. Any such action, e.g. a web request or browser action, that is not intended to be directly perceived by natural persons does not qualify as AI-generated content within the scope of Article 50(2) AI Act. If an AI agent takes an action that is perceptible by natural persons in the form of audio, image, video or text, those outputs fall within the scope of Article 50(2) AI Act and must be marked and detectable as described below.

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

(60) Certain outputs of AI systems do not qualify as synthetic content under Article 50(2) AI Act, since they merely present already existing content or their output is not capable of being perceived by humans or the risks of misinformation and manipulation at scale, fraud, impersonation and consumer deception are insignificantly low.

(61) 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 recommender systems that only select or rank existing content, based on user preferences or activity, without creating anything new or manipulating the content itself.

(62) 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 (e.g. by robots or other AI-enabled sensors) and/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; e.g., consumption recorded by AI-enabled smart meters, grid frequency and voltage measurements; recording of GPS location data from vehicles etc.

(63) 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 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. For example, content falling outside the scope of Article 50(2) includes:

(64) Outputs generated in the form of a short sequence of numbers, symbols or letters;
- Source code;
- Outputs of an AI system intended to be exclusively communicated from machine to machine and processed automatically without any exposure to humans;
- Outputs that are only used in closed loop industrial and product development environments, for example for film production, do not fall within the scope of Article 50(2) unless they are the final output provided it constitutes AI-generated or manipulated text, audio, image or video content.

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

(65) 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. For every marking solution deployed, providers should provide corresponding means for its detection (as outlined below) to enable natural persons exposed to the content and other relevant actors to identify and distinguish it 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.

(66) 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 ensuring their detectability by providing means for their detection) will not suffice[25].

([25]) Recital 133 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 refers to the obligation regarding detection.

4.2.1. The marking element

(67) The scope of the marking obligation is limited to implementing machine-readable marks. Perceptible marks and labels are not excluded as a complementary measure, where appropriate, and to facilitate the compliance of deployers with their obligation to label deep fakes pursuant to Article 50(4) AI Act.

(68) To comply with the marking element, 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, reliability, robustness and interoperability as required by the second sentence of Article 50(2) AI Act (see Section 4.2.3. below).

(69) 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, 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) and be conducive in enabling natural persons to distinguish AI-generated or manipulated content.

(70) Providers may implement the marking solution for the outputs at different stages of the value chain (e.g. at the level of the AI system or at the level of the underlying AI model) and rely on the marking solution implemented by an upstream model providers or third party providing the solution (e.g. as an open standard or a specialised service), without prejudice to their responsibility to demonstrate compliance with Article 50(2) AI Act.

4.2.2. The detection element

(71) 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. This means that the provider is obliged to make the means of detection available to the persons potentially exposed to the content and that – pursuant to Article 50(5) AI Act – such detection methods can produce human-readable results whether the content has been AI-generated or manipulated.

(72) 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. The provider must either provide their own detection solution or rely on a third-party or a publicly available detection solution as long as they ensure interoperability (see section 4.2.3.).

(73) Article 50(5) AI Act further specifies that 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.

4.2.3. Compliance with the requirements for technical solution(s): effective, interoperable, robust and reliable

(74) The second sentence of Article 50(2) AI Act provides that the technical solution(s) 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 the provider 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 for a variety of content generated or manipulated by the AI system.
- **Robustness** refers to the capability of the technical solution to maintain intended performance levels 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 deployed in different AI systems.

(75) Providers must ensure that the combination of the technical solutions for marking and detection holistically meet all 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 as far 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.

(76) '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, 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.

(77) 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. '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, which is based on the relevant consolidated findings of science, technology and experience and is accepted as good practice in technology. The state-of-the-art does not necessarily imply the latest scientific research still in an experimental stage or with insufficient technological maturity.

(78) Since under the current state-of-the-art there is no single technique for marking and detection that meets all four requirements at the same time to the legally required degree, a technical solution is required to combine different marking techniques to fulfil those

requirements. The specific solutions are however fast evolving, which requires alignment with the state of the art that may also evolve over time. Providers may therefore in the future be able to demonstrate compliance also through a technical solution utilising a single marking technique and detection solution if they can demonstrate they fulfil the four quality criteria simultaneously to the lawfully required degree representing the state of the art.

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

(80) There may therefore be narrowly defined cases where a technical solution based on one marking technique is sufficient for the purpose of complying with Article 50(2) AI Act, especially with regard to the inherent lower risks of deception, manipulation or negative effects on the information ecosystem. This may be the case when a generative AI system is embedded in physical products generating outputs in a technically controlled and closed environment that is mainly instructive in nature without the output leaving the product and not being exported, for example an AI system embedded in navigation systems in vehicles.

(81) There may also be limited cases of "industrial AI applications" or "business to business applications" where, due to a proportionate implementation and limited benefits for the transparency objectives, no marking and detection of AI generated 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 by a limited pre-defined number of natural persons acting in a professional capacity within the organisation of the provider/deployer, and it is not intended to be shared outside the company or to be usable or verifiable by external persons, with appropriate safeguards in place to avoid reasonably foreseeable misuse.

(82) Real-time content generation that is ephemeral and consumed at the moment without being stored or disseminated further (e.g. in video games) may also be exempted where the persons are aware that the content is AI-generated or manipulated and the content is consumed immediately without the AI outputs being recorded and enabling further verification.

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

(83) Article 50(2) AI Act provides three explicit exceptions from the transparency obligation regarding AI-generated or manipulated content examined below.

(84) 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.

(85) 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, 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.

(86) 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, format conversions, technical compression, noise reduction; minor cropping; minor colour adjustments or corrections; limited lightening or darkening; limited sharpening; removal of dust spots caused by a dirty lens or sensor; removal of red-eye caused by flash photography; rotating an image; rescaling of a video clip; limited video stabilisation; minor adjustments to playback speed, minor corrections to level the horizon of an image.

- 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 translations and summaries of text, adding objects or information not present in the original image or video; deleting or obscuring backgrounds, objects or other information visible in the original file; pixelation or blurring of faces; altering the body shape or the skin colour of a person; extreme lightening, darkening, colour and contrast adjustments using any editing software (e.g., making a grey sky blue, a blue sky orange, etc.); creating an extreme silhouette from a correctly exposed image or video file; converting a black & white to colour image or video; creation of composite images or video clips, any other substantial alteration of the content.

(87) Finally, if a generative AI system is authorised by law to generate or manipulate synthetic content to detect, prevent, investigate or prosecute criminal offences, it will be exempted from the marking and detection requirements under Article 50(2) AI Act. The clarifications provided for this exception in points 47-48 above are also relevant in this case.

4.4. Interplay with other Union legal acts

(88) The marking and detection obligation 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 falling within that provision must be therefore compliant with applicable EU data protection law, including data protection principles and obligations, such as data protection by design and by default, data minimisation, limited storage period, security and confidentiality of the information, etc.

(89) 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[26]. This applies in particular as regards the obligations of providers of very large online platforms ('VLOPs') or very large online search engines ('VLOSEs') to 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 misinformation and disinformation.

(90) The obligations to mark and label content generated 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[27]. 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 the content.

(91) For example, if a labelled deep fake image is flagged as potentially child pornography or infringing on trademark or copyright law, then the application of a label does not affect the potential illegality of the content. That assessment should be conducted solely on the basis of applicable law (e.g. criminal law, trademark or copyright rules). Similarly, unmarked deep fakes or other unmarked AI-generated content may be considered illegal content within the meaning of Article 3(h) DSA[28] where the content does not comply with the transparency obligations in Article 50 AI Act.

(92) Providers of online platforms, search engines and other 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.

([26]) Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act) (Text with EEA relevance), PE/30/2022/REV/1, OJ L 277, 27.10.2022, pp. 1–102.

([27]) See also recital 11 AI Act.

([28]) Article 3(h) DSA defines 'illegal content' as any information that, in itself or in relation to an activity, including the sale of products or the provision of services, is not in compliance with Union law or the law of any Member State which is in compliance with Union law, irrespective of the precise subject matter or nature of that law.

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

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

(94) Recital 132 AI Act explains that the purpose of Article 50(3) AI Act 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 the privacy of the persons concerned. The obligation therefore applies regardless of whether the persons are exposed to such systems in real-time or they 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

(95) 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 (under preparation), including the practical examples set out therein, are equally relevant in the context of Article 50(3) AI Act.

(96) 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

(97) 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 (under preparation), 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.

(98) Importantly, Article 50(3) AI Act applies to any biometric categorisation systems, regardless of whether they fall in scope of high-risk AI systems under the AI Act.

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

(99) **The scope of the information to be provided**: Deployers are required to inform persons concerned that they are exposed to an emotion recognition system or to a biometric categorisation system, respectively, which is being operated/used. The AI Act does not require including the information about the reasons for the system's operation.

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

(101) **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 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), and the possible existence of a relationship with the addressees (e.g., if there is an existing communication channel with the individual). Depending on the use case, the information could be provided in writing, by standardised icons (also when presented electronically), orally, or by using combinations of such ways.

Examples:
- 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.

(102) **The timeline for providing the information**: The information must be provided at the latest at the time of the first 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. Out of scope

(103) The 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 47-48 above are also relevant in this case.

5.3. Interplay with other Union legal acts

(104) Deployers need to comply with the information requirement in Article 50(3) AI Act, in addition to any applicable information requirements under EU data protection law[29]. In certain cases, deployers can consider adding notifications pursuant to Article 50(3) AI Act to the privacy statement provided to data subjects under EU data protection law (for example, where appropriate, providing it at the moment of consent collection where consent is relied on as a legal basis for personal data processing).

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

(105) Article 50(4) AI Act establishes two obligations addressed to deployers of generative AI systems requiring clear and distinguishable disclosures (i) of deep fakes and (ii) of 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

([29]) 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).

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[30].

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

(106) Article 50(4) subparagraph 1 AI Act applies if several conditions are fulfilled:

i. The system must be an AI system;

ii. Used by deployers for non-personal 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 where its use is authorised by law to detect, prevent, investigate or prosecute criminal offence.

6.1.1. The notion of 'deep fake'

(107) 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. For the purpose of clarifying this definition, four elements are examined below: (i) resemblance of (ii) existing (iii) persons, objects, places, entities or events that (iv) would falsely appear to a person to be authentic or truthful.

i. **Resemblance:** First, 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 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 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.

ii. **Existing:** Second, 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 can exist or could have existed in reality to be considered a deep fake. 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

([30]) See Recital 143 that clarifies that those transparency obligations for deployers apply '[f]urther to the technical solutions employed by the providers of the AI system'.

elephants driving cars) are considered unrealistic and therefore fall outside the scope of the transparency obligation.

iii. **Persons, objects, places, entities or events:** Third, 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 natural, human beings[31]. "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 (such as e.g. historical events or the depiction of 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 to potentially deceive or mislead a person regarding the content's authenticity or truthfulness[32]. Content authenticity refers to whether the content is what it purports to be in terms of its source or creation process. Truthfulness pertains to the veracity of the content.

(108) Importantly, the assessment of the last criterion does not consider the intention of the deployer to deceive or mislead the natural persons exposed to the content. Rather, it should duly take into account the possible (diverse) composition of the audience that may be exposed to the deep fake content and cannot be based on a hypothetical "average" person expected to be exposed to the content (as opposed to Article 50(1) AI Act, see Section 3.3.1). Due consideration of audience composition is especially important if it is likely that the deep fake content may be at some point perceived by children, elderly persons 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.

(109) In a similar vein, AI-supported manipulation of minor, 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, lighting adjustments, adapting audio parameters, colour correction, noise reduction, improving accessibility or file compression. Such evaluation is, however, inherently case-specific and needs to consider the context and the impact on the persons' perception of the content's authenticity or truthfulness in the specific case. For example, substantive AI-powered editing of background details of journalistic images likely negatively affects the content's authenticity and truthfulness, while AI-powered colour correction or background extensions of existing content or re-scaling of images applied in product advertisements is likely to have only a minor impact on a person's perception of authenticity and truthfulness of the ad and the product.

([31]) This could cover, for example, digital replicas of real persons, as well as realistic AI-generated human avatars or personas.

([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

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

- AI-manipulated image of a scene involving two 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.

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-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 an historical event

6.1.2. The disclosure obligation under Article 50 (4), subparagraph 1 AI Act

(110) Article 50(4), subparagraph 1, 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 clarifies that deployers using AI systems to create 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), subparagraph 1, 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.

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

(111) Article 50(4), subparagraph 1, 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 content in an appropriate manner that does not hamper the display or enjoyment of the work.

(112) **Relevant content categories**: First, to benefit from attenuated treatment, the deep fake (image, audio or video) content should form part of, at least, one of the following five content categories: artistic, creative, satirical, fictional, or analogous works or programmes. Whether deep fake content 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 display creative choices by the deployer of the AI system. 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 works that involve persons, objects, places, entities or events, in an imaginary setting.
- Analogous works are to be understood as works sharing core traits or serving similar expressive or functional purposes with the above categories, but not fitting neatly into one (e.g. because they have a secondary informative or commercial purpose).

(113) The above categories of content 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' or 'programmes' in the context of Article 50(4) AI Act does not imply that such content is or may be protected under Union law on copyright and related rights.

(114) **Evidently:** Second, the fact that a deep fake falls within one (or more) of the five content categories should be evident. Content whose nature is potentially unclear or ambiguous to the audience is excluded from scope. 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 etc.); (ii) the context in which the content is presented (e.g. if the platform, medium or place where the content is presented 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 five content categories also excludes content from the scope of those content categories if it serves primarily an informative or commercial purpose and is recognisable as such. In this respect, some kinds of content (e.g. advertisements or documentaries) containing deep fake content might be regarded as evidently creative or fictional works in certain situations, but not in others, since the assessment is case-specific.

(115) **Appropriate disclosure not hampering the display or enjoyment of the work**: Deep fakes that form part of evidently artistic, creative, satirical or fictional works are not excluded from the transparency obligation of Article 50(4), subparagraph 1, 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 works. 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 (incl. nature of the work, audience, context, etc.).

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

(116) **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 is "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 personal image 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 is therefore not a justification for falling to respect the rights of the rightsholders under Union law on intellectual property or EU data protection law.

Examples of creative, satirical, fictional or analogous works

- *Artistic/fictional work:* AI-generated special effects in movie scenes that constitute deep fakes such as simulations of actors, de-aging of existing actors, digital replicas of dead persons etc.

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

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

– *Analogous creative/fictional work:* AI-generated gaming imagery involving deep fake simulations of existing persons or locations.

Examples that do not constitute artistic, creative, satirical, fictional or analogous work:- AI-manipulated video involving deep fake simulation of humans advertising a product in an AI-generated scene depicting the 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, lacking any fictional, satirical or analogous purpose.

6.1.4. Exception for law enforcement

(117) If the use of a deep fake is authorised by law to detect, prevent, investigate or prosecute criminal offences, deployers are fully exempted from the transparency obligation under Article 50(4) AI Act. The clarifications provided in points 47-48 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

(118) The transparency obligation under Article 50(4), subparagraph 1, AI Act has an important interaction with Article 35(1)(k) DSA.[34] The latter provision lists as a risk mitigation measure that providers of VLOPs and of VLOSEs may put in place to prominently mark "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" in order to ensure that such content is distinguishable when displayed on their online interfaces. In addition, they could also provide an easy-to-use functionality which enables recipients of the service to indicate

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

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**: Article 35(1)(k) DSA covers a wider range of "false" content than Article 50(4) AI Act, regardless of the technologies used to create the content.
- **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. Generally, deployers of AI systems integrated into a VLOP or VLOSE under the AI Act will be considered 'recipients of the service' under Article 3(b) DSA[35] if they disseminate such content. 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 content's dissemination on the VLOP or VLOSE used.

(119) If deep fake content includes personal data relating to a directly or indirectly identifiable natural person (who is alive), then deployers as controllers under EU data protection law will also need to comply with the respective personal data protection obligations[36]. This includes, for example, reliance on an appropriate legal basis for the processing of the personal data, transparency towards the data subjects depicted, etc.

(120) 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 law on copyright or related rights) if their deep fake incorporates subject matter protected under those laws.

(121) 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.

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

(122) Article 50(4), subparagraph 2, 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.

([35]) Recital 2 DSA.

([36]) 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).

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

(123) For the purpose of clarifying the scope of Article 50(2), 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.[37]

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 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 text documents.

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, text having an entirely distinct objective (e.g. entertainment) is not covered by the transparency obligation in Article 50(4), subparagraph 2, AI Act.

iii. **On matters of public interest:** Importantly, 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, European 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 including, but not limited to, public administration and services, fundamental rights (incl. the administration of justice and law enforcement), public health, environmental protection, consumer safety and any economic, political, scientific, or cultural development with potentially important public implications. This list is not exhaustive; 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:

- 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 published academic paper 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.

([37]) In situations where all three requirements are fulfilled, deployers will not be able to rely on the exception for purely personal, non-professional activity from the scope of application of the AI Act (Article 2 (10) AI Act, see also section 2.4.1).

([38]) Inspired by the notion of 'public' as construed by the CJEU case-law (see e.g. C-89/04, C-192/04, C-306/05, C-135/25).

- 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 under Article 50(4) AI Act:

- AI-generated fictional novels or poems in any genre.

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

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

6.2.2. The disclosure obligation under Article 50(4) subparagraph 2

(124) Article 50(4), subparagraph 2, 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 (incl. disclaimers) applied in accordance with Article 50(4), subparagraph 2, 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.

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

(125) Article 50(4), subparagraph 2, 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 and appropriate, deployers may also rely on relevant applicable professional or deontological standards.

i. Text under human review or editorial control

(126) The first condition for the exception 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 competence and professional judgement pertaining to the subject matter under scrutiny. 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).

(127) Superficial, solely formal or procedural checks (e.g. spell-checking or grammatical correction), the mere existence of an editorial policy 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.

ii. Editorial responsibility

(128) The second condition for the exception 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 or the natural person 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 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.

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 EU 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.

6.2.4. Interplay with other Union legal acts

(129) For European media law, editorial responsibility is a crucial concept defined in Article 2(8) of the European Media Freedom Act (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), subparagraph 2, AI Act, in particular where the deployers of the AI system also qualify as media service providers

([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) (Text with EEA relevance) PE/4/2024/REV/1 OJ L, 2024/1083, 17.4.2024.

under the EMFA. Editorial responsibility as used in Article 50(4) AI Act should therefore be interpreted in line with the existing EU 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.

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

(130) Pursuant to Article 50(5) AI Act, the information to be provided in accordance with Article 50(1)-(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.

7.1. Information provision in a clear and distinguishable manner

(131) Information will be considered to be provided in a clear manner where it is noticeable and easy to understand by the person concerned (including persons with accessibility needs). 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 [40]. To provide information in a clear and distinguishable manner, providers or deployers may *inter alia* present the information 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 is only included as part of a manual or hidden under layers of menu options on an online interface.

7.2. First interaction or exposure

(132) 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 AI-generated or manipulated content, the information obligation therefore 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. Moreover, providers and deployers are allowed to inform natural persons earlier than the first actual interaction or exposure, which may be implemented as appropriate, taking

([40]) This should notably ensure that those notifications are: (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 require the use of graphics, videos, and/or characters or other techniques; (vi) given to children gradually and overtime to maximise retention by the user.

into account the objective of the obligations to ensure an effective provision of the information. For example, as natural persons can start watching later than the beginning of a live broadcast that features deep fakes, disclosure should not only be done at the beginning of the broadcast but also at later stages or persistently.

- **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. Disclosure when ending interaction does not comply with Article 50(5) AI Act.
- *First exposure:* at the start of a video featuring deep fake content, at the start of an AI-generated or manipulated text publication on matters of public interest or when encountering AI-manipulated deep fake content when scrolling on social media. Disclosure as part of end credits does not comply with Article 50(5) AI Act.

7.3. Compliance with applicable accessibility requirements

(133) 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 Directives 2016/2102[41] and Directive 2019/882[42]) and assess whether their product, service or contents needs to comply with them. 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

(134) This section aims to clarify (1) the effects of adhering to an approved code of practice as regards compliance with the transparency obligations for AI generated content, (2) the role of the competent supervisory authorities in the enforcement of Article 50 AI Act, and (3) the applicable penalties for infringements of those obligations.

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

(135) 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 deemed adequate by the AI Office[43]. While providers and deployers may also demonstrate compliance with those obligations through adequate alternative means,

([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]) Article 50(7) AI Act. See relevant information about the ongoing process on development of a Code of Practice on marking and labelling of AI-generated content | Shaping Europe's digital future.

adherence to a code of practice that is deemed adequate is a straightforward way of demonstrating compliance. For signatories of a code of practice that is deemed adequate, the Commission and the competent market surveillance authorities should focus their supervisory activities on assessing whether those signatories have adhered to the code of practice. 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.

(136) Providers and deployers that are not signatories to a code of practice that is deemed adequate by the AI Office 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 by carrying 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. In that respect, providers may also be subject to a larger number of requests for information and requests for access to assess, for example, the effectiveness, interoperability, robustness and reliability of the technical solutions used in accordance 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. 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 are ensuring compliance with their obligations under Article 50(2), (4) and (5) AI Act, they will likely need more detailed information when monitoring for compliance.

(137) 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[44].

(138) If a code of practice is not deemed adequate by the AI Office, 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[45].

8.2. Market Surveillance Authorities

(139) Market surveillance authorities designated by the Member States, the AI Office[46], and the European Data Protection Supervisor[47] are responsible for enforcing the rules for AI systems falling within their competence, including the transparency obligations laid

([44]) Article 99(7)(e) AI Act.
([45]) Article 50(7) AI Act.
([46]) 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.
([47]) As the market surveillance authority for the EU institutions, agencies and bodies.

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. The enforcement powers of market surveillance authorities in relation to AI systems are laid down in the AI Act and in Regulation (EU) 2019/1020. 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[48].

8.3. Penalties

(140) 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.

8.4. Entry into application

(141) 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, regardless of their date of placement on the market or putting into service[49]. The AI Omnibus proposal which is currently examined by the EU co-legislators envisages a targeted grandfathering rule only with regard to the marking and detection obligations under Article 50(2) for generative AI systems placed on the market or put into service before 2 August 2026, giving providers of those systems a transitional period to bring their systems in conformity.

(142) AI-generated or manipulated outputs (including deep fakes) within the scope of Article 50(2) and (4) AI Act, which are generated and already made available before 2 August 2026 do not need to be marked or labelled retroactively. However, 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 possession of and/or disseminating such content are encouraged to do so.

9. REVIEW AND UPDATE OF THE COMMISSION GUIDELINES

(143) 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,

([48]) Article 85 AI Act.

([49]) 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.

the research community, and civil society organisations to contribute to this process by responding to future calls for public consultation.