STAKEHOLDER CONSULTATION ON TRANSPARENCY REQUIREMENTS FOR CERTAIN AI SYSTEMS UNDER ARTICLE 50 AI ACT

SUBMISSION PREPARED BY TECHNOLOGY AND RIGHTS CLUSTER, SCHOOL OF LAW, UNIVERSITY OF GALWAY

KWARKYE, THOMPSON GYEDU; KENNEDY, RÓNÁN; PUTRANTI, DESLAELY; ZHANG, MINGZHU; LIU, XINPENG; BAYINDIR, ANIL SENA; REKAS, ABIGAIL

Section 1. Questions in relation to Article 50(1) AI Act

Question 1. Please provide practical examples of AI systems that directly interact with natural persons, as well as examples for which there is doubt and you would seek clarification or consider out of scope.

Practical Examples

a. **Large Language Models (LLMs), Conversational AI developed by tech companies:** These LLMs are typically trained through deep learning methods on massive corpora, enabling them to identify patterns in the data. Based on these recognised patterns, they can generate real-time responses to user inputs that align with the detected patterns. Typical examples include ChatGPT, developed by OpenAI, Gemini, developed by Google, and Grok, developed by xAI.

b. **AI virtual assistants embedded in appliances:** Unlike LLMs that primarily engage in direct conversation with users, this category of AI virtual assistants is often integrated into hardware devices such as smartphones, smart speakers, or even refrigerators, where they serve to recognise and execute user commands. Namely, the primary function of such AI assistants is not to converse with users but rather to receive commands and perform tasks on the appliance. A typical example is Siri, the voice assistant on iPhones. These AIs are generally built on several core components such as Natural Language Processing (NLP), Speech Recognition, and Machine Learning. It is worth noting that AI virtual assistants may also employ deep learning techniques similar to those used in training LLMs. For example, Apple is reportedly attempting to rebuild Siri using LLM-based technologies.[1]

c. **Task-oriented AI:** In various industries, task-oriented AI systems are trained to accomplish specific functions. These AIs may employ different techniques during training, including Natural Language Processing, Deep Learning, and Planning Algorithms.[2] Many such systems directly interact with users while executing their tasks. For instance, to achieve health-monitoring tasks, Apple has developed the Apple Watch, which interacts with users' bodies through sensors to perform health checks. Another example is driver-assistance AI, which engages in user interaction by alerting drivers to fatigue or issuing risk warnings. It should be noted that, given the complexity of AI applications in real-world contexts, these three categories may

overlap. For example, an online customer service chatbot on an e-commerce platform may simultaneously fall into category (a) as an LLM-based conversational AI and serve as category (c) by addressing customer inquiries as a task-oriented AI.

Examples with Doubt

a. **Algorithmic recommendation AI:** This type of AI analyses user preferences and adjusts the content recommended to users. It is commonly found on online platforms such as TikTok. There is doubt, however, as to whether users are in direct interaction with such AI systems. One view is that users do not interact directly with the recommendation algorithm, since they cannot explicitly instruct the AI on what content to recommend; instead, they only influence the algorithm indirectly through their choices on the platform. By contrast, it can also be argued that users interact directly with the AI since every click on content constitutes feedback to the system. At the very moment the user clicks, the recommendation AI updates its judgment and produces new recommendations, which amounts to a direct interaction with the user. Therefore, clarification is needed as to whether algorithmic recommendation AI falls under the notion of "direct interaction."

b. **Decision-making AI:** Decision-making AI is a subcategory of task-oriented AI. It refers to systems that can unilaterally affect the user, without the user being able to influence the AI in return. Typical examples include automated facial recognition access control systems, where the AI unilaterally analyses the user and decides whether to grant access. If the system denies entry, the user generally has no means to contest the AI's decision. Similarly, some driver-assistance systems can override the driver and take control of the vehicle in dangerous situations. In such cases, the driver has no choice but to accept the AI's unilateral intervention. Thus, clarification is needed on whether such AI systems, which exert one-way influence over users, should be regarded as engaging in "interaction" with natural persons.

Question 2. Please provide practical examples of AI systems that directly interact with natural persons and that can be authorised by law to detect, prevent, investigate, or prosecute criminal offences. For each system, provide the law that can authorise the use and describe appropriate safeguards for the rights and freedoms of third parties.

a. **Facial Recognition:** Facial recognition can be used to identify criminals and victims. For example, devices equipped with facial recognition technology can be installed in public spaces such as airports and metro stations to locate suspicious criminals. Once a victim has been found, the police can also use facial recognition to confirm the victim's identity online. However, facial recognition may raise concerns about violations of fundamental rights, especially for marginalised groups. For instance, in Ireland, the police have attempted to use facial recognition technology (FRT) in law enforcement, which has sparked widespread debate.[3]

A clear legal basis is required to prevent facial recognition from infringing on human rights. In the EU, the use of facial recognition must comply with both the EU AI Act and the General Data Protection Regulation (GDPR). Under Article 5(1)(h) of the AI Act, the use of "real-time" remote biometric identification systems is prohibited, with limited exceptions, including locating victims, preventing major security threats, and locating individuals in the context of criminal investigations. The use of facial recognition must therefore be restricted to these exceptions. According to Annexe III(1), legitimate biometric identification systems are classified as high-risk AI systems, which means facial recognition must also comply with the requirements for high-risk AI under the AI Act. In addition, under GDPR Article 9, facial data are considered a special category of personal data. They may only be processed under specific conditions, such as for reasons of substantial public interest. Thus, the use of facial recognition cannot go beyond the circumstances permitted under Article 9 GDPR.

b. **AI Camera:** In smart cities, AI cameras can help prevent crime.[4] For example, if an AI camera detects unusual movements, the police may receive a real-time alert. In the UK, such AI cameras have already been deployed—for instance, in Devon and Cornwall, where the police use AI cameras to detect drivers who may be under the influence of alcohol or drugs.[5]

In the EU context, AI cameras must comply with both the EU AI Act and the GDPR. According to Annexe III(6) of the AI Act, AI systems used for law enforcement purposes are classified as high-risk AI, meaning that their use must meet the AI Act's requirements for high-risk systems. Furthermore, the processing of data by AI cameras must comply with GDPR Article 6. For example, processing can be based on Article 6(1)(e), where data are processed to perform a task in the public interest. If the AI camera captures biometric data such as human faces, the processing must also meet the conditions under GDPR Article 9, which restricts the processing of biometric data to the situations explicitly permitted therein.

c. **Fraud Detection:** Fraud detection AI can identify patterns of fraudulent behaviour. Businesses and financial institutions can use fraud detection AI to flag suspicious transactions and reduce losses.[6] In Ireland, banks and other financial institutions are already using AI to detect fraud.[7] According to a survey by insurance broker and risk management company Gallagher, 94% of businesses reported using AI, with fraud detection accounting for 34% of AI applications.[8]

In the EU, fraud detection must comply with GDPR Article 6, particularly Article 6(1)(f). This allows companies, as data controllers, to process data where it is necessary to pursue their legitimate interests, provided these interests are not overridden by the interests or fundamental rights and freedoms of the data subject.

d. **Cybercrime Detection:** Many crimes, such as drug trafficking, child exploitation, and human trafficking, are carried out through online platforms. Authorities can use AI to identify suspicious keywords on these platforms and thereby uncover criminal activity.[9]

In the EU, when authorities use AI to search online platforms for potential criminal activity, they must comply with Article 8 of the Law Enforcement Directive (Directive (EU) 2016/680). This provision requires that the processing of personal data is lawful only if it is carried out by competent authorities for the purposes of the prevention, investigation, or detection of criminal offences, or the prosecution of such offences.

Question 3. If you are aware of any examples of notification techniques that can be employed with interactive AI systems, including embedded in their design, to duly inform natural persons that they are interacting with an AI system, please provide them in your response.

Examples of Notification

a. **Textual prompts in the user interface:** In many AI products, developers inform users that the product is AI-powered through textual prompts embedded in the interface. For example, one textual prompt in OpenAI's ChatGPT is "ChatGPT: smart and simple AI". Similarly, Google introduces its Gemini system with "Meet Gemini, Google's AI assistant". On WhatsApp, users interacting with Meta's AI assistant are also shown introductions reminding them that Meta is an AI. Textual prompts are highly effective in ensuring transparency in user–AI interaction because they are embedded directly into the interaction platform, meaning that each time users engage with the system, they are reminded that they are interacting with AI.

b. **Watermarks:** Unlike textual prompts that appear within the interaction interface of an AI platform, watermarks are usually embedded directly into AI-generated visual outputs. For example, a watermark might appear in the background or corner of an AI-generated image, stating that the picture was created by artificial intelligence. Such watermarks provide an additional layer of transparency, reminding users that they are interacting with AI. Importantly, watermarks are not only valuable for direct users of AI platforms but also for individuals who did not use the AI system themselves but come into contact with AI-generated images.

c. **Voice prompts:** Where AI applications involve conversational interactions, developers may use spoken prompts to inform users. For instance, when asked about its identity, Apple's iPhone assistant Siri explicitly states that it is a "virtual assistant", thereby making clear to users that they are interacting with an AI system. Where AI applications involve conversational interactions, developers may use spoken prompts to inform users. For instance, when asked about its identity, Apple's iPhone assistant Siri explicitly states that it is a "virtual assistant", thereby making clear to users that they are interacting with an AI system. Their effectiveness depends on their clarity and regularity.

Question 4. Are there aspects related to the scope or practical implementation of the transparency obligation for interactive AI systems under Article 50(1) for which you would seek further clarification?

a. **Greater specificity on cases where transparency need not be implemented:** According to Article 50(1), when, from the point of view of a natural person who is reasonably well-informed, observant and circumspect, it is obvious that they are interacting with an AI system, the AI provider is not required to implement the transparency obligation. However, determining what constitutes a well-informed, observant and circumspect natural person is subjective; therefore, guidelines are needed to provide more concrete criteria. For instance, an elderly individual may be cautious but unfamiliar with AI technologies, making it difficult to recognise whether they are interacting with AI. Whether such a person should be considered well-informed, observant and circumspect under Article 50(1) could be contested. In addition, guidelines should, to the greatest extent possible, list scenarios in which the transparency obligation applies, to assist AI providers in understanding their duties.

b. **More clarity on "direct interaction":** Guidelines are needed to provide concrete examples clarifying what qualifies as "direct interaction." For example, when a user engages with ChatGPT, it is generally undisputed that this constitutes direct interaction. However, in other cases, direct interaction may be debatable. For instance, some search engines are AI-powered, yet users may not consciously be aware that they are interacting with AI when performing searches, even though the engine operates with AI in the background, or when a person uses language editing tools like Grammarly, it is not always clear where to draw the line. Whether such use counts as direct interaction requires clarification. Similarly, suppose a person encounters and uses an AI-generated image without directly using the AI model to create it. In that case, the question arises as to whether engaging with AI outputs alone constitutes direct interaction. Such situations call for further clarification.

c. **Greater standardisation of notification methods:** Given the complexity of user–AI interactions, it is necessary to standardise the methods for implementing transparency. AI models may interact with users through text, voice, or images. It should be made clearer which form of notification is required for each type of interaction. In the case of multi-modal systems that engage with users through multiple channels, such as text, speech, and images, it is unclear how providers can fully implement transparency, and further clarification is needed. Moreover, users themselves are highly diverse, including vulnerable groups such as the blind, the deaf, or individuals with other disabilities. Standardisation should specify how notification obligations must be fulfilled in relation to these vulnerable populations. Beyond user diversity, AI systems are deployed across a wide range of industries, including law, finance, and healthcare. Standardisation guidelines should clarify what conditions notifications must meet in different sectors for the transparency obligation to be considered fulfilled.

Section 2. Questions in relation to Article 50(2) AI Act

Question 5. Please provide practical examples of AI systems that generate synthetic text, audio, image, or video content as well as examples of systems for which there is doubt and you would seek clarification or consider them out of scope.

Several AI systems are quite popular and frequently used to assist and generate content as follows:

a. Gemini (Google), GPT (OpenAI), and Perplexity. It usually generates text and images based on the prompt provided by users.

b. Muse AI and Suno AI. This type of AI generates Music/Audio.

c. DALL-E, Stable Diffusion, OpenArt, and Gamma. This type of GenAI generates images and a slideshow (Gamma).

d. Midjourney, Sora (OpenAI): This type of GenAI generates Video.

e. Deepseek. Deepseek generates text only.

Based on the examples above, some borderline cases need to be clarified since not every content with AI assistance can be considered AI-generated content. For example, tools that merely enhance or modify existing material, such as upscaling software that improves image resolution, noise reduction in audio processing, or automatic colour correction in audio and video, do not create substantially new content and may reasonably be considered out of scope. More applications lie in a grey area and require further analysis; some of the examples are as follows:

a. **Text – Writing Assistants or Content Generation (e.g. Grammarly):** Spelling and grammar checkers have long supported writing improvement, but AI-driven tools like Grammarly extend these functions beyond basic corrections. The platform provides real-time feedback, explaining errors and suggesting edits, while allowing users to accept changes. Its premium version adds features such as sentence restructuring, readability enhancements, and tone suggestions, allowing the software to analyze complex sentences and propose clearer rewrites. As such tools become more prevalent in educational settings, their impact on writing instruction and student learning warrants careful consideration.

b. **Audio – Voice Conversion:** Voice conversion (VC) stands as a crucial research area in speech synthesis, enabling the transformation of a speaker's vocal characteristics to resemble another while preserving the linguistic content. This technology has broad applications, including automated movie dubbing, speech-to-singing conversion, and assistive devices for pathological speech rehabilitation. Voice conversion and similar audio transformation technologies occupy a somewhat ambiguous position in the context of Gen-AI. Systems that extract voice embedding or other biometric features and synthesise a new audio waveform that mimics the original speaker introduce

opacity. The output might be considered partially generated content, as it is derived from the input but involves new elements such as the new vocal print and tract.

c. **Image Filter:** Many filters, such as adjusting brightness, colour, and saturation contrast, these actions primarily modify existing images and do not create substantially new content, in terms of the standards and principles from copyright law, they typically lack originality. By contrast, filters that produce distinctive visual transformations or artistic effects, especially when driven by unique stylistic combinations, may generate content that is sufficiently original, for example, to qualify for intellectual property protection. It is therefore important to distinguish between simple functional adjustments and transformations that produce perceptibly new or creative outputs.

Question 6. Please provide examples of marking and detection solutions, including combinations of techniques, that can be employed to mark in a machine-readable format AI-generated or manipulated content and enable detection whether the content has been generated or manipulated by AI.

Some marking techniques have been employed in work even before the arrival of GenAI. However, with the advancement of GenAI, there are several examples of marking and detection techniques that can be applied to ease the user/audience's distinct between human-made and AI-generated content as follows:

a. **Watermarking/Cryptographic signature.** Watermarking is the process of adding a visible or invisible identifier to a digital work or document to mark ownership, prevent misuse, or provide authentication. The goal is to protect the authenticity of the content from unauthorized claims or use by others. Cryptographic Signature is an advanced version of watermarking for the abovementioned content. It uses encryption to enhance security and authentication as well as detect data manipulation. Both Watermarking and Cryptographic Signature are suitable for image, text, and AI-generated video.

b. **Metadata tagging:** A descriptive label applied to digital data will make it easier to manage, search, and understand without changing its appearance. Metadata tagging means attaching descriptive labels, such as author, date, and keywords to digital content so that it can be stored, organized, and searched more easily[10].It does not change how the file looks, but it helps both users and machines understand where the content comes from. For AI-generated works, metadata can include the model's name, the prompt used, or a flag showing that it is synthetic. It is suitable for text, images, and videos.

c. **Digital Labelling:** A labelling system using digital data that can be accessed through a machine. The broad form of labelling involves watermarking and disclosure of the provenance of the users post. Suitable for image, video, audio, text, or some combination (gives AI-generated label)[11]

d. **Blockchain:** Blockchain is a decentralized digital ledger technology that records data in a chain of cryptographically linked blocks. Each block contains a timestamp and a hash of the previous block, ensuring data immutability, consistency, and transparency. In digital content marking, blockchain can combine with digital watermarking to provide tamper proof copyright verification and content tracking.[12] Blockchain is suitable for all types of content

Question 7. For each of the solutions included in the previous question, please clarify whether there is relevant information that can help you competently assess their effectiveness, interoperability, robustness and reliability 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. Please also assess to what extent the detection mechanisms are accessible and enable people exposed to the AI generated or manipulated content to identify its origin.

Each of the following solutions has specific strengths and limitations in terms of robustness, interoperability, reliability, cost, and accessibility. Their effectiveness often depends on the type of content (text, image, video) and the technical means available for detection. While some solutions (e.g., cryptographic signatures, blockchain) offer stronger guarantees of authenticity, they may involve higher costs or reduced interoperability. Others (e.g., watermarking, metadata tagging, digital labelling) are more accessible and easier to deploy but often suffer from weak robustness or lower reliability. The detection mechanisms vary in ease of access and usability for end-users, affecting how effectively individuals can identify the origin of AI-generated or manipulated content.

Solution	Robustness	Cost	Interoperability	Accessibility	Reliability	Suitability
Watermarking	Weak, easy to remove/edit	Cheap	Easy	Easy	Low	Text, Image, Video
Metadata Tagging	Weak, easy to remove	Cheap	Easy	Easy	Low	Text, Image, Video
Digital Labelling	Weak	Medium	Unsure	Unsure	Low	Video (e.g., AI-generated label)
Cryptographic Signature	Strong	Medium (not cheap, not very	Easy	Easy	Strong	Broad applicability across

		expensive)				content types
Blockchain	Strong	Expensive	Difficult	Depends on blockchain (public vs private)	Strong	Broad applicability (but with limitations)

Labelling AI-generated content is an important tool for increasing transparency and helping users identify the source of content, which improves their ability to access its authenticity. Mandatory labelling, including visible or invisible markers, can also support platforms in fulfilling governance responsibilities and maintaining the integrity of public discourse. However, such labelling faces several challenges. Labels can be removed or altered, especially if they are hidden or machine-readable. Different platforms often use varying standards and formats, making cross-platform recognition difficult. Implementing and maintaining automated labelling systems requires technical development and server resources, which increases operational costs. Enhancing robustness may also affect content quality, and visible labels can interfere with the user experience while hidden labels may not be easily accessible or understood by all users. Finally, labelling alone cannot be easily accessible or understood by all users.[13]

Blockchain offers a robust method for securing the provenance of AI-generated or manipulated content by providing immutable, timestamped records that verify authorship and detect tampering. When combined with digital watermarks, it can enhance copyright protection, prevent unauthorized use, and enable transparent tracking of content distribution. However, blockchain-based solutions also face practical challenges. Implementing and maintaining such systems requires significant technical development and infrastructure, which can be costly and complex. Scalability and performance limitations may restrict real-time processing of large volumes of images, videos, or audio files. Different platforms often lack standardized protocols, making cross-platform verification difficult. Moreover, while blockchain ensures backend verification, end-users typically cannot directly access or interpret these records, meaning it mainly supports institutional or platform-level checks rather than immediate public identification. [14]

Question 9. Are there any other aspects related to the scope or the practical implementation of the transparency obligation for generative AI systems under Article 50(2) for which you would seek clarification?

Yes, several aspects of the transparency obligation under Article 50(2) for generative AI systems would benefit from further clarification. Transparency in generative AI operates on two interrelated levels.

Firstly, from a technological perspective, it requires clear disclosure of the data sources used to train the model, the methods of data collection, and the processes involved in data curation and model development. This includes specifying whether the data was publicly available,

licensed, or obtained through other means, and whether it includes personal or copyrighted content. It should also indicate the degree of AI involvement, distinguishing fully AI-generated works from AI-assisted creations, and disclose stylistic influences or creative parameters, such as artistic styles, musical genres, or literary conventions referenced by the AI models.

Secondly, from the external oversight and accountability perspective, third-party auditors or regulators must provide sufficient information to assess compliance. This includes access to documentation outlining the purpose of data use, the legal basis for processing, and evidence of consent or permission where applicable. It must be demonstrable that the data used does not infringe on individuals' personal rights or commercial interests.

Further regulatory guidance is needed to clarify:

a. Whether transparency refers primarily to the technical mechanisms (e.g., model architecture, training data lineage) or the governance principles (e.g., fairness, accountability, consent).

b. What constitutes sufficient disclosure for compliance purposes?

c. How transparency obligations interact with trade secrets and intellectual property protections.

d. The role of standardized documentation or auditable reporting frameworks in ensuring consistent implementation across providers.

For AI users, platforms and institutions should provide comprehensive training and clear guidelines on how to assess the extent of AI involvement in a work, including how to distinguish between fully AI-generated content, AI-assisted creation, and human-authored elements. Users should also be guided on how to ethically disclose AI contributions in different contexts and how to explain stylistic influences or references drawn from AI models. This guidance could include practical examples, checklists, or standardized disclosure templates to ensure consistent and understandable communication to audiences. Platforms should also provide educational resources on the legal and ethical implications of using AI-generated content, helping users make informed decisions while respecting intellectual property and moral rights.

For content uploaders or data providers, transparency mechanisms should allow flexible options over whether, how, and under what conditions their works can be used for AI training or model refinement. Instead of a binary agree or reject choice in terms of services, creators should select specific usage types, licensing conditions, or temporal limits, and be able to withdraw or modify permissions over a certain period. Platforms might also implement clear dashboards showing how uploaded works are used in AI systems, offering visibility and accountability. Such measures empower creators to protect their rights, maintain control over their artistic contributions, and participate in AI development in a transparent and informed

manner, while simultaneously supporting ethical and responsible AI practices across the creative ecosystem.

Such clarification should consider not only stressing creators' rights but also promoting informed participation, ethical engagement, and trust in AI-generated artistic works. By integrating these technical, ethical, and user-centered considerations, transparency in generative AI can move beyond formal compliance to become a practical and meaningful safeguard for creators, users, and affected communities.

Section 3. Questions in relation to Article 50(3) AI Act

Question 10. Please provide practical examples of AI systems that may be considered emotion recognition and biometric categorisation systems.

Before providing practical examples of AI systems that may be classified as emotion recognition or biometric categorisation systems, it is vital to highlight the distinctions between these systems. Although these categories may appear similar due to their shared reliance on the processing of personal biometric data, they differ in both their operational mechanisms and the intended purposes. According to Article 3(35) of the EU AI Act, biometric identification refers to processing physical, physiological, behavioural, or psychological human features to establish a natural person's identity by comparing their biometric data to the stored database. However, Recital 15 of the Act clarifies that AI systems intended for biometric verification are excluded from the scope of biometric identification defined in the Act.

In contrast, emotion recognition is defined in Article 3(39) as the processing of physical, physiological, or psychological biometric data for the purpose of identifying or inferring the emotions or intentions of natural persons. Similarly, biometric categorisation is addressed in Article 3(40) to describe the assignment of individuals to specific categories based on their biometric data, unless such categorisation is ancillary to another commercial service and strictly necessary for technical reasons. Recital 16 further elaborates that these categories may pertain to characteristics such as sex, age, hair colour, eye colour, tattoos, behavioural or personality traits, language, religion, membership of a national minority, or sexual or political orientation.

Based on these explanations, AI systems that might be considered as emotion recognition can be listed below under five categories depending on the type of personal input data:

a. Oral input data: AI-based chatbots in call centres categorise calls according to the callers' emotions, depending on their voice and word choice, by processing voice tone detection to triage the calls.

b. Written input data: Large language models (like ChatGPT) adjust their tone according to the texts written by the user or enable parental control by sending notifications to parents of minors when the system detects their teen is in a moment of acute distress.

c. Audio input data: AI-based song recommendation systems conduct emotion perception from the musical choices of a user to recommend new music or generate new music with AI-based tools.

d. Visual input data

i. AI-based camera systems in museums process visitors' facial expressions to determine the attention rate of each piece in the gallery.

ii. AI-based video call tools like Zoom—to assess candidates' emotional skills during recruitment processes by processing facial expressions, pupil dilation, and body postures.

iii. AI-based camera systems in classrooms—to monitor students' attention levels by processing facial expressions, pupil dilation, and body postures.

iv. AI-based camera systems in gyms—to assess the mood influence of gym members and assess the success of a gym session by processing facial expressions, pupil dilation, and body postures.

v. AI-based camera systems in cars—to monitor drivers' attention for road safety by processing facial expressions, pupil dilation, and body postures.

vi. AI-based camera systems in airports, railways, and subways to provide early warning through threat detection and terror prevention via emotion recognition by processing suspects' facial expressions, pupil dilation, and body postures.

vii. AI-based camera systems at border controls to prevent illegal migration by processing people's facial expressions, pupil dilation, and body postures to determine their emotional state.

viii. AI-based camera systems will be used in police interrogation rooms and courtrooms to conduct emotion recognition and decide on the truthfulness of suspects' defences.

ix. AI-based camera systems in therapy rooms or AI-based video call tools for online therapy sessions monitor a patient's state or detect mental health issues during therapy sessions by processing facial expressions, pupil dilation, and body postures.

x. AI-based camera systems in houses to monitor domestic peace by processing facial expressions, pupil dilation, and body postures of family members to prevent domestic abuse by calling law enforcement agencies in case of an emergency.

xi. Automated billboards check consumers' attention to better align with the market's needs, increasing product or service sales. They also evaluate voters' attention to design a political agenda.

e. Hormonal, cardiologic or neurologic input data

i. AI-based clothing to conduct emotion recognition through hormones via sweat or saliva for various purposes.

ii. AI-based accessories which can conduct cardiologic or neurologic tests, such as an electroencephalogram or an electrocardiogram, for emotion perception and various purposes.

Emotion recognition can be important in marketing, education, human–robot interaction, healthcare, mental health monitoring, and security.

Secondly, AI systems that might be considered for biometric categorisation can be listed below:

a. AI-based camera systems embedded in stores to analyse gender, age, language, religion, and other demographic factors for better product design and marketing strategies.

b. AI-based camera systems embedded in billboards to classify viewers according to their gender, age, language, and religion to decide which advertisement to display at that moment.

c. AI-based camera systems embedded in billboards to classify viewers according to their political orientation to decide which political campaigns to display.

d. AI-based camera systems embedded on streets to analyse gender, age, language, religion, and disability demographics of pedestrians, enabling local authorities to understand the needs of the region and design their task force accordingly.

e. AI-based chatbots at border control to categorise individuals' language levels for migration purposes.

f. AI-based camera systems at public transport stations, especially subways, are used to classify movements toward the rails to prevent suicide.

Emotion recognition and biometric categorisation can coincide when emotions detected through biometric data are used to classify individuals. For instance, AI-based camera systems in billboards could track voter attention to inform political strategies and categorise individuals into political groups, revealing the demographic distribution of political views in each region.

Question 11. If you are aware of any examples of transparency measures that can be employed with emotion recognition or biometric categorization systems to duly inform natural persons exposed thereto of the operation of the system, please provide them in your response.

When biometric data is processed, the person's consent becomes vital according to the general principles of personal data protection law. For consent to be validly given, informing

the individuals becomes important. Transparency should be a core value and principle for informing individuals. Similarly, since biometric input data will be processed to conduct emotion recognition and biometric categorisation, transparency should lie at the core of these practices.

The principle of transparency can be put into practice by taking specific actions and observing certain considerations in their implementation. For example, transparency in the context of emotion recognition and biometric categorisation can be achieved by providing multi-modal indicators and consent prompts that are as accessible and inclusive as possible. Multi-modal indicators may include written text, audio messages, and captions, while consent prompts should clearly communicate the choices available to users. To ensure these indicators and prompts are effective, they should be clear, precise, and engaging. This can be achieved by carefully addressing content-related and formal communication aspects. These aspects are listed below:

a. **Content-related Requirements:** The indicators should include the purpose of utilising emotion recognition and biometric categorisation. The individuals' right to obtain human review of automated decisions should be highlighted. The way the AI system uses input data to generate the output, such as emotion recognition or categorisation, should be explained in simple terms. The accuracy of the systems should be provided to give individuals an idea about possible false results. Details on what emotional or biometric data is stored, for how long, and who it is shared with should be provided.

b. **Formalistic Requirements:** Indicator texts should be compatible with screen readers. High color contrast should be prioritized between text and background. Sufficient font size and scalable text should be provided. Flashing elements or animations that could cause problems for photosensitive people should be avoided. Plain language free from legal and technical jargon should be used. The context in which emotion recognition or biometric categorisation is performed, whether in public or non-public settings, must be carefully considered because it determines the extent and manner of transparency required for these practices.

The EU AI Act 3(44) defines 'publicly accessible space' as any publicly or privately owned physical place accessible to an undetermined number of natural persons, regardless of whether certain conditions for access may apply, and regardless of the potential capacity restrictions. And this definition is elaborated in Recital 19, and accordingly the activity for which the space may be used is also irrelevant when a place is determined as publicly accessible or not it can be either a bank, café, swimming pool etc. Even access is subject to certain predetermined condition, such as the purchase of a ticket or title of transport, prior registration or having a certain age will not change the status of the place as non-publicly accessible.

So, if emotion recognition and/or biometric categorisation are conducted in a publicly accessible place, all the above requirements would need to be fulfilled to a greater extent. Therefore, the threshold for satisfying the requirements will be higher for emotion recognition and/or biometric categorisation in publicly accessible places, since the stakes associated with freedom of expression, freedom of assembly, the right to privacy, non-discrimination, and democracy will be much more under threat.[15] Accordingly, the purpose of these practices should be clearly stated and justified on valid grounds, the reliability of the systems should be ensured, and the individuals should be warned of the risks of false results and their effect on them with clarity.

Question 12. Are there any other aspects related to the scope or the practical implementation of the transparency requirement for emotion recognition and biometric categorisation systems for which you would seek clarification?

The other aspects related to the scope and the practical implementation of the transparency requirement for emotion recognition and biometric categorisation systems that need more clarification can be categorised under five categories below:

a. **Scope of Biometric Categorisation:** As regards scope, it is important to clarify what differs biometric categorisation from biometric identification, and what conditions must be met for these kinds of conduct to be biometric categorisation and not identification. In addition, the extent to which biometric categorisation and biometric identification can be separated, and how categorisation does not constitute a step towards identification, should also be clarified.

b. **Certainty on the Legality of Biometric Categorisation:** More clarity on the legality of biometric categorisation is necessary. In particular, explaining whether it is a prohibited practice, the grounds for permitting it, the valid reasons for conducting it legally and ethically, and the safeguards required to address risks such as discrimination and surveillance would be valuable.

c. **Scope and Sufficiency of Biometric Data as Input for Such Technologies:** Also, the extent of the biometric data within emotion recognition should be clarified, specifically whether behavioural signals fall within this category, to prevent any form of emotion recognition based on human feature inputs. In other words, the scope of the prohibition with regard to emotion recognition needs clarification. It would be valuable to clarify whether emotion recognition through the processing of biometric data is prohibited, or whether emotion recognition through the processing of human feature input data is prohibited. For example, it is clear that facial recognition technologies process biometric data for emotion recognition purposes, but it is less clear whether processing a person's texts on ChatGPT for emotion recognition purposes falls within the scope of emotion recognition under the EU AI Act, since it is uncertain whether an individual's 'texts' can be regarded as biometric, even though they can be used for emotion recognition.

d. **Emotion Recognition and Biometric Categorisation on Online spaces:** Online spaces are not considered publicly accessible places according to Recital 19 of the EU AI Act. As a result, the legality of emotion recognition and biometric categorisation in online spaces, and the consequences of their use, become unclear when these practices are applied in such contexts. In addition, the risk of passive data collection is greater in these environments; therefore, more detailed guidance on regulating emotion recognition and biometric categorisation in online spaces appears necessary.

e. **Pre-launch transparency approvals:** Transparency requirements should be integrated into AI systems used for emotion recognition or biometric categorisation and be subject to pre-launch approval before such technologies are introduced to the market. The responsibility for ensuring compliance with these requirements should rest with the actors bringing such systems to market. For instance, producers of AI systems designed for emotion recognition or biometric categorisation should be legally accountable for fulfilling all transparency obligations before they introduce the systems to the market. Clear guidance on this liability and the approval procedures is essential.

Section 5. Other horizontal questions in relation to the implementation of Article 50 AI Act

Question 21. Are there aspects related to the AI Act's horizontal requirements in Article 50(5), including their interplay with the requirements in Article 50(1)–(4), for which you would seek clarification?

Article 50(5) introduces horizontal transparency obligations that apply across all AI systems covered under Article 50(1)–(4). While Articles 50(1)–(4) address specific categories such as interactive AI, generative AI, emotion recognition and biometric categorisation, and deep fakes/public interest content, Article 50(5) ensures the foundation of transparency that cuts across these domains. This includes obligations such as informing users about the presence and functioning of AI systems, the nature of the content generated or manipulated, and the potential impact on individuals. However, the practical implementation of these horizontal requirements can be challenging, especially when they overlap or diverge from the more targeted provisions.

Therefore, there is a need to clarify how deployers should interpret and apply Article 50(5) in conjunction with the specific obligations in Articles 50(1)–(4). For instance, should deployers treat Article 50(5) as a foundational layer that supplements the more specific requirements, or should it be considered a standalone set of obligations? There is also the need to understand whether compliance with Article 50(1)–(4) automatically satisfies the horizontal transparency requirements, or whether additional disclosures and safeguards are necessary. A harmonised framework or guidance document from the Commission or national regulators could help clarify this. Such a framework could include practical examples, sector-specific

interpretations, and transparency disclosure templates that integrate horizontal and vertical obligations.

Question 22. Are there any further aspects related to the transparency obligations under Article 50(1)-(5) for which you would seek clarification regarding their interplay with other obligations in the AI Act?

The transparency obligations outlined in Article 50 are interconnected with other requirements of the Act, such as risk management, data governance, human oversight, and robustness. For example, the obligation to inform users that they are interacting with an AI system in Article 50(1) is not only a transparency issue but also relates to accountability, user autonomy, and ethical deployment. Similarly, the requirement to mark AI-generated content mentioned in Article 50(2) intersects with obligations around accuracy, traceability, and preventing misuse. These overlaps raise questions about how deployers should structure their compliance strategies to address multiple, interrelated obligations without confusion.

It is important to clarify whether fulfilling transparency obligations under Article 50 contributes to compliance with other parts of the AI Act, or whether separate documentation and procedures are required. For instance, should the same risk assessment report include transparency disclosures or be documented independently? Furthermore, how should deployers prioritise obligations when resource constraints make full compliance challenging? A cross-referenced compliance framework or checklist could help organisations understand how transparency fits into the broader regulatory landscape. This would also support the development of internal governance structures that align legal, technical, and ethical responsibilities. Regulators could also consider issuing sector-specific guidance that maps transparency obligations to other requirements to help deployers from other sectors navigate the regulatory terrain.

Question 23. Are there any further aspects related to the transparency obligations under Article 50(1)-(5) for which you would seek clarification regarding their interplay with obligations in other Union or national legislation (e.g. data protection regulation such as Regulations (EU) 2016/679 and (EU) 2018/1725 and Directive (EU) 2016/680, Regulation (EU) 2024/900 on the transparency and targeting of political advertising or Regulation (EU) 2022/2065 on a Single Market For Digital Services)?

The transparency obligations under Article 50 must be interpreted in the context of existing EU legislation, particularly the General Data Protection Regulation (GDPR – Regulation EU 2016/679), the Law Enforcement Directive (EU 2016/680), and newer regulations like the Digital Services Act (EU 2022/2065) and the Political Advertising Regulation (EU 2024/900). For example, informing users about AI interactions (Article 50(1)) overlaps GDPR requirements for data processing notices and consent. Similarly, marking deep fake content (Article 50(4)) intersects with obligations under the political advertising regulation, especially during election periods. These overlaps can create legal uncertainty, particularly for deployers across multiple jurisdictions or sectors.

To address this, regulators and policymakers should provide joint guidance or interoperability frameworks that clarify how Article 50 obligations align with or differ from other legal instruments. For instance, a unified consent mechanism could satisfy GDPR and AI Act requirements, while a standardised disclosure format could meet transparency obligations under multiple laws. Additionally, clarification is needed on handling conflicts between statutes, for example, if transparency under the AI Act requires disclosure that might compromise privacy under the GDPR. Sector-specific guidelines would be invaluable in high-risk domains such as healthcare and law enforcement, where multiple regulations may apply simultaneously. The main point in responding to this question is that a coordinated approach to legal interpretation and enforcement will be essential to ensure that AI deployers can meet their obligations efficiently and ethically while safeguarding fundamental rights and freedoms.

Question 24. Are there any recommendations or good practices you would like to share as input for the Code of Practice to operationalise the implementation of the transparency obligations regarding interactive and generative AI systems?

To effectively operationalise the transparency obligations under Article 50, the Code of Practice should promote clarity, consistency, and user empowerment. Good practices could include standardised disclosure formats such as AI labels, metadata tags, user-friendly explanations of AI system capabilities, and real-time notifications when interacting with AI. Watermarking and provenance tracking should be encouraged for generative systems and clear disclaimers for synthetic content. Visual indicators and opt-out mechanisms should be an essential requirement for emotion recognition and biometric categorisation. These practices should be adaptable to different contexts and user groups, ensuring that transparency is meaningful and not merely procedural.

The Code of Practice should also emphasise inclusivity and accessibility. Transparency measures must be understandable to diverse audiences, including those with limited digital literacy or people living with some form of disability. Multilingual support, visual cues, and simplified language can enhance user comprehension. Additionally, the Code should encourage participatory design, where stakeholders, including civil society, end-users, and marginalised communities, contribute to shaping transparency norms. This would foster trust and accountability while ensuring AI systems serve the public interest. Additionally, public education could go a long way in building trust and influencing the understanding of transparency norms. Finally, the Code should include mechanisms for continuous improvement, such as feedback loops, periodic reviews, and alignment with emerging standards and technologies. By embedding these principles, the Code of Practice can become a living document that evolves with the AI ecosystem and supports responsible innovation.