Journal officiel
de l'Union européenne

FR
Série C

C/2026/6005

12.11.2026

COMMUNICATION DE LA COMMISSION

Lignes directrices relatives à la mise en œuvre des obligations de transparence pour certains systèmes d'IA au titre de l'article 50 du règlement (UE) 2024/1689 (règlement sur l'IA)

(C/2026/6005)

Table des matières

Page

1. CONTEXTE ET OBJECTIFS 3
2. VUE D'ENSEMBLE DES OBLIGATIONS DE TRANSPARENCE ET DES THÈMES HORIZONTAUX 3
   - 2.1. Vue d'ensemble des obligations de transparence prévues à l'article 50 du règlement sur l'IA 3
   - 2.2. Justification et objectifs 4
   - 2.3. Acteurs responsables au titre de l'article 50 du règlement sur l'IA 5
   - 2.4. Exclusion du champ d'application du règlement sur l'IA 7
     - 2.4.1. Déployeurs agissant dans le cadre d'une activité strictement personnelle à caractère non professionnel 7
     - 2.4.2. Recherche et développement 8
     - 2.4.3. Systèmes d'IA publiés dans le cadre de licences libres et ouvertes 8
   - 2.5. Interaction avec les pratiques interdites, les exigences applicables aux systèmes d'IA à haut risque et la maîtrise de l'IA 9
   - 2.6. Interaction avec les règles applicables aux modèles et/ou systèmes d'IA à usage général 9
3. ARTICLE 50, PARAGRAPHE 1, DU RÈGLEMENT SUR L'IA: TRANSPARENCE POUR LES SYSTÈMES D'IA INTERACTIVE 9
   - 3.1. Principaux éléments, concepts et obligation(s) de transparence s'y rapportant 10
     - 3.1.1. Systèmes d'IA destinés à interagir directement avec des personnes physiques 10
     - 3.1.2. Obligation d'information au titre de l'article 50, paragraphe 1, du règlement sur l'IA 12
   - 3.2. Exceptions à l'obligation d'information prévue à l'article 50, paragraphe 1, du règlement sur l'IA 14
     - 3.2.1. Exception en cas d'interaction manifeste avec un système d'IA 14
     - 3.2.2. Exception pour les systèmes d'IA dont l'utilisation est autorisée par la loi à des fins répressives 16
   - 3.3. Interaction avec d'autres actes juridiques de l'Union 17
4. ARTICLE 50, PARAGRAPHE 2, DU RÈGLEMENT SUR L'IA: MARQUAGE ET DÉTECTION DES CONTENUS GÉNÉRÉS OU MANIPULÉS PAR UNE IA 18
   - 4.1. Principaux éléments et concepts de l'article 50, paragraphe 2, du règlement sur l'IA 18
     - 4.1.1. Systèmes d'IA générant ou manipulant des contenus de synthèse 18
     - 4.1.2. Formats des contenus de synthèse relevant du champ d'application 19
     - 4.1.3. Contenus ne relevant pas du champ d'application de l'article 50, paragraphe 2, du règlement sur l'IA 20
   - 4.2. L'obligation de marquage et de détection prévue à l'article 50, paragraphe 2, du règlement sur l'IA 21
     - 4.2.1. L'obligation de marquage 21
     - 4.2.2. L'obligation de détection 22
     - 4.2.3. Respect des exigences relatives aux solutions techniques: efficacité, interopérabilité, solidité et fiabilité 22
   - 4.3. Exceptions aux obligations prévues à l'article 50, paragraphe 2, du règlement sur l'IA 24
   - 4.4. Interaction avec d'autres actes juridiques de l'Union 25

5. ARTICLE 50, PARAGRAPHE 3, DU RÈGLEMENT SUR L'IA: SYSTÈMES DE RECONNAISSANCE DES ÉMOTIONS ET SYSTÈMES DE CATÉGORISATION BIOMÉTRIQUE 26

5.1. Principaux composants, concepts et obligations de transparence s'y rapportant au titre de l'article 50, paragraphe 3, du règlement sur l'IA 27

5.1.1. La notion de système de reconnaissance des émotions 27

5.1.2. La notion de système de catégorisation biométrique 27

5.1.3. L'obligation d'information prévue à l'article 50, paragraphe 3, du règlement sur l'IA 27

5.2. Exception à des fins répressives 28

5.3. Interaction avec d'autres actes juridiques de l'Union 28

6. ARTICLE 50, PARAGRAPHE 4, DU RÈGLEMENT SUR L'IA: ÉTIQUETAGE DES HYPERTRUCAGES ET DE CERTAINES PUBLICATIONS TEXTUELLES 28

6.1. Principaux composants, concepts et obligation(s) de transparence s'y rapportant pour les hypertrucages au titre de l'article 50, paragraphe 4, du règlement sur l'IA 28

6.1.1. La notion d'«hypertrucage» 29

6.1.2. L'obligation d'information prévue à l'article 50, paragraphe 4, premier alinéa, du règlement sur l'IA 31

6.1.3. Transparence des hypertrucages artistiques, créatifs, satiriques, fictifs ou analogues 32

6.1.4. Exception à des fins répressives 34

6.1.5. Interaction avec d'autres actes juridiques de l'Union 34

6.2. Principaux composants, concepts et obligations de transparence s'y rapportant pour les textes générés ou manipulés par une IA au titre de l'article 50, paragraphe 4, du règlement sur l'IA 35

6.2.1. Texte publié dans le but d'informer le public sur des questions d'intérêt public 36

6.2.2. L'obligation d'information prévue à l'article 50, paragraphe 4, deuxième alinéa, du règlement sur l'IA 37

6.2.3. Exception à l'obligation de transparence pour les textes faisant l'objet d'un examen humain ou d'un contrôle éditorial et dont la publication relève de la responsabilité éditoriale d'une personne physique ou morale 37

6.2.4. Exception à des fins répressives 38

6.2.5. Interaction avec d'autres actes juridiques de l'Union 39

7. EXIGENCES HORIZONTALES APPLICABLES AUX INFORMATIONS FOURNIES EN VERTU DE L'ARTICLE 50, PARAGRAPHE 5, DU RÈGLEMENT SUR L'IA 39

7.1. Fourniture d'informations de manière claire et reconnaissable 39

7.2. Première interaction ou première exposition 40

7.3. Conformité avec les exigences applicables en matière d'accessibilité 40

8. APPLICATION DE L'ARTICLE 50 DU RÈGLEMENT SUR L'IA 40

8.1. Effets de l'adhésion à un code de bonne pratique jugé approprié 41

8.2. Autorités de surveillance du marché 42

8.3. Sanctions 42

8.4. Entrée en application 42

9. RÉVISION ET MISE À JOUR DES LIGNES DIRECTRICES 43

1. **CONTEXTE ET OBJECTIFS**

(1) Le règlement (UE) 2024/1689 du Parlement européen et du Conseil du 13 juin 2024 établissant des règles harmonisées concernant l'intelligence artificielle et modifiant certains règlements (ci-après le «règlement sur l'IA») est entré en vigueur le 1er août 2024 ([1]). Le règlement sur l'IA établit des règles harmonisées concernant la mise sur le marché, la mise en service et l'utilisation de systèmes d'intelligence artificielle (IA) dans l'Union ([2]). Son objectif est de promouvoir l'innovation dans le domaine de l'IA et l'adoption d'une IA digne de confiance, tout en garantissant un niveau élevé de protection de la santé, de la sécurité et des droits fondamentaux dans l'Union, ainsi que de la démocratie et de l'état de droit.

(2) Le règlement sur l'IA suit une approche fondée sur les risques, en classant les systèmes d'IA en quatre catégories de risque différentes, l'une d'elles concernant les systèmes d'IA présentant des risques en matière de transparence qui sont soumis aux obligations établies à l'article 50 dudit règlement. Ces obligations de transparence s'appliquent deux ans après l'entrée en vigueur du règlement sur l'IA, c'est-à-dire à partir du 2 août 2026.

(3) Conformément à l'article 96, paragraphe 1, point d), du règlement sur l'IA, les présentes lignes directrices de la Commission sont publiées dans le but de servir d'orientations pratiques pour aider les autorités compétentes, ainsi que les fournisseurs et les déployeurs de systèmes d'IA, à garantir le respect des obligations de transparence prévues à l'article 50 dudit règlement de manière cohérente, efficace, proportionnée et uniforme.

(4) L'élaboration des présentes lignes directrices s'est appuyée sur les contributions de diverses parties prenantes recueillies au cours d'une vaste consultation organisée par la Commission et sur les contributions des États membres au sein du Comité IA, ainsi que sur les contributions des parties prenantes au projet de lignes directrices qui a été publié pour consultation.

(5) Les présentes lignes directrices ne sont pas contraignantes. En définitive, seule la Cour de justice de l'Union européenne (CJUE) peut donner une interprétation faisant autorité du règlement sur l'IA.

2. **VUE D'ENSEMBLE DES OBLIGATIONS DE TRANSPARENCE ET DES THÈMES HORIZONTAUX**

2.1. **Vue d'ensemble des obligations de transparence prévues à l'article 50 du règlement sur l'IA**

(6) L'article 50 du règlement sur l'IA comprend quatre obligations de transparence, chacune s'appliquant à différents types de systèmes d'IA ou à leurs sorties.

Disposition	Type de système d'IA/sortie	Obligation de transparence	Exceptions ou régimes spéciaux
Article 50, paragraphe 1	Systèmes d'IA interagissant directement avec des personnes physiques	Les fournisseurs doivent développer et concevoir le système d'IA de manière à ce que les personnes physiques concernées soient informées qu'elles interagissent avec un système d'IA.	Exceptions 1) si l'origine artificielle de l'interaction est claire ou 2) si la loi autorise l'utilisation du système à des fins de prévention ou de détection des infractions pénales, d'enquêtes ou de poursuites en la matière, sauf si le système est mis à la disposition du public pour permettre le signalement d'une infraction pénale.

([1]) Règlement (UE) 2024/1689 du Parlement européen et du Conseil du 13 juin 2024 établissant des règles harmonisées concernant l'intelligence artificielle (règlement sur l'intelligence artificielle) (JO L, 2024/1689, 12.7.2024, ELI: http://data.europa.eu/eli/reg/2024/1689/oj).

([2]) Article 1er du règlement sur l'IA.

Disposition	Type de système d'IA/sortie	Obligation de transparence	Exceptions ou régimes spéciaux
Article 50, paragraphe 2	Systèmes d'IA générant ou manipulant des contenus de synthèse de type image, vidéo, audio ou texte	Les fournisseurs doivent veiller à ce que les sorties du système d'IA soient marquées dans un format lisible par machine et identifiables comme ayant été générées ou manipulées par une IA avec des solutions techniques efficaces, interopérables, solides et fiables.	Exceptions si 1) le système d'IA remplit une fonction d'assistance pour la mise en forme standard ou ne modifie pas de manière substantielle les données d'entrée fournies ou leur sémantique, ou 2) la loi autorise son utilisation à des fins de prévention ou de détection des infractions pénales, d'enquêtes ou de poursuites en la matière.
Article 50, paragraphe 3	Systèmes d'IA de reconnaissance des émotions ou de catégorisation biométrique	Les déployeurs doivent informer les personnes physiques exposées au système d'IA du fonctionnement de ce dernier.	Exception si la loi autorise l'utilisation du système d'IA à des fins de prévention ou de détection des infractions pénales ou d'enquêtes en la matière.
Article 50, paragraphe 4	Systèmes d'IA générant ou manipulant des hypertrucages ou des textes publiés pour informer le public sur des questions d'intérêt public	Les déployeurs doivent indiquer que les contenus ont été générés ou manipulés par une IA.	Exceptions si 1) la loi autorise l'utilisation du système à des fins de prévention ou de détection des infractions pénales, d'enquêtes ou de poursuites en la matière ou 2) si la publication du texte a fait l'objet d'un examen humain ou de contrôle éditorial et qu'une personne physique ou morale assume la responsabilité éditoriale. Un régime spécial de divulgation s'applique aux hypertrucages faisant partie d'œuvres ou de programmes artistiques, créatifs, fictifs, satiriques ou analogues.

(7) Chacune de ces quatre obligations est examinée aux sections 3 à 6 des présentes lignes directrices. En outre, l'article 50, paragraphe 5, du règlement sur l'IA prévoit que les informations visées dans ses paragraphes 1 à 4 sont fournies aux personnes physiques concernées de manière claire et reconnaissable au plus tard au moment de la première interaction ou de la première exposition et sont conformes aux exigences applicables en matière d'accessibilité. La section 7 donne des précisions sur ces exigences.

(8) Les différentes obligations de transparence établies à l'article 50 du règlement sur l'IA peuvent s'appliquer de manière cumulative à (la sortie d')un seul système d'IA, engageant éventuellement la responsabilité de différents acteurs (fournisseurs ou déployeurs). Par exemple, si un système d'IA génère des images dans le cadre d'une interaction directe avec une personne physique, le fournisseur du système peut être tenu de se conformer à la fois aux paragraphes 1 et 2 de l'article 50 du règlement sur l'IA. Lorsque ce système est également en mesure de générer des images constituant des hypertrucages ou des textes publiés sur des questions d'intérêt public, le déployeur peut en outre être tenu de se conformer à l'article 50, paragraphe 4, du règlement sur l'IA.

2.2. **Justification et objectifs**

(9) Comme expliqué aux considérants 132 à 136 du règlement sur l'IA, l'objectif des obligations de transparence est de réduire les risques d'usurpation d'identité, de tromperie, de désinformation, de manipulation à grande échelle et de fraude et d'atténuer les incidences négatives potentielles sur les processus démocratiques et la confiance de la société causées par des contenus générés ou manipulés par l'IA ou par des interactions avec celle-ci. Informer les personnes de l'origine artificielle de l'interaction et du contenu les aidera à prendre des décisions éclairées, tout en contribuant à préserver l'intégrité de l'écosystème informationnel et la confiance en celui-ci. La transparence est non seulement une exigence de conformité technique, mais aussi une condition préalable pour permettre aux personnes physiques de reconnaître les interactions et les contenus faisant intervenir l'IA ou l'utilisation de systèmes de reconnaissance des émotions ou de catégorisation biométrique. Garantir une utilisation transparente et en toute confiance des systèmes d'IA interactive et générative permettra également de veiller au respect de l'autonomie personnelle et des valeurs européennes, de la dignité humaine et d'autres droits fondamentaux (par exemple, la liberté de pensée, le respect de la vie privée, l'égalité et la non-discrimination, la protection des consommateurs, la propriété intellectuelle, etc.), tout en préservant la liberté d'expression et la crédibilité des contenus médiatiques.

2.3. **Acteurs responsables au titre de l'article 50 du règlement sur l'IA**

(10) Conformément à l'article 3, point 3), du règlement sur l'IA, les **fournisseurs** sont les personnes physiques ou morales, les autorités publiques, les agences ou les autres organismes qui développent ou font développer des systèmes d'IA et les mettent sur le marché de l'Union, ou les mettent en service sous leur propre nom ou leur propre marque, à titre onéreux ou gratuit. Les fournisseurs doivent respecter les obligations de transparence pertinentes prévues à l'article 50, paragraphes 1 et 2, du règlement sur l'IA qui leur sont applicables, qu'ils soient établis ou situés dans l'Union ou dans un pays tiers ([3]). Les fournisseurs établis ou situés en dehors de l'Union sont également soumis au règlement sur l'IA si les sorties de leur système d'IA sont utilisées dans l'Union ([4]). Toutefois, une utilisation en aval fortuite, imprévisible ou non autorisée ne devrait pas entraîner à elle seule l'application des obligations aux fournisseurs de systèmes d'IA interactive ou générative établis ou situés dans un pays tiers dont les systèmes ne sont pas mis sur le marché ou mis en service dans l'Union, sans préjudice des pouvoirs des autorités de surveillance du marché compétentes de considérer leurs sorties non conformes au règlement sur l'IA.

(11) Les fournisseurs doivent veiller à ce que leurs systèmes d'IA interactive ou générative intègrent la transparence dans leur conception et leur fonctionnement et à ce que toutes les obligations de transparence pertinentes prévues à l'article 50, paragraphes 1, 2 et 5, du règlement sur l'IA soient remplies au plus tard lorsque ces systèmes sont mis sur le marché ou mis en service. Toutes les personnes physiques qui interagissent avec le système d'IA ou sont exposées à des sorties de ce système qui relèvent du champ d'application doivent être informées de manière efficace et cohérente tout au long du cycle de vie du système d'IA, tel qu'il est mis sur le marché ou mis en service par le fournisseur (quels que soient les déployeurs ou les autres utilisateurs du système).

> Par exemple, une entreprise fournit une application d'IA générative ou interactive (par exemple, un dialogueur, un générateur d'images, un agent d'IA) sur le marché de l'Union sous son propre nom ou sa propre marque à des consommateurs, à des utilisateurs professionnels ou à d'autres entités juridiques qui peuvent utiliser directement le système d'IA à des fins d'interaction ou de production de contenus (y compris en intégrant le système dans leurs processus existants et en l'utilisant en conséquence sans qu'il soit nécessaire de le modifier ou de le compléter par des composants supplémentaires). L'entreprise est un fournisseur responsable du respect des obligations de transparence prévues à l'article 50, paragraphes 1 et/ou 2, du règlement sur l'IA, que le système d'IA soit fourni gratuitement ou à titre onéreux et que le fournisseur soit établi ou situé dans l'Union ou dans un pays tiers.
>
> Un autre exemple de fournisseur est une entreprise ou une autre organisation (par exemple, une autorité publique) qui a développé un système d'IA interactive (par exemple, un dialogueur) en interne et le met en service dans l'Union pour son propre usage et sous son propre nom ou sa propre marque.
>
> Si une entreprise reprend un système d'IA générative existant mis sur le marché par un autre fournisseur et modifie ce système (par exemple avec de nouvelles données d'entraînement), qu'elle met ensuite en service sous son propre nom ou sa propre marque, cette entreprise devient un fournisseur du nouveau système sans préjudice de la responsabilité du fournisseur du système d'IA initial pour ce dernier.

(12) Les déployeurs sont des personnes physiques ou morales, des autorités publiques, des agences ou d'autres organismes utilisant sous leur propre autorité un système d'IA sauf lorsque ce système est utilisé dans le cadre d'une activité personnelle à caractère non professionnel ([5]). L'«autorité» sur un système d'IA devrait s'entendre comme le fait d'assumer la responsabilité de la décision de déployer le système et de la manière dont ce dernier est effectivement utilisé (y compris ses sorties). Elle ne passe pas nécessairement par un contrôle technique sur le fonctionnement du système d'IA, pour autant que le déployeur prenne la décision quant aux finalités et à la manière d'utiliser le système d'IA (y compris dans les flux de travail décentralisés et les structures de groupes de société). Les déployeurs intervenant dans des chaînes de valeur complexes concernant la production et la distribution de contenus devraient prendre des mesures proportionnées pour veiller à ce que l'étiquetage des contenus qu'ils ont mis en œuvre conformément à l'article 50, paragraphe 4, du règlement sur l'IA soit affiché de manière claire et reconnaissable pour le public ciblé et prévisible au moment de la première exposition conformément à l'article 50, paragraphe 5, du règlement sur l'IA [par exemple, au moyen de conditions contractuelles avec les partenaires de distribution, de paramètres d'expérience utilisateur (UX) et d'interfaces à afficher].

([3]) Article 2, paragraphe 1, point b), et article 3, points 3), 9) et 11), du règlement sur l'IA. Pour de plus amples informations sur les notions de mise sur le marché, de mise en service et d'utilisation, voir les lignes directrices de la Commission sur les pratiques interdites en matière d'intelligence artificielle, C(2025) 5052, et la communication de la Commission – le «Guide bleu» relatif à la mise en œuvre de la réglementation de l'UE sur les produits 2022, point 2 (JO C 247 du 29.6.2022, p. 1).

([4]) Article 2, paragraphe 1, point c), et considérant 22 du règlement sur l'IA.

([5]) Article 3, point 4, du règlement sur l'IA.

(13) Les déployeurs relèvent du champ d'application des obligations de transparence prévues à l'article 50, paragraphes 3 et 4, du règlement sur l'IA si leur lieu d'établissement ou leur localisation se situe dans l'Union, ou s'ils sont établis ou situés dans un pays tiers et que les sorties de leur système d'IA sont utilisées dans l'Union ([6]). Dans ce dernier cas, les obligations de transparence s'appliquent aux entités situées ou établies en dehors de l'Union lorsque le déployeur prévoit lui-même la diffusion et l'utilisation des sorties de l'IA dans l'Union (c'est-à-dire en ordonnant ou en autorisant la distribution au sein de l'Union, y compris en publiant des hypertrucages sur l'internet accessible à l'échelle mondiale). Toutefois, les déployeurs de pays tiers ne sont pas liés par les obligations de transparence lorsque le contenu du système d'IA atteint des publics dans l'Union par des canaux qui sont imprévisibles et échappent à leur contrôle.

(14) Lorsque le déployeur d'un système d'IA est une personne morale sous l'autorité de laquelle le système est utilisé (par exemple, une entreprise de publicité), les employés qui agissent selon les instructions et sous le contrôle de cette personne morale (par exemple, les animateurs numériques, les concepteurs de sites web, les créateurs de contenus ou les journalistes) ne devraient pas être considérés comme des déployeurs distincts de ce système. Une personne morale reste un déployeur même si elle fait intervenir des tiers (par exemple, des sous-traitants ou des prestataires indépendants) dans l'exploitation du système en son nom et sous sa responsabilité et son contrôle.

Par exemple, un organe de presse établi dans l'Union, dont les rédacteurs utilisent un système d'IA pour leur traitement par écrit des événements d'actualité au moyen de publications textuelles générées par une IA sur des questions d'intérêt public puis publiées sur un site web accessible à l'échelle mondiale, est un déployeur relevant du champ d'application du règlement sur l'IA. De même, une entreprise établie dans un pays tiers qui utilise un système d'IA pour générer un hypertrucage d'une célébrité figurant dans une publicité affichée dans l'Union est également un déployeur relevant du champ d'application du règlement sur l'IA.

En revanche, une entreprise qui se contente de charger une agence de publicité de produire une publicité, sans prendre de décisions et sans exercer de contrôle sur le fait que l'agence de publicité utilise ou non une IA et sur la manière dont elle l'utilise dans le processus de production, n'est pas un déployeur.

(15) Les opérateurs (par exemple, les fournisseurs et les déployeurs) peuvent remplir simultanément plus d'un rôle en ce qui concerne un système d'IA. Par exemple, une entreprise pourrait développer un système interne d'IA générative et l'utiliser pour produire des hypertrucages. Dans ce cas, elle sera qualifiée à la fois de fournisseur et de déployeur qui doit garantir le respect des obligations de marquage lisible par machine prévues à l'article 50, paragraphe 2, et des exigences relatives à un étiquetage clair et perceptible des hypertrucages prévues à l'article 50, paragraphe 4, du règlement sur l'IA.

(16) À l'inverse, d'autres acteurs (tels que les fournisseurs de services d'hébergement, y compris les plateformes en ligne, ou les radiodiffuseurs) dont le rôle se limite à diffuser ou à transmettre des contenus générés ou manipulés par une IA créés par des tiers, ou qui reçoivent des contenus générés ou manipulés par une IA ou y sont exposés sans avoir autorité sur l'utilisation du système d'IA aux fins de la production ou de la manipulation de contenus par IA, ne sont pas des déployeurs au sens du règlement sur l'IA. Sans être responsables au titre dudit règlement, ces acteurs peuvent néanmoins jouer un rôle important dans la chaîne de valeur afin de préserver l'intégrité de l'écosystème informationnel et la confiance en celui-ci. Ils sont donc vivement encouragés à préserver le marquage et l'étiquetage mis en œuvre conformément à l'article 50 du règlement sur l'IA et à appliquer des mesures proportionnées et techniquement réalisables pour permettre aux personnes physiques de détecter les contenus générés ou manipulés par une IA, sans préjudice de la responsabilité des fournisseurs et des déployeurs de systèmes d'IA générative ou de toute obligation ou responsabilité au titre du règlement (UE) 2022/2065 (règlement sur les services numériques ou «DSA») ([7]) ou d'autres dispositions applicables du droit de l'Union ou du droit national.

([6]) Article 2, paragraphe 1, points b) et c), du règlement sur l'IA.

([7]) Règlement (UE) 2022/2065 du Parlement européen et du Conseil du 19 octobre 2022 relatif à un marché unique des services numériques; JO L 277 du 27.10.2022, p. 1.

(17) Les autres acteurs agissant à titre professionnel et diffusant des contenus (même s'ils ne sont pas directement concernés par l'article 50 du règlement sur l'IA) sont également encouragés à prendre des mesures proportionnées et appropriées, de sorte que les personnes physiques qui se retrouvent finalement exposées à ces contenus soient effectivement informées et en mesure de percevoir que ceux-ci ont été générés ou manipulés par une IA. En outre, ces parties peuvent être tenues d'étiqueter les hypertrucages ou les publications textuelles sur des questions d'intérêt public en vertu d'autres règles juridiques ou déontologiques applicables (par exemple, un radiodiffuseur étiquetant et montrant, dans le cadre d'une émission télévisée, un hypertrucage créé par un particulier). Toutefois, le simple fait qu'un acteur étiquette un contenu ou utilise un contenu étiqueté ne signifie pas que cet acteur est un «déployeur» au sens du règlement sur l'IA. Le règlement sur l'IA n'étend ni ne modifie la responsabilité juridique de ces acteurs en vertu des cadres juridiques applicables s'ils ne sont pas qualifiés en même temps de fournisseurs ou de déployeurs de systèmes d'IA.

2.4. **Exclusion du champ d'application du règlement sur l'IA**

(18) L'article 2 du règlement sur l'IA prévoit un certain nombre d'exclusions générales du champ d'application dudit règlement qui ont déjà été clarifiées dans les lignes directrices de la Commission sur les pratiques interdites en matière d'intelligence artificielle ([8]). La présente section ne traitera donc que des aspects pertinents pour les obligations de transparence prévues à l'article 50 du règlement sur l'IA.

2.4.1. *Déployeurs agissant dans le cadre d'une activité strictement personnelle à caractère non professionnel*

(19) Conformément à l'article 2, paragraphe 10, du règlement sur l'IA, ledit règlement «ne s'applique pas aux obligations incombant aux déployeurs qui sont des personnes physiques utilisant des systèmes d'IA dans le cadre d'une activité strictement personnelle à caractère non professionnel» ([9]). Toute activité dont des personnes physiques tirent régulièrement un avantage économique ou par laquelle elles participent d'une autre manière à une activité professionnelle, commerciale ou indépendante doit être considérée comme une activité «professionnelle». La précision «strictement personnelle» qualifie l'activité non professionnelle, ce qui signifie que la personne doit agir à la fois à titre personnel et à titre non professionnel pour relever du champ d'application de l'exception. Toute utilisation par des personnes physiques agissant pour le compte ou sous l'autorité d'un déployeur agissant à titre professionnel relèvera du champ d'application des obligations de transparence du déployeur professionnel énoncées à l'article 50, paragraphes 3 et 4, du règlement sur l'IA.

Exemples d'utilisations personnelles non professionnelles de systèmes d'IA exclues du champ d'application:

— particuliers utilisant un système d'IA pour produire des hypertrucages des membres de leur ménage;

— particuliers utilisant un système d'IA à titre personnel pour produire du texte sur divers sujets (par exemple, les élèves qui utilisent l'IA pour écrire leurs devoirs);

— particuliers utilisant un système d'IA à titre personnel pour produire des hypertrucages diffusés sur les médias sociaux.

(20) L'exclusion prévue à l'article 2, paragraphe 10, du règlement sur l'IA ne s'applique qu'en ce qui concerne les obligations des déployeurs lorsqu'ils utilisent le système pour des activités strictement personnelles à caractère non professionnel. Le système en tant que tel reste dans le champ d'application du règlement sur l'IA en ce qui concerne les obligations des fournisseurs qui mettent le système sur le marché ou le mettent en service et les obligations des déployeurs qui utilisent le système à des fins professionnelles. L'exception est également sans préjudice de l'application d'autres dispositions pertinentes du droit de l'Union et du droit national.

([8]) Voir la section 2.5 des lignes directrices de la Commission sur les pratiques interdites en matière d'intelligence artificielle au sens du règlement (UE) 2024/1689 (règlement sur l'IA), Bruxelles, 29.7.2025, C(2025) 5052 final.

([9]) La définition du terme «déployeur» exclut également les utilisateurs participant à de telles activités (voir la section 2.3 ci-dessus).

Par exemple, le fournisseur du système d'IA utilisé pour produire des hypertrucages à afficher sur une carte de Noël personnelle devrait marquer ces hypertrucages d'une manière lisible par machine, conformément à l'obligation de transparence prévue à l'article 50, paragraphe 2, du règlement sur l'IA. Parallèlement, l'utilisation du système par des personnes physiques dans le cadre d'activités strictement personnelles et non professionnelles ne relèvera pas de l'obligation incombant aux déployeurs prévue à l'article 50, paragraphe 4, dudit règlement.

2.4.2. *Recherche et développement*

(21) L'article 2, paragraphe 6, du règlement sur l'IA prévoit une exclusion pour les systèmes d'IA ou les modèles d'IA spécifiquement développés et mis en service uniquement à des fins de recherche et développement scientifiques, et pour leurs sorties. Cette exclusion est sans préjudice des normes éthiques applicables en matière de recherche, de la législation de l'UE en matière de protection des données et d'autres actes législatifs. Dans le contexte de l'article 50 du règlement sur l'IA, cette exclusion s'applique non seulement aux systèmes d'IA interactive, mais aussi aux sorties des systèmes d'IA générative (y compris les contenus de type hypertrucage) utilisées dans le cadre de la recherche scientifique. Toutefois, si ces systèmes sont mis en service ou si leurs sorties sont utilisées à d'autres finalités qu'aux seules fins de la recherche et du développement scientifiques, les obligations de transparence pertinentes prévues à l'article 50 du règlement sur l'IA devraient quand même être respectées.

Par exemple, si des chercheurs développent et mettent sur le marché ou mettent en service un système d'IA interactive et souhaitent examiner si les personnes physiques sont en mesure de faire la distinction entre les systèmes d'IA et les êtres humains lorsqu'elles interagissent oralement avec ce système d'IA, ils ne sont pas tenus de mettre en œuvre les obligations de transparence correspondantes prévues à l'article 50, paragraphes 1 et 2, du règlement sur l'IA. Toutefois, si le même système d'IA interactive est également mis en service à des fins autres que la recherche et le développement, les obligations de transparence pertinentes prévues à l'article 50 du règlement sur l'IA devraient être respectées.

(22) L'article 2, paragraphe 8, exclut en outre les activités de recherche, d'essai et de développement relatives aux systèmes d'IA ou aux modèles d'IA avant leur mise sur le marché ou leur mise en service. Les essais en conditions réelles (dans les bacs à sable réglementaires de l'IA ou en dehors de ces derniers) ne relèvent pas de cette exclusion. Par ailleurs, une fois qu'un système d'IA est mis sur le marché ou mis en service à la suite de ces activités de recherche et d'essai, il devrait néanmoins être conforme aux obligations de transparence pertinentes prévues à article 50 du règlement sur l'IA ([10]).

2.4.3. *Systèmes d'IA publiés dans le cadre de licences libres et ouvertes*

(23) Conformément à l'article 2, paragraphe 12, du règlement sur l'IA, les systèmes d'IA publiés dans le cadre de licences libres et ouvertes ne relèvent pas du champ d'application dudit règlement, pour autant qu'ils ne soient pas visés par les interdictions prévues à l'article 5 du règlement sur l'IA ou les obligations de transparence prévues à l'article 50 du règlement sur l'IA, ou qu'ils ne soient pas classés à haut risque ([11]). Cela signifie que les fournisseurs et les déployeurs de systèmes d'IA ouverts relevant du champ d'application de l'article 50 du règlement sur l'IA doivent encore veiller au respect de leurs obligations respectives en matière de transparence.

(24) En revanche, les fournisseurs de composants d'IA libres et ouverts qui englobent les logiciels, les données, les modèles, les outils, les services ou les processus à personnaliser et à intégrer dans un système d'IA et qui ne constituent pas en eux-mêmes (en partie ou en totalité) un système d'IA ne sont pas soumis aux obligations de transparence prévues à l'article 50 du règlement sur l'IA. Les fournisseurs de ces systèmes sont néanmoins encouragés à concevoir ces modèles, composants et services de la manière décrite au point 27 ci-dessous, dans la mesure où cela est proportionné et réalisable.

([10]) Considérant 25 du règlement sur l'IA.

([11]) Le considérant 102 du règlement sur l'IA explique que la publication de logiciels et de données dans le cadre d'une licence libre et ouverte leur permet d'«être partagés librement et [...] permet aux utilisateurs de librement consulter, utiliser, modifier et redistribuer ces logiciels et données ou leurs versions modifiées».

2.5. Interaction avec les pratiques interdites, les exigences applicables aux systèmes d'IA à haut risque et la maîtrise de l'IA

(25) Le considérant 137 du règlement sur l'IA précise que le respect des obligations de transparence prévues à l'article 50 du règlement sur l'IA pour un système d'IA donné ne peut pas être interprété comme indiquant que l'utilisation de ce système d'IA ou de ses sorties est licite en vertu dudit règlement. En particulier, les systèmes d'IA soumis aux obligations énoncées à l'article 50 du règlement sur l'IA peuvent relever de l'article 5 dudit règlement et donc être interdits (par exemple, l'utilisation de systèmes de reconnaissance des émotions sur le lieu de travail ou dans les établissements d'enseignement) ([12]). En outre, les systèmes d'IA relevant du champ d'application de l'article 50 du règlement sur l'IA peuvent également être classés à haut risque ([13]) et être soumis aux exigences et obligations applicables à ces systèmes. Le respect de l'article 50 du règlement sur l'IA ne remplace, ne limite et n'affecte pas les obligations découlant d'autres dispositions du règlement sur l'IA ([14]). Dans le même temps, les systèmes d'IA peuvent également relever du champ d'application de l'article 50 du règlement sur l'IA sans être classés à haut risque en vertu de l'article 6 dudit règlement. Les exigences en matière de maîtrise de l'IA prévues à l'article 4 du règlement sur l'IA s'appliquent à la fois aux fournisseurs et aux déployeurs de systèmes d'IA relevant du champ d'application de l'article 50 du règlement sur l'IA.

2.6. Interaction avec les règles applicables aux modèles et/ou systèmes d'IA à usage général

(26) L'article 50 du règlement sur l'IA s'applique à certains systèmes d'IA, qui comprennent les systèmes d'IA à usage général ([15]), tels que les dialogueurs qui favorisent l'interaction directe avec des personnes physiques ou les systèmes d'IA capables de générer ou de manipuler des contenus de synthèse (y compris les hypertrucages) ([16]).

(27) L'article 50 du règlement sur l'IA ne s'applique pas explicitement aux modèles d'IA à usage général. Toutefois, si un système d'IA interactive ou générative relevant du champ d'application de l'article 50, paragraphes 1 ou 2, du règlement sur l'IA repose sur un modèle d'IA à usage général fourni par le même fournisseur, les mesures de transparence énoncées dans les présentes lignes directrices pourraient également être mises en œuvre au niveau du modèle (voir également la section 4 ci-dessous). En outre, bien qu'ils ne relèvent pas du champ d'application de l'article 50 du règlement sur l'IA, les fournisseurs d'autres modèles d'IA à usage général et d'autres modèles d'IA générative sont également encouragés à mettre en œuvre des mesures de transparence appropriées au niveau du modèle pour l'identification des systèmes d'IA interactive et le marquage des contenus générés et manipulés par une IA. Ces mesures sont particulièrement utiles pour faciliter le respect, par les fournisseurs en aval de systèmes d'IA, des obligations qui leur incombent en vertu de l'article 50, paragraphes 1 et 2, du règlement sur l'IA ([17]) et pour contribuer à préserver l'intégrité de l'écosystème informationnel et la confiance en celui-ci. Pour les modèles d'IA à usage général présentant un risque systémique, les mesures de sécurité mises en œuvre au niveau du modèle pour l'identification des interactions avec l'IA et le marquage des contenus générés ou manipulés par une IA peuvent, le cas échéant, faire partie d'un ensemble plus large de mesures d'atténuation que les fournisseurs peuvent prendre dans le cadre de leur obligation d'atténuer les risques systémiques conformément à l'article 55, paragraphe 1, point b), du règlement sur l'IA.

3. ARTICLE 50, PARAGRAPHE 1, DU RÈGLEMENT SUR L'IA: TRANSPARENCE POUR LES SYSTÈMES D'IA INTERACTIVE

(28) L'article 50, paragraphe 1, du règlement sur l'IA s'adresse aux fournisseurs de systèmes d'IA interagissant directement avec des personnes physiques, qui doivent concevoir et développer leurs systèmes de manière que les personnes physiques concernées soient informées qu'elles interagissent avec un système d'IA. Comme expliqué au considérant 132 du règlement sur l'IA, l'objectif de cette obligation d'information est de permettre à ces personnes physiques de prendre des décisions éclairées concernant les sorties du système, d'éviter une dépendance excessive à l'égard de ces systèmes et de permettre à ces personnes physiques d'ajuster en conséquence leur degré de confiance dans les contenus générés par ces systèmes et les interactions avec ces derniers.

([12]) Pour de plus amples informations sur l'article 5 du règlement sur l'IA, voir les lignes directrices de la Commission sur les pratiques interdites en matière d'intelligence artificielle au sens du règlement (UE) 2024/1689 (règlement sur l'IA), C(2025) 5052 final.

([13]) Voir l'article 6 du règlement sur l'IA.

([14]) En particulier, l'article 26, paragraphe 11, du règlement sur l'IA impose une obligation de transparence aux déployeurs de systèmes d'IA à haut risque qui prennent des décisions ou facilitent les prises de décision concernant l'information des personnes physiques sur le fait qu'elles sont soumises à l'utilisation du système d'IA à haut risque. Ces obligations de transparence sont complémentaires de celles prévues à l'article 50 du règlement sur l'IA, étant donné qu'elles concernent des informations différentes à fournir aux personnes physiques, mais les informations requises peuvent être combinées dans une même notification.

([15]) L'article 3, point 66), du règlement sur l'IA définit un système d'IA à usage général comme un système d'IA qui est fondé sur un modèle d'IA à usage général et qui a la capacité de répondre à diverses finalités, tant pour une utilisation directe que pour une intégration dans d'autres systèmes d'IA.

([16]) L'article 50, paragraphe 2, du règlement sur l'IA contient une référence explicite aux systèmes d'IA à usage général.

([17]) Considérant 133 du règlement sur l'IA.

3.1. **Principaux éléments, concepts et obligation(s) de transparence s'y rapportant**

(29) L'obligation de transparence prévue à l'article 50, paragraphe 1, du règlement sur l'IA s'applique aux fournisseurs de systèmes d'IA lorsque le système est destiné à interagir directement avec des personnes physiques et que cette interaction ne relève d'aucune des exceptions suivantes: 1) il ressort clairement du point de vue d'une personne physique normalement informée et raisonnablement attentive et avisée qu'elle interagit avec un système d'IA, ou 2) l'utilisation du système est autorisée par la loi à des fins de prévention ou de détection des infractions pénales, d'enquêtes ou de poursuites en la matière, sauf si le système est mis à la disposition du public pour permettre le signalement d'une infraction pénale.

3.1.1. *Systèmes d'IA destinés à interagir directement avec des personnes physiques*

(30) Quatre éléments doivent être présents pour qu'un système relève du champ d'application de l'article 50, paragraphe 1, du règlement sur l'IA: i) le système doit être un système d'IA, ii) destiné à interagir, iii) directement, iv) avec des personnes physiques.

i) **Un système d'IA:** le système doit réunir les éléments de la définition du système d'IA ([18]), qui exclut les simples mécanismes de réponse automatisée ne reposant pas sur l'IA (par exemple, les messages automatiques d'absence traditionnels ou les réponses rapides aux messages fondées sur des règles).

ii) **Destiné à interagir:** le système d'IA doit être destiné à interagir avec des personnes physiques. L'interaction suppose la capacité à permettre un échange bidirectionnel d'informations ou d'actions entre des personnes physiques et le système d'IA, qui présente un véritable caractère conversationnel ou réactif. Elle peut se présenter sous diverses formes (auditive, visuelle et physique, par exemple). Les personnes physiques devraient être en mesure de fournir des données d'entrée (par exemple, du texte écrit ou d'autres formes de contenu, des indications vocales ou des actions physiques) au système d'IA (par exemple, par des moyens de communication techniques, y compris des interfaces en ligne, un téléphone ou un courrier électronique, ou directement), tandis que le système d'IA devrait être capable de répondre par des sorties contextuelles (y compris tout type de contenu ou d'actions) et inversement. L'interaction peut concerner tout type de sortie ou de format qui peut être perçu et compris par des êtres humains. Elle ne doit pas nécessairement être déclenchée par un être humain; elle peut également englober les contenus ou les actions de l'IA suscitant une réponse humaine. L'interaction peut prendre la forme d'un échange unique (par exemple, une seule invite suivie d'une seule réponse) ou se dérouler sur une certaine période de temps (avec plusieurs échanges successifs). Les systèmes qui collectent des données uniquement de manière passive et qui ne sont pas en mesure de participer à un échange avec des personnes physiques (par exemple, les contrôles d'accès automatisés à reconnaissance faciale) ou qui ne recueillent qu'un retour d'information ponctuel (par exemple, les filtres anti-spam) ne sont pas considérés comme interagissant avec des personnes physiques.

iii) **Directement:** l'interaction entre le système d'IA et les personnes physiques doit être directe. Cela suppose généralement une interaction en temps réel ou quasi réel (y compris grâce aux capacités du système à écrire et à envoyer des messages à des personnes physiques ou à interagir d'une autre manière avec l'environnement physique ou virtuel qui peut être perçu par les personnes physiques). L'interaction directe exclut l'interaction humaine indirecte ou passant par un intermédiaire lorsqu'une personne physique est exposée à des sorties d'IA sans interagir directement avec le système (par exemple, si les représentants du service à la clientèle utilisent des outils d'assistance par IA pour les aider à communiquer avec des personnes physiques ou si les sorties d'IA ne sont pas mises à la disposition de la personne par le système d'IA lui-même, mais par une autre personne qui diffuse le contenu). Toutefois, la simple possibilité pour les êtres humains d'intervenir ou d'examiner les sorties du système d'IA ne devrait pas être utilisée pour contourner l'application de l'obligation d'information prévue à l'article 50, paragraphe 1, du règlement sur l'IA. Les cas de produits et services combinant des réponses générées par une IA avec des contenus sélectionnés par un être humain relèvent du champ d'application de l'article 50, paragraphe 1, du règlement sur l'IA et nécessitent une indication sur le fait que les sorties concernées de l'interaction sont générées par une IA, à moins que ces sorties n'aient été soigneusement examinées et envoyées par des êtres humains en tant que principaux interlocuteurs des personnes physiques.

iv) **Avec des personnes physiques:** l'article 50, paragraphe 1, du règlement sur l'IA impose que le système d'IA interagisse avec des personnes physiques. Ces personnes peuvent être des déployeurs professionnels, d'autres utilisateurs (utilisant également le système pour des activités strictement personnelles) ou d'autres personnes utilisant le système pour leur compte. Les systèmes d'IA qui fonctionnent dans un environnement physique fermé (tel qu'une partie d'une configuration de machines industrielles), les appels de machine à machine effectués en arrière-plan entre des systèmes d'IA (y compris les agents d'IA) dont les sorties ne sont pas destinées à atteindre des personnes physiques, ou les environnements virtuels sans aucun contact direct avec des personnes physiques sont exclus du champ d'application de l'obligation.

([18]) Voir l'article 3, paragraphe 1, du règlement sur l'IA. Pour de plus amples orientations sur la notion de système d'IA, voir les lignes directrices de la Commission sur la définition d'un système d'intelligence artificielle, 29.7.2025, C(2025) 5053 final.

(31) Les agents d'IA sont visés par l'article 50, paragraphe 1, du règlement sur l'IA s'ils sont capables d'interagir avec les personnes qui leur donnent des instructions ou avec d'autres personnes physiques pour l'exécution de tâches (par exemple, effectuer des réservations, gérer la correspondance, négocier ou conclure des contrats, exécuter des achats, etc.). Les agents d'IA doivent être conçus et développés de manière à indiquer à la fois leur nature artificielle et la personne pour le compte de laquelle ils agissent, en tenant compte de la nécessité de transparence quant à l'origine et à la délégation de l'autorité, ainsi qu'à la responsabilité concernant les conséquences de leurs actions. Cela inclut les cas dans lesquels les agents d'IA sont capables d'interagir dans des architectures multiagents complexes avec d'autres agents d'IA qui interagissent directement avec les personnes physiques lors de l'exécution de la tâche. Néanmoins, il n'est pas toujours possible pour le fournisseur de l'agent d'IA de reconnaître à l'avance les cas particuliers d'interaction directe d'un agent d'IA avec d'autres personnes physiques une fois que l'agent d'IA est déployé dans différents contextes. Lorsque le fournisseur ne peut pas déterminer de manière fiable si l'agent d'IA interagira directement avec une personne physique avant de le mettre sur le marché ou de le mettre en service, l'agent devrait être conçu au niveau de l'architecture et avoir pour instruction de se présenter comme un agent d'IA dans chaque situation où il est raisonnablement susceptible d'interagir avec une personne physique (y compris lorsqu'une personne physique peut représenter une entité juridique ou agir au nom de celle-ci). En outre, les agents d'IA devraient également indiquer leur nature artificielle aux personnes qui leur donnent des instructions à des étapes clés (par exemple, au moment d'une autorisation, d'une déclaration, d'une validation, etc., y compris lorsque l'agent d'IA déployé reçoit, traite ou utilise des sorties générées par d'autres systèmes d'IA plutôt que directement par une personne physique) et à chaque nouvelle interaction.

Exemples de systèmes d'IA destinés à interagir directement avec des personnes physiques relevant du champ d'application de l'article 50, paragraphe 1, du règlement sur l'IA:

— assistants vocaux et dialogueurs/agents conversationnels dotés d'IA dans divers contextes (par exemple, service public, assistance à la clientèle, gestion des plaintes, commerce électronique, finance, soins de santé, éducation, etc.);

— assistance téléphonique fondée sur l'IA pour les incidents, les fraudes ou d'autres signalements, robots/cobots (humanoïdes), compagnons IA; animaux de compagnie robotiques; avatars d'IA (par exemple, dans des environnements de réalité virtuelle);

— bots d'IA sur les réseaux et les médias sociaux, agents de codage et autres agents d'IA capables d'interagir directement avec des personnes physiques (y compris dans des architectures multiagents complexes).

3.2.1. **Exemples de systèmes d'IA n'interagissant pas directement avec des personnes physiques et ne relevant pas du champ d'application de l'article 50, paragraphe 1, du règlement sur l'IA:**

— les robots industriels traditionnels fondés sur l'IA qui fonctionnent dans un environnement industriel fermé et qui ne sont pas destinés à interagir avec des êtres humains (par exemple, les robots d'assemblage électronique);

— les systèmes de recommandation, les filtres anti-spam, les outils de traduction ou de transcription automatiques fondés sur l'IA;

— la recherche et l'extraction d'informations fondées sur l'IA sans production ou modification de contenu;

— les techniques de saisie automatique de texte et de code, les systèmes d'authentification/de reconnaissance biométrique ou les systèmes dorsaux d'aide à la décision fondés sur l'IA dans lesquels l'utilisateur ne voit que la sortie d'IA sans avoir la possibilité d'interagir directement avec le système;

— les systèmes de maintenance prédictive et d'optimisation fondés sur l'IA dans les usines et les applications industrielles, la visualisation de produits assistée par l'IA (par exemple, essayage virtuel, aménagement de pièces, rendu de produits), les outils de planification d'itinéraires à requête unique fondés sur l'IA, la transcription d'appels de clients traités à des fins d'analyse comportementale, etc.

3.1.2. *Obligation d'information au titre de l'article 50, paragraphe 1, du règlement sur l'IA*

(32) L'article 50, paragraphe 1, du règlement sur l'IA oblige les fournisseurs de systèmes d'IA qui interagissent directement avec des personnes physiques à veiller à ce que leurs systèmes d'IA (y compris les agents d'IA et les systèmes d'IA à usage général) soient conçus et développés de manière que les personnes physiques concernées soient informées qu'elles interagissent avec un système d'IA.

(33) Alors que le mécanisme d'information doit être intégré dans la conception et le développement du système et dans la manière dont celui-ci interagit avec les personnes physiques, l'information doit être fournie pendant le fonctionnement du système d'IA et au plus tard au moment de la première interaction avec la personne physique (comme l'exige l'article 50, paragraphe 5, du règlement sur l'IA, voir la section 7 ci-dessous). Il incombe aux fournisseurs de veiller à ce que, après la mise sur le marché ou la mise en service du système d'IA, les personnes physiques qui interagissent avec ce dernier soient informées de la nature de l'interaction de manière effective et cohérente tout au long du cycle de vie du système, tel qu'il est mis sur le marché ou mis en service par ce fournisseur.2.2

(34) **Modalités d'information:** l'article 50, paragraphe 1, du règlement sur l'IA ne précise aucune technique particulière de communication de l'information. Les fournisseurs peuvent donc choisir n'importe quelle technique appropriée, pour autant qu'elle respecte les deux critères suivants: i) les exigences énoncées à l'article 50, paragraphe 5, du règlement sur l'IA sont remplies en ce qui concerne la fourniture d'informations claires et reconnaissables au plus tard au moment de la première interaction (voir la section 7 ci-dessous); et ii) les caractéristiques des personnes physiques appartenant à des groupes vulnérables en raison de leur âge (par exemple, les enfants ou les personnes âgées) ou d'un handicap sont prises en compte dans la mesure où le système d'IA est destiné à interagir ou est raisonnablement susceptible d'interagir avec des personnes physiques appartenant à ces groupes (par exemple, y compris lorsque le système d'IA est raisonnablement susceptible d'interagir avec le grand public) ([19]). Les informations doivent également être fournies dans un format accessible aux personnes physiques handicapées, comme précisé à l'article 50, paragraphe 5, et au considérant 132 du règlement sur l'IA (voir la section 7). Lorsque le système peut interagir avec des enfants, ces notifications doivent être adaptées aux enfants, adaptées à leur âge, faciles à comprendre et facilement accessibles à tous les enfants, y compris ceux présentant un handicap ou ayant des besoins supplémentaires en matière d'accessibilité. Il est recommandé aux fournisseurs d'utiliser les pratiques de communication d'information et les mentions normalisées, lorsqu'elles sont disponibles et appropriées, afin de garantir l'efficacité et la cohérence des notifications et une meilleure prise de conscience des utilisateurs.

(35) **Contenu de la notification:** l'article 50, paragraphe 1, du règlement sur l'IA impose aux fournisseurs de systèmes d'IA interactive de les concevoir et de les développer de manière que toutes les personnes physiques qui interagissent avec ces systèmes d'IA soient explicitement informées de la nature artificielle et non humaine de leur interlocuteur. Si le fournisseur peut raisonnablement prévoir que le système d'IA puisse interagir avec des personnes physiques appartenant aux groupes vulnérables susmentionnés, la notification devrait être adaptée en conséquence et fournie dans un format accessible, comme l'exige le droit applicable de l'Union, afin de garantir également la pleine compréhension de ces personnes physiques ([20]).

(36) Les notifications devraient fournir des informations sur l'origine artificielle de l'interaction d'une manière adaptée au contexte dans lequel cette dernière a lieu. Cette communication devrait se faire de manière à garantir que les personnes concernées sont correctement et efficacement informées, en évitant tout risque de tromperie ou de manipulation avant l'interaction et pendant toute sa durée, tout en veillant à ce que les modalités de communication de l'information soient réalisables et proportionnées. Parmi les exemples de fourniture d'informations figurent les notifications par écrit (par exemple, un dialogueur qui entame une conversation en mentionnant qu'elle est fondée sur une technologie de l'IA), les moyens visuels (par exemple, un courrier électronique généré par un agent d'IA envoyé à une personne physique qui porte une étiquette d'IA dans son en-tête), les moyens auditifs (par exemple, un assistant vocal qui déclare au début d'une session qu'il est alimenté par une IA), la communication des identifiants d'IA et de justificatifs (par exemple, des agents d'IA qui révèlent leur identité d'IA de manière vérifiable, dans la mesure du possible) ([21]).

([19]) Voir le considérant 132 du règlement sur l'IA. Pour ce faire, il n'est pas nécessaire que le système soit uniquement destiné à ces groupes vulnérables. Si un système est destiné à interagir avec n'importe quel citoyen, les membres des groupes vulnérables sont également inclus et devraient être pris en considération par les fournisseurs lors de la mise en œuvre des mesures de notification.

([20]) Voir la section 7.3 pour les exigences applicables en matière d'accessibilité.

([21]) Par exemple, les attestations électroniques d'attributs établies conformément au règlement (UE) nº 910/2014 et mises à disposition dans les portefeuilles européens d'identité numérique mis en place en vertu dudit règlement, ainsi que les portefeuilles européens d'identité numérique pour les entreprises tels que proposés, peuvent fournir des moyens sûrs et efficaces d'identifier les agents d'IA. Les portefeuilles peuvent stocker et gérer des attestations électroniques qui vérifient l'identité, les attributs et les autorisations de l'agent d'IA, permettant ainsi de communiquer ces informations aux personnes physiques de manière fluide et fiable.

(37) **Format des notifications:** la conception du système d'IA avec des approches multimodales et accessibles en matière de communication d'informations est recommandée en tant que bonne pratique. La communication de ces notifications peut combiner différentes techniques afin d'améliorer la compréhension des utilisateurs, lorsque cela est approprié et en fonction du contexte, des groupes cibles et prévisibles et des ressources dont disposent les fournisseurs, en particulier les petites et moyennes entreprises (PME). Parmi les techniques appropriées, on peut citer:

— **la communication textuelle** (fondée sur une interface utilisateur): des étiquettes ou des bannières en langage clair et bien visibles (par exemple, «Vous interagissez avec un système d'IA») et des messages d'accueil dès le début de l'échange dans les dialogueurs qui peuvent être combinés, le cas échéant, avec des badges ou des étiquettes permanents visibles tout au long de l'interaction, dans la mesure où cela est proportionné au contexte et aux risques spécifiques. En outre, il est recommandé de placer ces informations à proximité de l'interface d'interaction (par exemple, à proximité du champ d'entrée/de sortie) et d'employer une formulation simplifiée, en particulier pour les utilisateurs ayant un niveau de compétence numérique plus faible et pour les enfants;

— **la communication auditive:** dans les échanges vocaux ou téléphoniques, des déclarations orales explicites au début de l'interaction (par exemple, «Il s'agit d'un assistant alimenté par l'IA») combinées, le cas échéant, à des rappels périodiques lors des interactions plus longues, en particulier en cas d'interruption ou de changement du rôle du système d'IA au cours du parcours de l'utilisateur. Des signaux audio distincts (par exemple, des tonalités ou des carillons) peuvent faciliter la reconnaissance d'un système d'IA, en particulier pour les utilisateurs malvoyants, mais ne sont pas considérés comme suffisants à eux seuls;

— **les signaux visuels/graphiques:** conception du système d'IA intégrant des éléments permanents, tels que des icônes, des filigranes, des cadres colorés ou des symboles «IA» reconnaissables pour compléter les informations textuelles. L'utilisation d'indicateurs visuels normalisés, y compris entre les services d'un fournisseur, est recommandée, étant donné que ces indicateurs peuvent réduire considérablement la charge cognitive et accroître la reconnaissance du système d'IA;

— **les combinaisons multimodales:** combinaison de signaux textuels, audio et visuels (par exemple, un dialogueur affichant une étiquette, tout en fournissant une information textuelle au début de l'interaction) afin de garantir l'accessibilité et de renforcer la clarté parmi les différents groupes d'utilisateurs.

(38) Certaines techniques ne sont pas nécessaires et, lorsqu'elles sont utilisées seules, sont insuffisantes pour remplir efficacement l'obligation de transparence prévue à l'article 50, paragraphes 1 et 5, du règlement sur l'IA, notamment:

— les informations contenues uniquement dans les conditions d'utilisation, les URL ou la documentation (ces informations peuvent compléter, sans les remplacer, les informations communiquées dans le contexte de l'interaction);

— les marquages lisibles par machine (par exemple, métadonnées ou filigranes) qui ne sont pas perceptibles par les utilisateurs au moment de l'interaction. Cela ne compromet pas le caractère approprié de ces marquages pour se conformer à l'article 50, paragraphe 2, du règlement sur l'IA (voir la section 4);

— les signaux peu clairs ou ambigus (par exemple, des références génériques à l'«assistant») ou des représentations anthropomorphes susceptibles d'induire les utilisateurs en erreur;

— les informations générales qui ne sont pas suffisamment spécifiques aux sorties du système d'IA et aux interactions avec celui-ci (par exemple, sur une plateforme fournissant divers services ou traitant des données d'entrée individuelles, une notification générale telle que «Les services sur ce site web utilisent l'IA» est insuffisante);

— les descriptions techniques ou relatives aux capacités: les déclarations faisant uniquement référence aux technologies sous-jacentes (par exemple, «ce système utilise des LLM») sans expliquer la fonction du système, ce que cela implique pour l'utilisateur, ni son origine artificielle.

(39) Les communications de l'information devraient être proportionnées et adaptées en fonction des différents contextes et des publics cibles du système d'IA afin de rester efficaces et conviviales et d'éviter, dans la mesure du possible, d'éventuelles perturbations de l'expérience utilisateur ou des effets d'accoutumance (phénomène dit de «cécité envers les bannières»).

(40) Une seule notification bien visible avant la première interaction du système d'IA avec une personne physique donnée devrait suffire dans la plupart des cas pour satisfaire à l'obligation prévue à l'article 50, paragraphe 1, du règlement sur l'IA. Toutefois, dans certains contextes plus risqués, des rappels périodiques et des communications d'informations tenant compte du contexte seront probablement nécessaires pour garantir que l'utilisateur reste conscient de la nature de l'interaction. C'est notamment le cas dans le contexte des systèmes d'IA qui interagissent avec des personnes vulnérables (par exemple, les enfants, les personnes âgées, les personnes handicapées et les personnes ayant un faible niveau de lettrisme), des interactions prolongées et évolutives dans des contextes sensibles ou immersifs (par exemple, lorsque les utilisateurs peuvent exprimer ou ressentir une détresse ou une vulnérabilité émotionnelles ou être exposés à un risque de comportements de type addictif), lorsqu'il existe un risque accru que les utilisateurs soient induits en erreur (par exemple, conseils financiers, assurances, assistance juridique, conseils en matière de santé et traitement des plaintes), ou qu'ils développent des liens affectifs ou des dépendances (par exemple, envers des compagnons virtuels), ainsi que lorsqu'il est nécessaire de veiller à ce que les utilisateurs soient conscients des actions des agents d'IA et en gardent le contrôle. Les fournisseurs doivent également concevoir et développer le système d'IA afin de garantir la communication d'informations dans toutes les situations où le système est interrogé sur sa nature ou sur l'origine de l'interaction, ou lorsque les échanges avec la personne physique permettent raisonnablement de supposer que celle-ci est susceptible d'être induite en erreur ou laissée perplexe quant à l'origine artificielle de l'interaction.

3.2. **Exceptions à l'obligation d'information prévue à l'article 50, paragraphe 1, du règlement sur l'IA**

(41) L'article 50, paragraphe 1, du règlement sur l'IA prévoit deux exceptions à l'obligation d'information concernant les systèmes d'IA interactive, examinées ci-dessous.

3.2.1. *Exception en cas d'interaction manifeste avec un système d'IA*

(42) Pour pouvoir invoquer la première exception, les fournisseurs devront évaluer et démontrer: i) le caractère artificiel manifeste de l'interaction pour ii) une personne physique qui doit être normalement informée et raisonnablement attentive et avisée, compte tenu des circonstances et du contexte d'utilisation.

(43) Cette norme s'appuie sur le droit établi de l'UE en matière de protection des consommateurs en ce qui concerne la notion de «consommateur moyen» et devrait être interprétée de manière cohérente, tout en tenant compte des contextes et utilisations spécifiques et variés des systèmes d'IA qui interagissent directement avec des personnes physiques et des objectifs poursuivis par l'article 50, paragraphe 1, du règlement sur l'IA.

(44) Pour qu'une personne physique puisse être considérée comme un consommateur «moyen», deux étapes sont nécessaires. En premier lieu, les fournisseurs devraient tenir compte du public auquel le système d'IA est destiné et avec lequel il est raisonnablement susceptible d'interagir. En second lieu, le fournisseur devrait examiner à quel point un membre moyen de ce public est normalement informé et raisonnablement attentif et avisé, en tenant compte des circonstances et du contexte d'utilisation. Plusieurs facteurs doivent être pris en considération lors de cette évaluation:

— les fournisseurs devraient tenir compte du public cible du système d'IA, ainsi que d'un public potentiel plus large qui soit raisonnablement prévisible. Cela sera particulièrement le cas si le grand public peut facilement accéder au système d'IA ou si le système fonctionne dans des espaces publics virtuels ou physiques et est capable d'interagir avec des personnes physiques présentes dans ces espaces. Le public susceptible d'être exposé à un système d'IA ne correspond pas toujours au public cible.

— Les fournisseurs devraient tenir compte de la diversité potentielle de la composition du public cible et raisonnablement prévisible, conformément au considérant 132 du règlement sur l'IA. Lorsque des personnes handicapées, des personnes âgées ou des mineurs sont susceptibles de faire partie de ce public, les niveaux attendus d'information, d'attention et de circonspection d'un membre moyen d'un tel public seront inférieurs à ceux d'un membre moyen d'un public qui n'inclut pas ces catégories de personnes. En revanche, si le système d'IA interactive est uniquement mis à la disposition d'un public cible professionnel ou spécialisé, cela pourrait avoir une incidence positive sur les niveaux attendus de connaissances, d'attention et de circonspection, et un membre moyen de ce public peut être considéré mieux informé et plus attentif et avisé.

— Plus généralement, dans le cadre des circonstances et du contexte d'utilisation, les fournisseurs devraient tenir compte du niveau de compétence numérique et de maîtrise de l'IA du public cible et raisonnablement prévisible, ainsi que des facteurs pertinents sur les plans social, culturel et linguistique.

(45) **Caractère manifeste de la nature artificielle de l'interaction:** lorsqu'ils déterminent, à titre hypothétique, que la personne physique est normalement informée et raisonnablement attentive et avisée, les fournisseurs devraient évaluer s'il ressortira clairement pour cette personne qu'elle interagit directement avec un système d'IA. L'exception devrait être interprétée de manière restrictive étant donné qu'elle prive les personnes physiques de la protection et du droit d'être informées par des notifications claires et reconnaissables qu'elles interagissent avec des systèmes d'IA. Le fait que les consommateurs et les autres personnes physiques soient généralement conscients de l'existence de systèmes d'IA (y compris les dialogueurs et les agents) ne signifie pas qu'ils les reconnaissent lors des interactions. Plusieurs facteurs peuvent être pertinents et pris en considération dans le cadre de l'évaluation visant à déterminer si la nature artificielle de l'interaction est manifeste:

— le type d'interaction prévu: dans les cas où le système d'IA interagit physiquement avec des personnes physiques, les facteurs pertinents peuvent inclure la visibilité des composants mécaniques (ce qui augmenterait le caractère manifeste de la nature artificielle de l'interaction) ou la mesure dans laquelle le système d'IA reproduit son équivalent non artificiel de manière authentique (ce qui réduirait le caractère manifeste de sa nature artificielle). Pour les interactions de type visuel ou auditif, des éléments tels que le style d'écriture ou le mode d'expression orale, le ton de la voix (voix robotisée par opposition à une voix aux sonorités authentiquement humaines), la conception de l'interface utilisateur (par exemple, une image de profil liée à un dialogueur et représentant un être humain) et la capacité à permettre des interactions personnalisées avancées peuvent avoir une incidence sur cette évaluation;

— la composition du public cible et prévisible: si le système d'IA interactive est uniquement mis à la disposition d'un public professionnel ou spécialisé, sans exposition potentielle à des groupes vulnérables, cela pourrait accroître la probabilité que la nature artificielle de l'interaction soit manifeste pour un membre moyen du public concerné. En revanche, si le système d'IA est susceptible d'être accessible au grand public ou à des consommateurs incluant également des personnes vulnérables (par exemple, des enfants, des personnes âgées, des personnes présentant des handicaps cognitifs, physiques ou mentaux, ou des personnes ayant une faible maîtrise de l'IA) pour lesquelles la nature de l'interaction est moins manifeste, l'exception ne peut être invoquée afin de protéger efficacement ces utilisateurs vulnérables;

— le fait qu'il devient de plus en plus difficile pour les personnes physiques de savoir si elles interagissent avec un système d'IA ou un véritable être humain: compte tenu de l'utilisation d'avatars et d'agents, de l'anonymat général des interactions en ligne et des capacités des systèmes d'IA à écrire et à interagir de manière humaine, l'exception fondée sur le caractère «manifeste» devrait être limitée aux cas où il ne subsiste pratiquement aucun doute quant à la nature de l'interaction pour une personne moyenne du public ciblé et raisonnablement prévisible.

Exemples de cas dans lesquels la nature artificielle de l'interaction est manifeste et pour lesquels l'obligation de transparence prévue à l'article 50, paragraphe 1, du règlement sur l'IA ne s'applique pas:

— les dialogueurs d'assistance au codage et de révision de code alimentés par une IA uniquement accessibles aux développeurs professionnels qui, en raison de leur expertise et du contexte d'utilisation, ne peuvent raisonnablement pas s'attendre à interagir avec un être humain et peuvent facilement comprendre que les suggestions sont générées par un système d'IA;

— un assistant interne à destination des employés réservé à du personnel dûment formé qui maîtrise l'IA et sait qu'il utilise des systèmes d'IA à des fins organisationnelles internes (par exemple, RH, juridique, passation de marchés, conformité ou soutien informatique);

— les systèmes d'IA interactive destinés uniquement à être utilisés par des professionnels de la santé dûment formés pour soutenir le diagnostic médical et suggérer les traitements correspondants;

— les systèmes d'IA ambiante intégrés dans des appareils ménagers et dotés de fonctionnalités limitées à l'assistance et au fonctionnement des appareils ménagers;

— les interactions avec des personnages non joueurs (PNJ) fondés sur l'IA dans un jeu vidéo à joueur unique lorsqu'il ressort clairement du type de jeu vidéo qu'aucune autre personne physique ne peut participer et interagir avec l'utilisateur et que la nature artificielle de l'interaction est également manifeste pour tous les utilisateurs (y compris les personnes vulnérables).

Exemples de cas qui ne permettent pas d'invoquer l'exception fondée sur le caractère manifeste de la nature artificielle de l'interaction et pour lesquels l'obligation de transparence prévue à l'article 50, paragraphe 1, du règlement sur l'IA s'applique:

— un animal de compagnie robotique alimenté par l'IA, d'apparence très semblable à son équivalent naturel et conçu pour reproduire les interactions typiques entre l'être humain et l'animal de compagnie, créant une ambiguïté quant à la nature artificielle ou non de l'interaction;

— les systèmes d'IA intégrés dans des environnements immersifs (par exemple, réalité virtuelle ou augmentée) utilisant des avatars ou des voix réalistes ressemblant à des êtres humains, où les utilisateurs, et en particulier les enfants, les personnes âgées ou les personnes handicapées, peuvent ne pas être en mesure de faire facilement la distinction entre une interaction humaine et une interaction avec une IA;

— les dialogueurs dotés d'IA intégrés dans des plateformes en ligne ou des outils d'assistance (services d'assistance technique) permettant aux utilisateurs d'interagir directement et de recevoir des sorties d'IA (par exemple, des réponses à des requêtes ou d'autres contenus générés par l'IA) qu'ils peuvent percevoir comme ayant été générées par un être humain.

3.2.2. *Exception pour les systèmes d'IA dont l'utilisation est autorisée par la loi à des fins répressives*

(46) Les fournisseurs de systèmes d'IA interactive sont exemptés de l'obligation de transparence prévue à l'article 50, paragraphe 1, du règlement sur l'IA si la loi autorise l'utilisation de ces systèmes à des fins de prévention ou de détection des infractions pénales, d'enquêtes ou de poursuites en la matière, sous réserve de garanties appropriées pour les droits et libertés des tiers.

(47) Par «la loi», on entend le droit de l'Union, ainsi que le droit national des États membres adopté conformément au droit de l'Union. La loi autorisant l'utilisation du système d'IA doit préciser clairement les finalités et les circonstances de l'utilisation autorisée et prévoir des garanties appropriées pour protéger les droits et libertés des tiers. Compte tenu de l'approche neutre sur le plan technologique du règlement sur l'IA, il n'est pas nécessaire d'identifier ou d'autoriser par voie législative des systèmes d'IA spécifiques lorsque l'utilisation du système d'IA est généralement prévue dans le cadre de l'activité légalement autorisée et est conforme au cadre juridique applicable.

(48) Pour relever de cette exception, le système d'IA doit avoir pour finalité de prévenir ou de détecter des infractions pénales, d'enquêter sur celles-ci ou d'engager des poursuites en la matière (par exemple, agent d'IA infiltré, systèmes d'IA utilisés sous le contrôle humain d'enquêteurs pour des enquêtes en ligne ou des communications avec des personnes susceptibles d'être impliquées dans des activités criminelles). L'exception ne se limite pas à l'utilisation de ces systèmes d'IA uniquement par les autorités répressives au sens de l'article 3, point 48, du règlement sur l'IA, mais peut également s'appliquer aux systèmes d'IA interactive (ou générative) utilisés par d'autres autorités publiques nationales ou de l'Union, pour autant que leur utilisation soit autorisée par la loi à des fins de prévention ou de détection des infractions pénales, d'enquêtes ou de poursuites en la matière, et sous réserve de garanties appropriées pour protéger les droits et libertés des tiers. L'exemption n'exige pas que le système d'IA soit en mesure de mener les enquêtes de manière indépendante. Si un système donné peut être utilisé à la fois à des fins répressives et à d'autres fins ordinaires, l'applicabilité de l'exemption devrait être évaluée au regard du déploiement considéré. Les obligations de transparence continuent de s'appliquer lorsque le même système d'IA est déployé à des fins autres que celles pour lesquelles la loi en autorise l'utilisation, à savoir la prévention ou la détection des infractions pénales, les enquêtes ou les poursuites en la matière. La transparence doit toujours être assurée aux personnes physiques lorsqu'elles interagissent avec le système dans le cadre de ses finalités habituelles.

(49) L'exception ne s'applique pas si i) le système d'IA est mis à la disposition du public et ii) le système offre une fonctionnalité permettant aux personnes de signaler des infractions pénales. Ces systèmes peuvent être utilisés, mais ils restent soumis à l'obligation de transparence prévue à l'article 50, paragraphe 1, du règlement sur l'IA, de sorte que les personnes physiques doivent être informées qu'elles interagissent avec ces systèmes.

Pour plus de clarté, cette exception concerne uniquement l'obligation prévue à l'article 50, paragraphe 1, d'informer les personnes physiques qu'elles interagissent avec un système d'IA ([22]). Elle est sans préjudice du droit de l'Union et du droit national applicables, y compris les règles de procédure concernant la documentation, la divulgation, la recevabilité et l'évaluation des éléments de preuve.

([22]) Cela s'applique également aux exceptions similaires au bénéfice des activités répressives prévues à l'article 50, paragraphes 2 à 4, du règlement sur l'IA.

Exemples de systèmes d'IA mis à la disposition du public pour permettre le signalement d'infractions pénales qui ne sont pas exemptés et relèvent du champ d'application de l'article 50, paragraphe 1, du règlement sur l'IA:

— les dialogueurs de police fondés sur l'IA déployés sur des sites web ou des applications mobiles officiels de la police qui permettent aux personnes de signaler des infractions pénales ou de déposer plainte. Ces systèmes guident les utilisateurs au moyen de questions structurées, collectent les informations pertinentes et trient les signalements avant de les transmettre aux agents humains pour évaluation;

— les lignes téléphoniques d'assistance alimentées par l'IA gérées par la police ou une autre autorité répressive pour les enquêtes publiques générales;

— les lignes téléphoniques d'assistance ou portails numériques de signalement des fraudes assistés par l'IA exploités par des établissements financiers ou des autorités publiques, permettant aux utilisateurs de signaler des infractions financières présumées. Le système d'IA interagit avec l'utilisateur pour recueillir des renseignements, catégoriser le signalement et classer les cas par ordre de priorité, tandis que des enquêteurs humains valident les informations avant toute action;

— les assistants virtuels utilisés pour le recueil de témoignages, permettant aux victimes ou aux témoins de fournir des informations au moyen d'une interface fondée sur l'IA. Ces systèmes d'IA peuvent prendre en charge des saisies multilingues, offrir des fonctionnalités d'accessibilité et permettre la présentation structurée d'éléments de preuve, mais l'évaluation finale et les étapes de l'enquête restent du ressort des autorités humaines.

3.3. **Interaction avec d'autres actes juridiques de l'Union**

(50) L'obligation de transparence prévue à l'article 50, paragraphe 1, du règlement sur l'IA s'applique sans préjudice de la législation existante de l'Union en matière de protection des consommateurs et de protection des données, en matière numérique ou en matière de non-discrimination qui peut exiger d'autres communications d'informations non spécifiques à l'origine artificielle de l'interaction. En particulier, en vertu de la directive 2005/29/CE (directive sur les pratiques commerciales déloyales) ([23]), les actions ou omissions trompeuses en ce qui concerne les caractéristiques principales d'un produit ou d'un service sont interdites. En outre, la directive 2011/83/CE (la «directive relative aux droits des consommateurs») ([24]) impose aux professionnels de fournir des informations précontractuelles claires et compréhensibles sur les principales caractéristiques des biens ou des services. Lorsqu'un service est fondé sur l'IA (par exemple, un dialogueur disponible sur abonnement, une application de compagnon virtuel ou un service de coaching virtuel), la fonctionnalité d'IA peut être considérée comme une caractéristique essentielle qui doit être déclarée avant que le consommateur ne soit lié par un contrat. Les obligations d'information prévues par le droit de l'Union en matière de protection des consommateurs s'appliquent indépendamment du fait que la nature artificielle de l'interaction soit considérée comme «ressortant clairement» au sens de l'article 50, paragraphe 1, du règlement sur l'IA. Ces obligations sont également complémentaires en tant qu'obligations générales d'information en matière de protection des consommateurs, tandis que l'obligation de notification prévue par le règlement sur l'IA est plus contextuelle et vise à garantir une prise de conscience continue par toutes les personnes physiques du fait qu'elles ont une interaction directe avec un système d'IA au moins avant le début de l'interaction et tout au long de celle-ci.

(51) Le DSA prévoit également des exigences de transparence pour les systèmes de recommandation des plateformes en ligne, y compris l'obligation d'expliquer les principaux paramètres de classement. Ces obligations sont complémentaires des obligations de transparence énoncées à l'article 50, paragraphe 1, du règlement sur l'IA, étant donné qu'elles ont une portée et une finalité différentes et qu'elles concernent des systèmes qui, en principe, n'interagissent pas directement avec les utilisateurs au sens dudit article.

(52) En ce qui concerne les obligations d'information à l'égard des personnes concernées en vertu du droit de l'Union en matière de protection des données, l'obligation prévue à l'article 50, paragraphe 1, du règlement sur l'IA répond à un objectif différent de ces obligations et, en tant que telle, n'a pas d'incidence sur ces obligations d'information.

([23]) Directive 2005/29/CE du Parlement européen et du Conseil du 11 mai 2005 relative aux pratiques commerciales déloyales des entreprises vis-à-vis des consommateurs dans le marché intérieur (JO L 149 du 11.6.2005, p. 2).

([24]) Directive 2011/83/UE du Parlement européen et du Conseil du 25 octobre 2011 relative aux droits des consommateurs (JO L 304 du 22.11.2011, p. 64).

(53) De même, l'obligation de transparence prévue à l'article 50, paragraphe 1, du règlement sur l'IA s'applique sans préjudice de la législation nationale et du droit de l'Union concernant la fourniture d'informations aux personnes dans le cadre de procédures pénales [par exemple, la directive 2012/13/UE ou la directive 2012/29/UE, telle que modifiée par la directive (UE) 2026/1472] ([25]).

4. **ARTICLE 50, PARAGRAPHE 2, DU RÈGLEMENT SUR L'IA: MARQUAGE ET DÉTECTION DES CONTENUS GÉNÉRÉS OU MANIPULÉS PAR UNE IA**

(54) L'article 50, paragraphe 2, du règlement sur l'IA impose aux fournisseurs de systèmes d'IA qui génèrent des contenus de synthèse de mettre en œuvre des solutions techniques qui satisfont à certaines exigences de qualité pour le marquage lisible par machine et la détection des sorties de leurs systèmes d'IA. L'objectif est de permettre aux personnes physiques de distinguer les contenus générés ou manipulés par une IA des autres contenus (par exemple, les contenus créés par des êtres humains) et de vérifier leur origine, contribuant ainsi également à renforcer l'intégrité de l'écosystème informationnel et la confiance en celui-ci.

(55) Par conséquent, les bénéficiaires de cette obligation de transparence sont toutes les personnes physiques susceptibles d'être exposées à des contenus générés ou manipulés par une IA, ainsi que les acteurs clés qui jouent un rôle important dans l'intégrité de l'écosystème informationnel et la confiance en celui-ci, tels que les autorités de surveillance du marché compétentes et les autres autorités compétentes, les chercheurs indépendants, les organisations de la société civile, les médias, les signaleurs de confiance, les défenseurs des droits fondamentaux, les fournisseurs de plateformes en ligne et de moteurs de recherche en ligne, les titulaires de droits et les autres créateurs et distributeurs de contenus, etc.

4.1. **Principaux éléments et concepts de l'article 50, paragraphe 2, du règlement sur l'IA**

(56) L'article 50, paragraphe 2, du règlement sur l'IA s'applique si plusieurs conditions cumulatives sont remplies:

i) le système doit être considéré comme un système d'IA;

ii) le système d'IA doit être capable de générer ou de manipuler des contenus de synthèse;

iii) le contenu doit se présenter sous l'un ou plusieurs des formats suivants: audio, image, vidéo ou texte;

iv) le système d'IA ne doit relever d'aucune des exceptions suivantes: 1) le système d'IA remplit une fonction d'assistance pour la mise en forme standard; 2) le système d'IA ne modifie pas de manière substantielle les données d'entrée fournies par le déployeur ou leur sémantique; ou 3) l'utilisation du système d'IA est autorisée par la loi à des fins de prévention ou de détection des infractions pénales, d'enquêtes ou de poursuites en la matière.

4.1.1. *Systèmes d'IA générant ou manipulant des contenus de synthèse*

(57) L'article 50, paragraphe 2, du règlement sur l'IA s'applique aux systèmes d'IA ([26]), y compris les systèmes d'IA à usage général, qui génèrent des contenus de synthèse de type audio, image, vidéo ou texte. L'article 50, paragraphe 2, du règlement sur l'IA est applicable à un sous-ensemble de systèmes qui répondent à la définition du système d'IA figurant à l'article 3, point 1), du règlement sur l'IA et qui génèrent ce type de sorties. L'article 50, paragraphe 2, du règlement sur l'IA concerne également les systèmes d'IA utilisés pour manipuler des contenus.

(58) La production de contenus désigne la production d'éléments de synthèse par le système d'IA (par exemple, sur la base d'une invite humaine), telle que la génération d'une image de synthèse ou d'une chanson. Dans la pratique, cela signifie que l'article 50, paragraphe 2, du règlement sur l'IA peut s'appliquer aux systèmes d'IA générative conçus avec une finalité d'utilisation limitée pour produire des sorties spécifiques. Parallèlement, l'article 50, paragraphe 2, du règlement sur l'IA s'applique également aux systèmes d'IA qui peuvent répondre à plusieurs finalités, y compris les systèmes d'IA à usage général, et qui peuvent produire divers types de contenus, ainsi qu'aux systèmes d'IA agentique, pour autant qu'ils génèrent des contenus de synthèse de type audio, image, vidéo ou texte.

([25]) Directive 2012/13/UE du Parlement européen et du Conseil du 22 mai 2012 relative au droit à l'information dans le cadre des procédures pénales (JO L 142 du 1.6.2012, p. 1). Directive 2012/29/UE du Parlement européen et du Conseil du 25 octobre 2012 établissant des normes minimales concernant les droits, le soutien et la protection des victimes de la criminalité et remplaçant la décision-cadre 2001/220/JAI du Conseil (JO L 315 du 14.11.2012, p. 57).

([26]) Tels que définis à l'article 3, point 1), du règlement sur l'IA.

(59) En revanche, la manipulation de contenus désigne la modification de contenus (de synthèse ou non) déjà existants par un système d'IA (par exemple, une image ou un enregistrement vocal existants qui sont manipulés par un système d'IA conformément à des instructions humaines). Pour que ces contenus relèvent de l'article 50, paragraphe 2, du règlement sur l'IA, il suffit qu'ils soient générés ou manipulés par une IA d'une manière allant au-delà de la mise en forme standard au titre des exceptions énoncées à l'article 50, paragraphe 2, du règlement sur l'IA (voir la section 4.3). En outre, l'article 50, paragraphe 2, du règlement sur l'IA n'exige pas que les contenus soient uniquement générés ou manipulés par une IA pour qu'ils relèvent de l'obligation. Les contenus combinés à des éléments créés par un être humain sont également considérés comme des contenus de synthèse s'ils sont manipulés ou générés dans l'un des formats visés à l'article 50, paragraphe 2, du règlement sur l'IA.

4.1.2. *Formats des contenus de synthèse relevant du champ d'application*

(60) L'article 50, paragraphe 2, du règlement sur l'IA fournit une liste exhaustive des formats des contenus de synthèse visés par l'obligation de transparence: audio, image, vidéo ou texte. Cela inclut également les contenus multimodaux, c'est-à-dire les contenus composés d'une combinaison de ces formats. Aux fins de l'article 50, paragraphe 2, du règlement sur l'IA, les différents formats doivent être interprétés en fonction des avancées pratiques et technologiques susceptibles d'évoluer au fil du temps. Pour le moment, on entend par:

— texte, un contenu symbolique autonome composé de caractères, ou de chiffres présentés dans un ordre particulier, qui peuvent être lus et interprétés sémantiquement par les êtres humains;

— image, une représentation spatiale statique contenant des informations visuelles qui peuvent être vues par les êtres humains;

— audio, un signal variant dans le temps représentant un son pouvant être perçu par l'ouïe humaine. Il peut s'agir de paroles, de musique instrumentale ou d'autres signaux audio;

— vidéo, une suite d'images dans le temps, qui peut être synchronisée avec de l'audio et être vue par les êtres humains. Dans ce cas, tant le contenu vidéo que le contenu audio doivent être marqués et identifiables, comme indiqué ci-dessous.

(61) Les systèmes d'IA générant ou manipulant des images 3D, des vidéos 3D et des contenus audio 3D, ainsi que la réalité virtuelle, augmentée ou mixte, relèvent du champ d'application de l'article 50, paragraphe 2, du règlement sur l'IA. La «réalité virtuelle» englobe les technologies d'IA qui permettent aux utilisateurs de faire l'expérience d'un environnement généré par ordinateur simulant une présence physique et d'interagir avec celui-ci et qui permettent une interaction en temps réel entre l'utilisateur et cet environnement. Étant donné que la réalité virtuelle et la réalité augmentée sont constituées d'une suite d'images dans le temps, éventuellement synchronisée avec du contenu audio, qui montre une évolution visuelle au fil du temps, elles peuvent être considérées comme des contenus «vidéo». La «sortie» pertinente à marquer est la forme finale de chaque élément généré par l'IA, avec des exceptions spéciales pour les sessions interactives éphémères en temps réel visées au point 88 ci-dessous. Les artefacts numériques utilisés pour générer des images 3D, des vidéos 3D et du contenu audio 3D, tels qu'un nuage de points ou un maillage 3D, ne doivent pas être marqués.

(62) Les jumeaux numériques qui fournissent une réplique virtuelle de personnes, d'objets physiques ou de systèmes, intégrant plusieurs types de données, relèvent également du champ d'application de l'article 50, paragraphe 2, du règlement sur l'IA s'ils sont générés ou manipulés par une IA et prennent la forme d'un contenu audio, d'une image, d'une vidéo ou d'un texte, à moins qu'ils ne relèvent des exceptions et ne soient considérés comme des «applications industrielles» ou des «applications interentreprises» au sens du point 87 ci-dessous ou comme des contenus éphémères en temps réel au sens du point 88 ci-dessous.

(63) L'article 50, paragraphe 2, du règlement sur l'IA peut également s'appliquer aux agents d'IA si ces derniers effectuent une action dont la sortie est un contenu généré ou manipulé par une IA perceptible par des personnes physiques sous la forme d'un contenu audio, d'une image, d'une vidéo ou d'un texte. Ces sorties doivent donc être marquées et identifiables comme décrit ci-dessous. En revanche, toute action d'un agent d'IA qui n'inclut pas de contenu généré ou manipulé par une IA (par exemple, des étapes de traitement intermédiaires telles que le raisonnement et la chaîne de pensée), ou toute action effectuée par une IA qui n'est pas destinée à être directement perçue par des personnes physiques (par exemple, une demande sur le web ou une action sur le navigateur), ne constitue pas un contenu de synthèse relevant du champ d'application de l'article 50, paragraphe 2, du règlement sur l'IA et ne doit pas respecter les obligations de marquage et de détection.

4.1.3. *Contenus ne relevant pas du champ d'application de l'article 50, paragraphe 2, du règlement sur l'IA*

(64) Les contenus générés par un système d'IA fondé sur un simple traitement de données, qui n'est pas spécifiquement généré ou manipulé par une IA, par exemple une image issue d'un rendu, ne relèvent pas du champ d'application de l'article 50, paragraphe 2, du règlement sur l'IA.

(65) En outre, les sorties d'un système d'IA qui ne font que reproduire des contenus existants ou permettent la présentation ou l'agencement de contenus déjà existants ne sont pas couvertes par l'article 50, paragraphe 2, du règlement sur l'IA. Cela peut être le cas, par exemple, des listes de lecture musicales ou d'autres systèmes de recommandation qui sélectionnent ou classent des contenus existants uniquement en fonction des préférences ou de l'activité de l'utilisateur, sans rien créer de nouveau ni manipuler le contenu lui-même. Il en va de même pour les systèmes d'IA utilisés dans des processus analytiques internes qui extraient et structurent des données, mais ne les résument pas.

(66) Le même principe s'applique également aux entrées et aux sorties d'un système d'IA qui constituent de simples observations et enregistrements de données provenant d'environnements physiques ou virtuels, qui ne sont pas générées ou manipulées par une IA (par exemple, par des robots ou d'autres capteurs dotés d'IA) ou aux transmissions de données par des systèmes d'IA sans aucune modification. Cela concerne toute une série d'entrées et de sorties des systèmes d'IA, en particulier dans des contextes industriels tels que les données de fabrication destinées aux robots observant ces données; la consommation enregistrée par des compteurs intelligents dotés d'IA, les mesures de la fréquence et de la tension du réseau; et l'enregistrement des données de localisation GPS de véhicules, etc.

(67) Par ailleurs, étant donné que l'objectif de l'obligation de transparence prévue à l'article 50, paragraphe 2, du règlement sur l'IA est de permettre aux êtres humains de reconnaître les contenus générés ou manipulés par une IA, de manière à faire face aux risques de tromperie et de manipulation et à garantir l'intégrité de l'écosystème informationnel et la confiance en celui-ci, tout contenu qui n'est pas lié à cet objectif, qui n'est pas perceptible ou qui n'est pas destiné à être interprété, vérifié ou traité par des personnes physiques, n'est pas visé par cette disposition.

(68) En particulier, les contenus suivants devraient être considérés comme ne relevant pas du champ d'application de l'article 50, paragraphe 2, du règlement sur l'IA:

— les sorties générées sous la forme d'une courte suite de chiffres, de symboles ou de lettres (par exemple, mots uniques, légendes d'images, texte de remplacement, étiquettes d'interface utilisateur et graphiques à l'échelle d'une icône et autres étiquettes de données);

— le code source au sens d'un contenu écrit dans un langage de programmation, de script, de balisage, d'interrogation ou de configuration qui est destiné à être interprété, compilé ou exécuté par un système informatique, qu'il soit composé ou non de caractères de texte. Cela peut également inclure les commentaires en langage naturel et les informations contextuelles qui font partie intégrante du code source. Sont également exclus les kits de développement logiciel, le langage d'interrogation structuré (SQL), l'infrastructure en tant que code et «YAML Ain' t Markup Language» (YAML), la configuration en notation des objets JavaScript (JSON), les schémas, les scripts, les spécifications lisibles par machine, les interfaces de programmation d'application (API et bibliothèques de logiciels);

— les sorties d'un système d'IA destinées à être exclusivement communiquées de machine à machine et traitées automatiquement sans être exposées à des êtres humains, par exemple la communication d'agent à agent, les signaux anti-spam, pour autant que ces sorties ne soient pas perçues par des personnes physiques;

— les sorties uniquement utilisées dans des environnements en boucle fermé dans le cadre de flux de travail industriels et de développement de produits (par exemple pour la production de films, d'animations, de jeux ou de publicités), à moins qu'il ne s'agisse de la sortie finale du système d'IA sous la forme de contenus générés ou manipulés par une IA de type texte, audio, image ou vidéo. En pareils cas, seule la sortie finale des étapes du flux de travail doit être marquée et identifiable.

4.2. **L'obligation de marquage et de détection prévue à l'article 50, paragraphe 2, du règlement sur l'IA**

(69) L'article 50, paragraphe 2, du règlement sur l'IA établit une obligation de transparence pour les fournisseurs de systèmes d'IA générative relevant du champ d'application de cette disposition, comprenant deux critères distincts, mais intrinsèquement liés. Premièrement, les fournisseurs doivent veiller à ce que les sorties du système d'IA soient marquées dans un format lisible par machine. Deuxièmement, les fournisseurs doivent veiller à ce que les sorties soient identifiables comme du contenu généré ou manipulé par une IA. Pour chaque solution de marquage déployée, les fournisseurs devraient veiller à ce que des moyens de détection correspondants soient disponibles (comme indiqué ci-dessous) pour permettre aux personnes physiques exposées au contenu et aux autres acteurs concernés de reconnaître le contenu généré par le système d'IA et de le distinguer des autres contenus. Une solution technique s'entend comme une combinaison de techniques de marquage et de moyens de détection que le fournisseur a mis en œuvre pour satisfaire à l'obligation de transparence prévue à l'article 50, paragraphe 2, du règlement sur l'IA.

(70) Chacun des deux critères susmentionnés doit être rempli pour atteindre efficacement les objectifs des obligations de transparence prévues à l'article 50, paragraphe 2, du règlement sur l'IA. Le respect d'un seul critère (par exemple, le marquage lisible par machine des sorties sans que les moyens pour leur détection ne soient disponibles) ne suffira pas à garantir la conformité avec cette disposition ([27]).

4.2.1. *L'obligation de marquage*

(71) Le champ d'application de l'obligation de marquage est limité à la mise en œuvre des marquages dans un format lisible par machine. Un format lisible par machine signifie que les marquages sont structurés de manière à permettre aux applications logicielles de les identifier, de les reconnaître et de les extraire facilement sans intervention humaine. Les marquages et les étiquettes perceptibles ne sont pas exclus en tant que mesure complémentaire, le cas échéant, en vue de faciliter le respect par les déployeurs de leur obligation d'étiqueter les hypertrucages conformément à l'article 50, paragraphe 4, du règlement sur l'IA.

(72) Pour se conformer à l'obligation de marquage, les fournisseurs peuvent recourir à une seule technique de marquage ou à une combinaison de techniques, pour autant que leur solution technique globale soit lisible par machine et réponde aux exigences en matière d'efficacité, d'interopérabilité, de solidité et de fiabilité au niveau requis par la loi, comme le prévoit la deuxième phrase de l'article 50, paragraphe 2, du règlement sur l'IA (voir la section 4.2.3 ci-dessous).

(73) Le considérant 133 du règlement sur l'IA fournit des exemples de ces techniques pour inclure les filigranes, les identifications de métadonnées, les méthodes cryptographiques permettant de prouver la provenance et l'authenticité du contenu, les méthodes d'enregistrement, les empreintes digitales ou d'autres techniques, ainsi qu'une combinaison de ces techniques. Si les méthodes permettant de prouver la provenance et l'authenticité sont mentionnées au considérant 133 du règlement sur l'IA, les fournisseurs ne sont pas tenus d'enregistrer ou de conserver une chaîne de provenance complète contenant des informations sur l'origine et les modifications du contenu ou toute autre indication pertinente concernant l'historique du contenu. Toutefois, ces méthodes de provenance peuvent également être utilisées aux fins du respect de l'article 50, paragraphe 2, du règlement sur l'IA et contribuer à permettre aux personnes physiques de distinguer les contenus générés ou manipulés par une IA des contenus authentiques.

(74) Les fournisseurs peuvent mettre en œuvre la solution de marquage pour les sorties à différents stades de la chaîne de valeur (par exemple, après la production de contenus au niveau du système d'IA, au niveau du modèle d'IA sous-jacent ou intégrée au processus d'inférence du système d'IA). Ils peuvent s'appuyer sur la solution de marquage mise en œuvre par un fournisseur de modèle en amont ou un tiers fournissant la solution (par exemple, sous la forme d'une norme ouverte ou d'un service spécialisé) dans la mesure où la solution de marquage est conforme à l'article 50, paragraphe 2, du règlement sur l'IA. Le recours à cette solution est sans préjudice de la responsabilité du fournisseur du système d'IA de démontrer la conformité de ce dernier avec l'article 50, paragraphe 2, du règlement sur l'IA.

([27]) Le considérant 133 du règlement sur l'IA précise qu'il convient d'exiger que les fournisseurs de ces systèmes intègrent des solutions techniques permettant le marquage dans un format lisible par machine et la détection du fait que les sorties ont été générées ou manipulées par un système d'IA, et non par un être humain. Le considérant 135 du règlement sur l'IA fait référence à l'obligation en matière de détection.

4.2.2. *L'obligation de détection*

(75) Les fournisseurs de systèmes d'IA relevant du champ d'application de l'article 50, paragraphe 2, du règlement sur l'IA doivent veiller à ce que les sorties de leurs systèmes soient identifiables comme du contenu généré ou manipulé par une IA. Cela signifie que le fournisseur est tenu de veiller à ce que les moyens de détection soient mis à la disposition des personnes potentiellement exposées au contenu. En outre, conformément à l'article 50, paragraphe 5, du règlement sur l'IA, ces solutions de détection devraient être en mesure de produire des résultats lisibles par un être humain indiquant si le contenu a été généré ou manipulé par une IA.

(76) Un outil de détection est un mécanisme qui détecte si du contenu a été généré ou manipulé par une IA, en repérant généralement des marqueurs techniques ou des signatures qui confirment son origine. Afin de garantir une interopérabilité totale, les fournisseurs doivent s'appuyer sur des solutions de détection conformes aux normes du secteur et mises à la disposition du public, qui permettent à tout tiers de mettre en œuvre une détection idéalement exécutable localement sur l'appareil numérique et qui permettent une intégration harmonieuse dans des réseaux autonomes et isolés. Lorsque de telles normes ne sont pas disponibles, en particulier au stade initial de la mise en œuvre de l'article 50, paragraphe 2, du règlement sur l'IA pour les technologies de filigrane, le fournisseur peut s'appuyer sur sa propre solution de détection ou sur une solution de détection tierce ou partagée, pour autant que ces solutions garantissent l'interopérabilité avec les solutions de détection utilisées par d'autres fournisseurs (voir la section 4.2.3). Cette possibilité devrait être limitée dans le temps jusqu'à la création de normes harmonisées et d'une solution de détection interopérable normalisée, indépendante des fournisseurs, qui soit sécurisée, respectueuse de la vie privée et exécutable localement.

(77) L'article 50, paragraphe 5, du règlement sur l'IA précise que les informations visées à l'article 50, paragraphe 2, dudit règlement devraient être fournies aux personnes physiques exposées au contenu de manière claire et reconnaissable au plus tard au moment de la première interaction ou de la première exposition. Ces informations portent notamment sur le résultat de la détection, qui indique si le contenu est généré ou manipulé par une IA, et non sur le contenu généré ou manipulé par l'IA lui-même. Par «au moment de la première exposition», il convient d'entendre le moment où des personnes physiques sont disposées à vérifier l'origine d'un contenu particulier et à accéder à la solution de détection pour en obtenir les résultats.

(78) Si le fournisseur du système d'IA dépend d'un autre acteur pour la fourniture d'une solution de détection ou l'accès à celle-ci, il doit veiller à ce que cette solution de détection soit conforme à l'article 50, paragraphes 2 et 5, du règlement sur l'IA, et à ce que ces acteurs affichent les résultats de la détection conformément à l'article 50, paragraphe 5, dudit règlement (voir la section 7).

4.2.3. *Respect des exigences relatives aux solutions techniques: efficacité, interopérabilité, solidité et fiabilité*

(79) L'article 50, paragraphe 2, deuxième phrase, du règlement sur l'IA dispose que les solutions techniques de marquage et de détection doivent être efficaces, interopérables, solides et fiables. Ces exigences doivent s'entendre comme suit:

— l'**efficacité** désigne la capacité de la solution technique mise en œuvre par les fournisseurs à détecter leurs marquages et à permettre aux personnes physiques de distinguer les contenus générés ou manipulés par une IA produits par leur système d'IA, contribuant ainsi à l'intégrité de l'écosystème informationnel et à la confiance en celui-ci;

— la **fiabilité** désigne la capacité de la solution technique à repérer avec précision les contenus générés ou manipulés par une IA et à les distinguer des autres contenus dans des conditions nominales pour l'ensemble des contenus générés ou manipulés par le système d'IA du fournisseur;

— la **solidité** désigne la capacité de la solution technique à identifier avec précision les contenus générés ou manipulés par une IA produits par le système d'IA du fournisseur et à les distinguer d'autres contenus dans des conditions variables, englobant à la fois les modifications courantes et les attaques par exemples contradictoires;

— l'**interopérabilité** désigne la capacité des différentes solutions techniques de marquage et de détection à fonctionner de manière fluide entre plusieurs systèmes, acteurs, contextes et mises en œuvre techniques afin de permettre la détection de contenus générés ou manipulés par une IA, quelle que soit la ou les techniques de marquage déployées par les fournisseurs de différents systèmes d'IA.

(80) Les fournisseurs doivent veiller à ce que la combinaison des solutions techniques de marquage et de détection réponde de manière globale à toutes les exigences susmentionnées au niveau requis par la loi, en tenant compte des limites et des complémentarités des différentes solutions déployées. Les solutions techniques doivent satisfaire à toutes les exigences pour autant que la technologie le permet, compte tenu des spécificités et des limites des différents types de contenus, des coûts de mise en œuvre et de l'état de la technique généralement reconnu, comme cela peut ressortir des normes techniques pertinentes.

(81) Les solutions «aussi efficaces, interopérables, solides et fiables que la technologie le permet» au sens de l'article 50, paragraphe 2, du règlement sur l'IA sont des solutions qui peuvent être mises en œuvre pour les formats relevant du champ d'application de la disposition, en utilisant les technologies, méthodes et pratiques d'ingénierie actuellement disponibles, au sein de l'architecture technique et de l'environnement opérationnel spécifiques concernés. Le fournisseur n'est pas tenu d'utiliser une solution technique qui n'est pas encore développée ou disponible sur le marché, ou dont la mise en œuvre est techniquement irréalisable. La faisabilité technique est une notion objective qui ne dépend pas des ressources et des capacités spécifiques des différents fournisseurs.

(82) L'article 50, paragraphe 2, du règlement sur l'IA impose aux fournisseurs de mettre en œuvre une ou des solutions techniquement réalisables qui correspondent à l'état de la technique généralement reconnu, afin de satisfaire aux quatre exigences de qualité que sont l'efficacité, la solidité, la fiabilité et l'interopérabilité.

(83) On entend par l'«état de la technique» un état d'avancement d'une capacité technique à un moment donné, en ce qui concerne un produit, un processus ou un service, fondé sur des découvertes scientifiques, techniques et expérimentales pertinentes, et accepté comme étant une bonne pratique dans le domaine de la technologie ([28]). L'état de la technique n'implique pas nécessairement les recherches scientifiques les plus récentes encore à un stade expérimental ou dont la maturité technologique est insuffisante. Les fournisseurs doivent adapter en permanence leurs solutions de marquage et de détection en temps utile et de manière proportionnée à mesure que la technologie et l'état de la technique évoluent.

(84) À la lumière de l'article 50, paragraphe 7, du règlement sur l'IA, les fournisseurs peuvent s'appuyer sur un code de bonne pratique, jugé approprié, en vertu de cette disposition, pour démontrer le respect des obligations en matière de marquage et de détection prévues à l'article 50, paragraphes 2 et 5, du règlement sur l'IA, et qui correspond à l'état de la technique. Pour de plus amples informations concernant les effets de l'application d'un tel code et la surveillance des autres déployeurs qui n'adhèrent pas à un code, voir la section 8.1 ci-dessous.

(85) Les coûts de mise en œuvre de certaines solutions techniques de marquage et de détection peuvent également être pris en considération. Dans certains cas, décrits aux points suivants, les coûts peuvent être disproportionnés par rapport aux gains marginaux, dont l'utilité serait plus limitée pour ce qui est de la capacité de permettre aux êtres humains de reconnaître les contenus générés ou manipulés par une IA et pour garantir l'intégrité de l'écosystème informationnel et la confiance en celui-ci.

(86) Il peut donc y avoir des cas strictement définis dans lesquels une solution technique fondée sur des marquages de métadonnées moins solides est suffisante pour se conformer à l'article 50, paragraphe 2, du règlement sur l'IA, en tenant compte des spécificités des sorties du système d'IA si elles présentent intrinsèquement moins de risques de tromperie, de manipulation ou d'effets négatifs sur l'écosystème informationnel. C'est notamment le cas lorsqu'un système d'IA générative est intégré dans des produits physiques générant des sorties dans un environnement techniquement contrôlé et fermé qui est principalement à caractère instructif (par exemple, un système d'IA intégré dans des systèmes de navigation de véhicules), pour autant que des mesures techniques efficaces soient intégrées dans le produit pour empêcher que les sorties ne quittent son environnement, par exemple en étant captées ou exportées et diffusées à l'extérieur.

([28]) Voir l'annexe II de la décision d'exécution de la Commission C(2025) 3871 relative à une demande de normalisation adressée au Comité européen de normalisation et au Comité européen de normalisation électrotechnique concernant les systèmes d'IA à haut risque à l'appui du règlement (UE) 2024/1689 du Parlement européen et du Conseil, et abrogeant la décision d'exécution C(2023) 3215.

(87) Il peut également y avoir des cas limités d'«applications d'IA industrielles» ou d'«applications interentreprises» (à l'exclusion des systèmes d'IA publics et destinés aux consommateurs) dans lesquels le principe de proportionnalité et les avantages limités au regard des objectifs de transparence signifient qu'aucun marquage et aucune détection des sorties générées ou manipulées par une IA ne sont requis. Cela ne s'applique que si les conditions suivantes sont remplies cumulativement:

i) La sortie générée par le système d'IA est de nature strictement technique, par exemple des plans d'ingénierie, des processus de production industrielle, des instructions techniques, des sorties générées dans le cadre de processus de maintenance prédictive du système dans des environnements industriels, des processus de documentation interne ou des étapes et processus de production avant que la sortie ne soit finalisée et mise à la disposition d'autres personnes extérieures ou du public.

ii) La sortie générée par le système d'IA est uniquement destinée à être perçue et traitée par un nombre limité prédéfini de personnes physiques agissant à titre professionnel au sein de l'organisation du fournisseur et du déployeur.

iii) La sortie générée par le système d'IA ne doit pas être destinée à être partagée en dehors de l'entreprise ou à être utilisable par des personnes extérieures, des garanties appropriées étant mises en place pour éviter une mauvaise utilisation raisonnablement prévisible (par exemple, isolation au sein d'un nuage, contrôles fondés sur les rôles).

(88) La production en temps réel de contenus qui sont éphémères et consommés immédiatement, sans être enregistrés, stockés ou diffusés plus largement (par exemple dans des jeux vidéo ou des applications de réalité virtuelle), peut également être exemptée de l'obligation lorsque la technologie ne permet pas le marquage et que les personnes exposées au contenu sont informées que celui-ci est généré ou manipulé par une IA (par exemple, information au cours de l'expérience, notifications au niveau de la session).

4.3. Exceptions aux obligations prévues à l'article 50, paragraphe 2, du règlement sur l'IA

(89) L'article 50, paragraphe 2, du règlement sur l'IA prévoit trois exceptions explicites aux obligations de transparence énoncées dans ladite disposition.

(90) La première exception concerne les systèmes d'IA dans la mesure où ils remplissent une fonction d'assistance pour la mise en forme standard. La mise en forme standard doit s'entendre comme le processus consistant à préparer du contenu existant en vue de sa publication ou de sa distribution (par exemple, les petites modifications visant à améliorer la lisibilité et la grammaire, la qualité et le format) et n'entraîne pas la production de nouveaux contenus. La mise en forme standard vise à garantir que le contenu est, entre autres, exempt d'erreurs linguistiques manifestes sur le plan technique ou grammatical, qu'il respecte les exigences applicables en matière de mise en page, de présentation, de formatage ou d'accessibilité, ou qu'il est conforme aux pratiques sectorielles. Il ne s'agit plus de mise en forme standard si le contenu est modifié d'une manière substantielle (modifications de fond, changements structurels, etc.) qui a une incidence sur sa signification, son style ou son intention.

(91) La deuxième exception s'applique lorsqu'un système d'IA ne modifie pas de manière substantielle les données d'entrée fournies par le déployeur ou leur sémantique. Une modification devrait être considérée substantielle si les données d'entrée ou leur sémantique ont été sensiblement manipulées par le système d'IA au cours de son processus de production de sorties. Cette appréciation doit se fonder sur une évaluation des facteurs pertinents, tels que le format, le type de contenu médiatique, le style et les modifications du contenu qui ont une incidence sur sa signification, son style ou son intention. Le caractère substantiel d'une modification doit être évalué au cas par cas.

(92) Si un système d'IA peut être utilisé à la fois pour la production ou la manipulation de contenus et pour des modifications mineures non substantielles des données d'entrée, l'obligation de transparence prévue à l'article 50, paragraphe 2, du règlement sur l'IA ne s'appliquera pas au contenu ayant fait l'objet de modifications mineures et non substantielles.

Exemples de mise en forme standard et de modifications mineures bénéficiant de l'exception prévue à l'article 50, paragraphe 2, du règlement sur l'IA:

— correction grammaticale et vérification de l'orthographe, retouches linguistiques et stylistiques mineures qui ne modifient pas le fond, la signification, le style ou le message du texte, traductions de texte générées par une IA;

— formatage, conversions de format, compression technique, réduction ou suppression du bruit visant à améliorer la clarté sans modifier la signification ou le fond du contenu;

— recadrage mineur, ajustements ou corrections mineurs des couleurs, éclaircissement ou assombrissement, accentuation des contours visant à améliorer la clarté ou autres opérations de mise en forme et corrections techniques standard;

— élimination de taches de poussière causées par un objectif ou un capteur sale, élimination des yeux rouges causés par la photographie au flash; suppression et masquage des arrière-plans visibles dans le fichier d'origine, pixellisation ou floutage de visages;

— redimensionnement d'un clip vidéo, compression et égalisation de la plage dynamique; stabilisation vidéo limitée;

— ajustements standard de la couleur et du contraste, ajustements mineurs de la vitesse de lecture, corrections mineures pour niveler l'horizon d'une image, application de filtres de pixels pour amplifier certaines parties d'une image, ou application de cartes de couleurs à des images à niveaux de gris, reconstitution des bords d'une image, conversion d'une image ou d'une vidéo en noir et blanc en une image ou vidéo en couleur et inversement;

— remplissage de pixels visant à adapter le rapport d'aspect ou à améliorer la clarté, clips de transition automatiques, autres modifications non substantielles à des fins esthétiques et techniques;

— traitement technique ou reconstruction d'images médicales au sein de dispositifs médicaux (par exemple, annotations, cadres, contours, cercles, cartes thermiques);

— transcriptions de conversations; contenus générés par une IA qui ne font que transformer les entrées humaines authentiques au moyen de technologies d'assistance permettant aux personnes handicapées de communiquer (par exemple, communication améliorée et alternative ou voix neuronales personnalisées), étant donné qu'ils ne modifient pas sémantiquement la signification du contenu.

Exemples de modifications sémantiques nécessitant un marquage au titre de l'article 50, paragraphe 2, du règlement sur l'IA:

— résumés de texte générés par une IA; paraphrase ou réécriture de textes qui modifie le style, la structure et la signification au-delà d'une simple correction grammaticale et stylistique mineure;

— suppression, remplacement ou insertion d'objets ou de personnes dans des images et vidéos existantes lorsque ces opérations modifient la signification et le fond du contenu; remplacement de visages ou modification importante des visages;

— synthèse de paroles réalistes avec la voix d'une personne spécifique, ou production d'une vidéo réaliste représentant des événements qui ne se sont pas produits; modification de la forme du corps ou de la couleur de peau d'une personne; éclaircissement ou assombrissement extrêmes, changements extrêmes de la couleur et des contrastes, qui modifient la signification, l'intention et le message du contenu;

— création d'images composites ou de clips vidéo qui modifient la représentation de personnes, d'objets, d'événements ou de faits, ou toute autre modification substantielle du contenu.

(93) La troisième exception s'applique aux systèmes d'IA générative autorisés par la loi à générer ou à manipuler des contenus de synthèse à des fins de prévention ou de détection des infractions pénales, d'enquêtes ou de poursuites en la matière. Les précisions fournies aux points 46 à 48 ci-dessus pour cette exception sont également pertinentes dans le cas présent.

4.4. Interaction avec d'autres actes juridiques de l'Union

(94) Les obligations de marquage et de détection prévues à l'article 50, paragraphe 2, du règlement sur l'IA sont axées sur la manière dont le contenu a été créé et sur son origine artificielle, et non sur l'identité de l'auteur du contenu. Toute solution de marquage et de détection utilisée par les fournisseurs de systèmes d'IA doit donc être conforme au droit de l'Union applicable en matière de protection des données, y compris aux principes et aux obligations en matière de protection des données, tels que l'obligation de protection des données dès la conception et par défaut, la minimisation des données et les principes de sécurité et de confidentialité. Par exemple, aux fins du marquage et de la détection, les informations relatives au créateur du contenu ne devraient pas être traitées et toute autre donnée à caractère personnel traitée devrait être supprimée dès que la finalité de la détection est atteinte, la sécurité et la confidentialité doivent être garanties, etc.

(95) Le considérant 136 du règlement sur l'IA souligne en outre l'importance particulière des obligations de transparence prévues à l'article 50, paragraphes 1, 2 et 4, du règlement sur l'IA pour faciliter la mise en œuvre effective du DSA. Cela vaut en particulier pour les obligations incombant aux fournisseurs de très grandes plateformes en ligne (ci-après les «TGPL») ou de très grands moteurs de recherche en ligne (ci-après les «TGMRL») consistant à recenser, analyser, évaluer et atténuer de manière diligente les risques systémiques susceptibles de découler de la diffusion de contenus qui ont été générés ou manipulés par une IA, en particulier le risque d'effets négatifs réels ou prévisibles sur les processus démocratiques, le débat public et les processus électoraux, notamment par le biais de la désinformation (articles 34 et 35 du DSA). Les marquages lisibles par machine peuvent permettre à ces fournisseurs de détecter les contenus générés ou manipulés par une IA et de fournir un étiquetage et des informations appropriés, réduisant ainsi le risque de désinformation générée par l'IA dans l'ensemble de leurs services.

(96) Les obligations de marquage et d'étiquetage des contenus générés ou manipulés par les systèmes d'IA prévues à l'article 50, paragraphes 2 et 4, du règlement sur l'IA sont sans préjudice de l'obligation, prévue à l'article 16, paragraphe 6, du DSA, imposée aux fournisseurs de services d'hébergement de traiter les notifications au sujet de contenus illicites reçues conformément à l'article 16, paragraphe 1, du DSA ([29]). En particulier, le marquage ou l'étiquetage appliqué aux contenus générés ou manipulés par une IA ne devraient pas influencer l'évaluation et la décision quant à l'illégalité du contenu spécifique en vertu d'autres cadres réglementaires. Cette évaluation ne devrait être effectuée qu'au regard des règles régissant la légalité du contenu conformément à ces cadres ([30]). Par exemple, si une image d'hypertrucage étiquetée est signalée comme étant potentiellement du matériel relatif à des abus sexuels sur enfants ou comme ne respectant pas le droit des marques ou la législation relative au droit d'auteur, l'utilisation d'une étiquette n'a pas d'incidence sur l'illégalité potentielle du contenu en vertu de ces cadres réglementaires. Cette appréciation devrait être effectuée uniquement sur la base des lois applicables respectives (par exemple, le droit pénal, les règles relatives aux marques ou au droit d'auteur).

(97) À l'inverse, le marquage lisible par machine intégré dans le contenu conformément à l'article 50, paragraphe 2, du règlement sur l'IA peut aider les fournisseurs de TGPL et de TGMRL à s'acquitter des obligations qui leur incombent au titre des articles 34 et 35 du DSA, consistant à recenser, à analyser et à évaluer les risques systémiques susceptibles de découler de la diffusion de contenus qui ont été générés ou manipulés par une IA, en particulier le risque d'effets négatifs réels ou prévisibles sur les processus démocratiques, le débat public et les processus électoraux, notamment par le biais de la désinformation ([31]), et de mettre en place des mesures raisonnables, proportionnées et efficaces pour atténuer ces risques, et adaptées aux risques systémiques spécifiques recensés. Les fournisseurs de TGPL et de TGMRL peuvent également être considérés comme des fournisseurs de systèmes d'IA s'ils remplissent les conditions pertinentes prévues par le règlement sur l'IA lorsqu'ils intègrent des systèmes d'IA dans leurs services (voir également la section 6.1.5).

(98) Les fournisseurs de plateformes en ligne, de moteurs de recherche en ligne et d'autres acteurs pertinents qui distribuent du contenu tout au long de la chaîne de valeur sont en particulier encouragés à préserver et à permettre l'identification des marquages mis en œuvre par les fournisseurs conformément à l'article 50, paragraphe 2, du règlement sur l'IA, afin que les personnes physiques exposées au contenu puissent être informées de son origine.

5. ARTICLE 50, PARAGRAPHE 3, DU RÈGLEMENT SUR L'IA: SYSTÈMES DE RECONNAISSANCE DES ÉMOTIONS ET SYSTÈMES DE CATÉGORISATION BIOMÉTRIQUE

(99) L'article 50, paragraphe 3, du règlement sur l'IA impose aux déployeurs de systèmes de reconnaissance des émotions et de systèmes de catégorisation biométrique l'obligation d'informer les personnes physiques qui y sont exposées du fonctionnement du système.

(100) Le considérant 132 du règlement sur l'IA explique que l'objectif de cette obligation est de veiller à ce que les personnes physiques (y compris les personnes handicapées) soient conscientes qu'elles sont exposées à des systèmes de reconnaissance des émotions et de catégorisation biométrique qui peuvent porter atteinte à leur vie privée. L'obligation s'applique donc indépendamment du fait que les personnes soient exposées à de tels systèmes en temps réel ou que ces systèmes soient exploités a posteriori.

([29]) Voir aussi le considérant 11 du règlement sur l'IA.
([30]) Voir aussi le considérant 137 du règlement sur l'IA.
([31]) Considérant 136 du règlement sur l'IA.

5.1. **Principaux composants, concepts et obligations de transparence s'y rapportant au titre de l'article 50, paragraphe 3, du règlement sur l'IA**

5.1.1. *La notion de système de reconnaissance des émotions*

(101) L'article 3, point 39), du règlement sur l'IA définit un «système de reconnaissance des émotions» comme «un système d'IA permettant la reconnaissance ou la déduction des émotions ou des intentions de personnes physiques sur la base de leurs données biométriques». L'explication détaillée de cette notion fournie dans les lignes directrices de la Commission sur la classification des systèmes d'IA à haut risque, y compris les exemples pratiques qui y sont cités, est tout aussi pertinente dans le contexte de l'article 50, paragraphe 3, du règlement sur l'IA.

(102) Étant donné que tous les systèmes de reconnaissance des émotions sont également classés à haut risque, à moins qu'ils ne soient interdits sur le lieu de travail et dans les établissements d'enseignement en vertu de l'article 5, paragraphe 1, point f), du règlement sur l'IA, l'obligation de transparence prévue à l'article 50, paragraphe 3, dudit règlement devrait s'appliquer en conjonction avec les autres garanties et exigences applicables aux systèmes d'IA à haut risque.

5.1.2. *La notion de système de catégorisation biométrique*

(103) L'article 3, point 40), du règlement sur l'IA définit un «système de catégorisation biométrique» comme «un système d'IA destiné à affecter des personnes physiques à des catégories spécifiques sur la base de leurs données biométriques, à moins que cela ne soit accessoire à un autre service commercial et strictement nécessaire pour des raisons techniques objectives». L'explication détaillée de cette notion fournie dans les lignes directrices de la Commission sur la classification des systèmes d'IA à haut risque, y compris ses éléments spécifiques ainsi que les exemples pratiques qui y sont cités, est tout aussi pertinente dans le contexte de l'article 50, paragraphe 3, du règlement sur l'IA.

(104) Il est important de noter que, sauf interdiction en vertu de l'article 5, paragraphe 1, point g), du règlement sur l'IA, l'article 50, paragraphe 3, dudit règlement s'applique à tout système de catégorisation biométrique (par exemple, classification par âge ou par genre fondée sur des données biométriques), qu'il relève ou non du champ d'application des systèmes d'IA à haut risque au titre du règlement sur l'IA.

5.1.3. *L'obligation d'information prévue à l'article 50, paragraphe 3, du règlement sur l'IA*

(105) **Étendue des informations à fournir:** les déployeurs sont tenus d'informer les personnes physiques concernées qu'elles sont exposées à un système de reconnaissance des émotions ou à un système de catégorisation biométrique, qui est en fonctionnement. Le règlement sur l'IA n'exige pas d'inclure des informations sur les raisons du fonctionnement du système ou, si le système traite également des informations à des fins autres que la reconnaissance des émotions ou la catégorisation biométrique, qui doivent être divulguées en vertu du droit de l'Union en matière de protection des données (voir la section 5.3.).

(106) **Destinataires des informations:** l'article 50, paragraphe 3, du règlement sur l'IA exige que les déployeurs informent toutes les personnes physiques exposées au fonctionnement du système d'IA, y compris les enfants.

(107) **Moyens de communication des informations:** le règlement sur l'IA ne prescrit pas de moyen de communication spécifique, si ce n'est que son article 50, paragraphe 5, exige que les informations soient fournies de manière claire et reconnaissable et conformément aux exigences applicables en matière d'accessibilité (en répondant aux besoins des personnes physiques handicapées). La manière dont les personnes physiques exposées au système doivent être informées peut dépendre du lieu de déploiement (par exemple, plateforme de jeu virtuelle, magasin physique, gare ferroviaire), des destinataires potentiels (par exemple, enfants, personnes âgées, personnes handicapées, clients), de la durée et du type de l'exposition (par exemple, longue durée, de manière ponctuelle, continue, intermittente) et de l'existence éventuelle d'un contact entre le déployeur et les personnes exposées (par exemple, s'il existe un canal de communication entre le déployeur et la personne). En fonction du cas d'utilisation, les informations peuvent être fournies par écrit, au moyen d'icônes normalisées (y compris lorsqu'elles sont présentées par voie électronique), oralement ou en utilisant une combinaison de tous ces moyens.

Exemples de méthodes d'information des personnes physiques:

— une fenêtre contextuelle placée au centre de l'écran avec un message d'accueil avant le lancement d'un jeu vidéo indiquant que le visage du joueur est enregistré et que ses émotions sont capturées;

— un panneau d'information visible à chaque entrée possible d'une salle d'exposition indiquant que les images faciales des visiteurs sont capturées à l'entrée de la salle afin de les affecter à un groupe d'âge spécifique.

(108) **Moment de la fourniture des informations:** les informations doivent être fournies au plus tard au moment de la première interaction avec le système d'IA ou de la première exposition à celui-ci, conformément à l'article 50, paragraphe 5, du règlement sur l'IA. Le cas échéant, il n'est pas exclu de fournir à l'avance les informations requises.

5.2. Exception à des fins répressives

(109) L'obligation d'information ne s'applique pas aux systèmes de reconnaissance des émotions et de catégorisation biométrique dont la loi autorise l'utilisation à des fins de prévention ou de détection des infractions pénales ou d'enquêtes en la matière, sous réserve de garanties appropriées pour les droits et libertés des tiers et conformément au droit de l'Union. Les précisions fournies pour cette exception aux points 46 à 48 ci-dessus sont également pertinentes dans le cas présent, à une différence près. Si les autres exceptions prévues à l'article 50 du règlement sur l'IA exigent une loi explicite autorisant l'utilisation non transparente des systèmes d'IA, l'article 50, paragraphe 3, dudit règlement rend cette utilisation possible lorsque les règles légales régissant les pouvoirs des autorités répressives les autorisent à utiliser des systèmes de reconnaissance des émotions et de catégorisation biométrique pour détecter ou prévenir les infractions pénales ou enquêter en la matière sans en informer les personnes concernées, sous réserve de garanties appropriées.

5.3. Interaction avec d'autres actes juridiques de l'Union

(110) L'obligation d'information prévue à l'article 50, paragraphe 3, du règlement sur l'IA ne rend pas, à elle seule, licite l'utilisation d'un système de reconnaissance des émotions et de catégorisation biométrique. Elle ne légitime pas non plus les utilisations intrusives ou discriminatoires qui pourraient être illégales et interdites en vertu de l'article 5 du règlement sur l'IA ou d'autres dispositions du droit de l'Union. Les déployeurs doivent se conformer à l'obligation d'information prévue à l'article 50, paragraphe 3, du règlement sur l'IA, ainsi qu'à toute exigence d'information applicable en vertu du droit de l'Union en matière de protection des données. Dans certains cas, les déployeurs peuvent envisager d'ajouter des notifications au titre de l'article 50, paragraphe 3, du règlement sur l'IA aux informations fournies aux personnes concernées en vertu du droit de l'Union en matière de protection des données.

6. ARTICLE 50, PARAGRAPHE 4, DU RÈGLEMENT SUR L'IA: ÉTIQUETAGE DES HYPERTRUCAGES ET DE CERTAINES PUBLICATIONS TEXTUELLES

(111) L'article 50, paragraphe 4, du règlement sur l'IA établit deux obligations distinctes destinées aux déployeurs de systèmes d'IA générative exigeant qu'ils indiquent de manière claire et reconnaissable i) les hypertrucages et ii) les textes générés ou manipulés par une IA publiés dans le but d'informer le public sur des questions d'intérêt public, sauf dans des cas précis. Ces obligations en matière d'étiquetage s'appliquent en outre et sans préjudice des obligations en matière de marquage et de détection lisibles par machine prévues à l'article 50, paragraphe 2, du règlement sur l'IA applicables aux fournisseurs de systèmes d'IA qui génèrent ou manipulent des contenus de synthèse de type images, vidéo, audio ou texte.

6.1. Principaux composants, concepts et obligation(s) de transparence s'y rapportant pour les hypertrucages au titre de l'article 50, paragraphe 4, du règlement sur l'IA

(112) L'article 50, paragraphe 4, premier alinéa, du règlement sur l'IA s'applique si plusieurs conditions sont remplies:

i) le système doit être un système d'IA;

ii) utilisé par les déployeurs à des fins professionnelles (voir les sections 2.3 et 2.4.1);

iii) pour générer ou manipuler des images ou des contenus audio ou vidéo qui constituent un hypertrucage;

iv) le système d'IA ne doit pas relever de l'exception relative à l'utilisation autorisée par la loi à des fins de prévention ou de détection des infractions pénales, d'enquêtes ou de poursuites en la matière.

6.1.1. *La notion d'«hypertrucage»*

(113) L'article 3, point 60), du règlement sur l'IA définit l'«hypertrucage» comme «une image ou un contenu audio ou vidéo généré ou manipulé par l'IA, présentant une ressemblance avec des personnes, des objets, des lieux, des entités ou événements existants et pouvant être perçu à tort par une personne comme authentiques ou véridiques». La différence entre la manipulation et la production a été clarifiée à la section 4.1 ci-dessus. Afin de clarifier la notion d'hypertrucage, quatre critères cumulatifs de cette définition sont examinés ci-dessous: i) la ressemblance avec ii) des personnes, objets, lieux, entités ou événements iii) existants iv) qui seraient perçus à tort par une personne comme authentiques ou véridiques.

i) **La ressemblance:** l'article 3, point 60), du règlement sur l'IA précise que les images et les contenus audio ou vidéo générés ou manipulés par une IA doivent «ressembler» à des sujets existants pour constituer un hypertrucage. Le considérant 134 du règlement sur l'IA ajoute qu'une telle ressemblance devrait être «sensible». Les images et les contenus audio ou vidéo générés ou manipulés par une IA ressemblent sensiblement à un sujet s'il existe un degré élevé de similitude entre le contenu de l'hypertrucage et le sujet (y compris n'importe lequel de ses éléments reconnaissables) simulé par l'hypertrucage. Le contenu ne doit pas nécessairement être identique au sujet. Le caractère sensible du niveau de ressemblance est à évaluer au cas par cas par le déployeur sur la base d'une comparaison objective entre le sujet simulé et le contenu de l'hypertrucage, tenant compte, entre autres, du degré de représentation des caractéristiques ou éléments distinctifs du sujet par l'hypertrucage.

ii) **Existants:** l'article 3, point 60), du règlement sur l'IA précise que les images ou les contenus audio ou vidéo générés ou manipulés par une IA doivent ressembler à un ou plusieurs sujets «existants» pour constituer un hypertrucage. Pour remplir efficacement l'objectif de l'article 50, paragraphe 4, du règlement sur l'IA, qui vise à réduire les risques d'usurpation d'identité, de tromperie, de désinformation, de manipulation et de fraude, le critère relatif à l'existence des sujets veut que l'image ou le contenu audio ou vidéo généré ou manipulé par une IA doive ressembler à des sujets réalistes (par exemple, des personnes, des objets, des lieux). Par conséquent, il suffit que des personnes, des objets, des lieux, des entités ou des événements simulés ressemblent à une personne ou à une chose qui existe, qui peut plausiblement exister ou qui aurait pu plausiblement exister dans la réalité pour que ce type de contenu soit considéré comme un hypertrucage, étant donné que de tels contenus peuvent être une source majeure de tromperie, de manipulation et de désinformation. En revanche, les personnes, objets, lieux, entités ou événements simulés qui, par exemple, défient les lois de la nature ou de la physique ou représentent des formes de vie qui ne sont pas communément acceptées en biologie (par exemple, les êtres humains volant sans aide mécanique, les dragons, ou les éléphants conduisant des voitures) et qui ne sont pas susceptibles d'induire une personne physique en erreur sont considérés irréalistes et ne relèvent donc pas du champ d'application de l'obligation de transparence.

iii) **Des personnes, des objets, des lieux, des entités ou des événements:** l'article 3, point 60), du règlement sur l'IA précise que les images ou contenus audio ou vidéo générés ou manipulés par une IA devraient ressembler à «des personnes, des objets, des lieux, des entités ou des événements» existants pour constituer un hypertrucage. Par «personnes», on entend des êtres humains réalistes (y compris les répliques numériques de personnes réelles, les avatars ou représentations humaines réalistes générés par une IA, et les caractéristiques ou expressions personnelles, telles que l'image, la voix, le comportement, les performances, etc.). Par «objets», on entend des objets matériels réalistes et inanimés, y compris des bâtiments, des œuvres d'art, des machines, des biens de consommation, etc. Par «lieux», on entend des endroits réalistes. Par «entités», on entend des êtres réalistes, non humains mais animés, y compris des animaux ou d'autres formes de vie biologiques. Par «événements», on entend des scènes ou des situations réalistes qui peuvent faire intervenir des personnes, des objets, des lieux et des entités (par exemple, l'évocation d'événements historiques ou la représentation de services professionnels ou de services aux consommateurs).

iv) **Perception à tort, par une personne, du contenu comme étant authentique ou véridique:** l'article 3, point 60), du règlement sur l'IA précise que l'image ou le contenu audio ou vidéo généré ou manipulé par une IA, présentant une ressemblance avec des personnes, des objets, des lieux, des entités ou des événements existants «[peuvent] être perçus à tort par une personne comme authentiques ou véridiques». Ce quatrième critère se rapporte à la caractéristique essentielle des hypertrucages et à leur capacité à potentiellement tromper ou induire en erreur une personne en ce qui concerne l'authenticité ou la véracité de leur contenu ([32]). L'authenticité du contenu désigne le fait que le contenu est réellement ce qu'il prétend être au regard de sa source ou de son processus de création (y compris, par exemple, la participation d'êtres humains ou d'animaux réels et leur comportement ou leurs actions réels, l'apparence ou l'utilisation réelles d'objets, la fourniture de services, le déroulement exact d'un événement ou son évolution précise). La véracité concerne la véracité du contenu (par exemple, l'exactitude factuelle des représentations dans l'hypertrucage).

([32]) Par conséquent, ce critère doit être compris indépendamment et de manière clairement distincte de la notion de tromperie telle que mentionnée par la directive 2005/29/CE relative aux pratiques commerciales déloyales des entreprises vis-à-vis des consommateurs dans le marché intérieur (directive sur les pratiques commerciales déloyales), https://eur-lex.europa.eu/eli/dir/2005/29/oj/fra.

(114) La possibilité qu'un contenu soit «perçu à tort par une personne comme authentique ou véridique» devrait être évaluée dans son ensemble, en tenant compte du niveau de ressemblance, du message substantiel potentiel du contenu, des contextes de déploiement prévus et prévisibles, de l'environnement dans lequel le contenu est présenté, ainsi que de la composition du public visé et raisonnablement prévisible et de ses attentes. Toutefois, cette évaluation est objective et n'exige pas que l'intention du déployeur soit de tromper ou d'induire en erreur les personnes physiques exposées au contenu pour que celui-ci constitue un hypertrucage. En ce qui concerne le contexte du déploiement et les attentes du public, cela signifie que si le contenu est utilisé dans des contextes de déploiement spécifiques où le public visé et raisonnablement prévisible ne s'attend pas à ce que le contenu soit authentique ou véridique, le contenu généré ou manipulé par une IA peut ne pas être perçu *à tort* comme authentique ou véridique même s'il n'est pas authentique ou véridique dans les faits. Dans de tels cas, le quatrième critère de la définition des hypertrucages ne sera pas rempli. Par exemple, la production ou la manipulation par une IA de scènes d'arrière-plan, d'effets spéciaux ou de prétraitement et de post-traitement techniques dans le cadre de processus de production cinématographique standard ne sont pas susceptibles de faire passer *à tort* le contenu comme authentique ou véridique auprès du public. À l'inverse, l'utilisation de l'IA pour produire ou manipuler des éléments essentiels ayant une incidence sur la perception du public est susceptible de le faire (tels que les acteurs entièrement générés par une IA, les répliques numériques d'acteurs réels ou décédés, le rajeunissement numérique, les performances simulées, y compris certaines parties de celles-ci, ou la représentation non authentique ou mensongère de personnes, d'objets, de lieux ou d'événements dans des documentaires). En ce qui concerne les images et les contenus vidéo, il est admis qu'un degré élevé de photoréalisme rend plus probable le fait que ces contenus soient considérés comme des hypertrucages (étant donné qu'ils ressemblent probablement à des sujets existants), mais le photoréalisme à lui seul n'est pas déterminant pour l'évaluation. Ce qu'il convient encore d'évaluer, c'est la capacité du contenu à tromper ou à induire en erreur une personne quant à son authenticité ou à sa véracité, tout en tenant compte, entre autres, du contexte de déploiement prévu et prévisible.

(115) Contrairement à ce que prévoit l'article 50, paragraphe 1, du règlement sur l'IA, l'évaluation de la possibilité qu'un contenu soit perçu à tort par une personne comme authentique ou véridique ne devrait pas se fonder sur une hypothétique personne «moyenne» censée être exposée au contenu (voir la section 3.2.1), mais devrait dûment tenir compte de la diversité possible de la composition du public raisonnablement prévisible, susceptible d'être exposé à l'hypertrucage. Il est particulièrement important de tenir dûment compte de la composition du public s'il est raisonnablement prévisible que l'hypertrucage puisse, à un moment donné, être montré à des enfants, des personnes âgées ou d'autres groupes de personnes ayant un faible niveau de compétence numérique et de maîtrise de l'IA ou de connaissances générales, étant donné qu'ils peuvent être plus facilement trompés ou induits en erreur quant à l'authenticité ou à la véracité du contenu. Dans de tels cas, le fait que le contenu soit perçu à tort comme authentique ou véridique par cette partie du public prévisible peut suffire à le qualifier d'hypertrucage. Toutefois, les déployeurs ne sont pas censés envisager une éventuelle diffusion ultérieure de l'hypertrucage par des tiers au-delà du public raisonnablement prévisible par ces déployeurs, compte tenu des canaux de distribution prévus (par exemple, si l'hypertrucage est présenté uniquement sur la partie d'un site web réservée aux abonnés ou dans une lettre d'information d'entreprise, cela ne signifie pas que le déployeur doive, par défaut, considérer que le contenu est largement accessible au public).

(116) Dans le même ordre d'idées, la manipulation, assistée par une IA, d'aspects matériels ou techniques insignifiants d'un contenu préexistant peut avoir une importance mineure pour l'appréciation par une personne de l'authenticité ou de la véracité du contenu, sans pour autant faire du contenu qui en résulte un hypertrucage. Ce type de manipulation pourrait inclure, par exemple, la modification de détails de l'arrière-plan (par exemple, la suppression de passants), les ajustements d'éclairage, l'adaptation des paramètres audio, la correction des couleurs, la réduction ou la suppression du bruit, l'amélioration de l'accessibilité ou la compression de fichiers, ainsi que les ajustements et améliorations esthétiques. Le classement de la manipulation dans cette catégorie dépend du contexte et de l'incidence qu'elle a sur la perception par les personnes de l'authenticité ou de la véracité du contenu dans un cas donné. Par exemple, les opérations fondées sur l'IA comme la correction des couleurs, les extensions de l'arrière-plan de contenus existants, les ajustements ou les remplacements d'arrière-plan à des fins clairement esthétiques, la composition et l'agencement de produits existants, ou le redimensionnement d'images utilisées dans des annonces publicitaires de produits ou sur leurs emballages, sont susceptibles de n'avoir qu'une incidence mineure sur la perception par une personne de l'authenticité et de la véracité de la publicité et du produit. À l'inverse, une modification substantielle, assistée par une IA, des détails de l'arrière-plan d'images journalistiques allant au-delà des pratiques techniques et éditoriales habituelles peut avoir une incidence négative sur la perception par une personne de l'authenticité et de la véracité du contenu.

Exemples d'hypertrucages au sens de l'article 3, point 60), du règlement sur l'IA:

— image manipulée par une IA d'une scène faisant intervenir deux véritables footballeurs professionnels devant un bâtiment ressemblant à un stade de football;

— contenu audio généré par une IA comportant le clonage vocal des présentateurs habituels d'un podcast de presse et d'un interlocuteur invité discutant de certains événements d'actualité;

— vidéo générée par une IA d'une personne ressemblant à une personnalité politique, tenant un discours devant un public;

— vidéo générée par une IA incluant une représentation générée par une IA d'un influenceur célèbre dans un contexte publicitaire ou promotionnel;

— vidéo générée par une IA présentant un avatar de synthèse réaliste d'un PDG d'entreprise félicitant ses salariés pour leur travail et pour les résultats de l'entreprise au cours de l'année écoulée;

— image générée par une IA d'un produit dans une annonce publicitaire ou sur un emballage, susceptible d'influencer la perception du public et de l'induire en erreur quant à l'apparence, aux caractéristiques ou à l'utilisation réelles du produit (par exemple, en donnant au produit une apparence différente de celle du produit réel, en le rendant plus attrayant ou en lui conférant une qualité supérieure à celle qu'il a dans la vie réelle).

Exemples ne constituant pas des hypertrucages au sens de l'article 3, point 60), du règlement sur l'IA:

— image générée par une IA d'une scène présentant un sphinx survolant la tour Eiffel;

— vidéo générée par une IA de souris se disputant, en langage humain, pour savoir quel est le meilleur type de fromage dans le cadre d'une campagne publicitaire pour un fabricant de fromage;

— reproduction vocale par une IA de personnages de fiction (par exemple, pour des audiolivres, des jeux ou des animations) lorsqu'il n'y a pas de tromperie quant à l'identité des narrateurs;

— manipulation par une IA d'une émission de radio consistant à régler les paramètres audio techniques (par exemple, normalisation des niveaux de volume, réduction du bruit, compression audio) sans modifier les mots prononcés par les intervenants ou leur manière de parler;

— dessin animé généré par une IA à partir d'une image préexistante représentant un événement historique;

— vidéo d'un présentateur d'une émission télévisée scientifique qui traverse une région montagneuse tout en parlant du recul des glaciers, lequel est illustré par une animation d'arrière-plan manipulée par une IA montrant comment les glaciers ont reculé depuis leurs niveaux historiques jusqu'à leurs niveaux actuels;

— environnements fictifs générés par une IA (par exemple, forêts et châteaux fictifs) dans des jeux vidéo;

— films dans lesquels de véritables acteurs existants (non manipulés par une IA) jouent sur un arrière-plan généré par une IA (par exemple, représentant une étendue fictive de prairie, des bâtiments historiques de villes antiques, etc.);

— présentation, dans une annonce publicitaire, d'un produit réel (par exemple, une voiture) sur un arrière-plan et dans un environnement générés par une IA, pour autant que la publicité ne soit pas susceptible d'induire le public en erreur sur la représentation du produit et sur ses caractéristiques et son utilisation réelles.

6.1.2. *L'obligation d'information prévue à l'article 50, paragraphe 4, premier alinéa, du règlement sur l'IA*

(117) L'article 50, paragraphe 4, premier alinéa, du règlement sur l'IA oblige les déployeurs de systèmes d'IA qui génèrent ou manipulent des images ou des contenus audio ou vidéo constituant un hypertrucage à «[indiquer] que les contenus [de type hypertrucage] ont été générés ou manipulés par une IA». Le considérant 134 du règlement sur l'IA précise que les déployeurs utilisant des systèmes d'IA pour générer ou manipuler des contenus de type

hypertrucage devraient déclarer de manière claire et reconnaissable que ces contenus ont été créés ou manipulés par une IA en étiquetant les sorties d'IA en conséquence et en mentionnant leur origine artificielle. Les méthodes d'étiquetage ou d'information appliquées conformément à l'article 50, paragraphe 4, premier alinéa, du règlement sur l'IA devraient être compréhensibles et perceptibles par les personnes physiques (par exemple, au moyen d'étiquettes visibles ou sonores), sans qu'elles aient besoin d'utiliser des outils techniques particuliers ou d'effectuer des actions spécifiques. Par conséquent, les déployeurs ne peuvent pas s'appuyer sur le marquage lisible par machine intégré dans le contenu par le fournisseur en vertu de l'article 50, paragraphe 2, du règlement sur l'IA, étant donné que ces marquages ne sont pas immédiatement clairs et reconnaissables pour les personnes physiques exposées à des contenus de type hypertrucage.

(118) À la lumière de l'article 50, paragraphe 7, du règlement sur l'IA, les déployeurs peuvent adhérer à un code de bonne pratique, jugé approprié conformément audit article, afin de démontrer le respect de cette obligation d'information. Pour de plus amples informations concernant les effets de l'application d'un tel code et la surveillance des autres déployeurs qui n'adhèrent pas à un code, voir la section 8.1 ci-dessous.

6.1.3. *Transparence des hypertrucages artistiques, créatifs, satiriques, fictifs ou analogues*

(119) L'article 50, paragraphe 4, premier alinéa, du règlement sur l'IA prévoit une obligation de transparence assouplie pour les hypertrucages faisant partie d'œuvres ou de programmes manifestement artistiques, créatifs, satiriques, fictifs ou analogues. Pour ce type de contenu, l'obligation de transparence se limite à la divulgation de l'hypertrucage d'une manière appropriée qui n'entrave pas l'affichage ou la jouissance de l'œuvre.

(120) **Catégories de contenu pertinentes**: pour bénéficier de l'obligation de transparence limitée, les hypertrucages (image, contenu audio ou vidéo) doivent faire partie, au moins, de l'une des catégories de contenu suivantes: œuvres ou programmes artistiques, créatifs, satiriques, fictifs ou analogues. Le classement d'un hypertrucage dans l'une ou plusieurs de ces catégories doit être évalué au cas par cas par le déployeur. Aux fins de l'article 50, paragraphe 4, du règlement sur l'IA, ces catégories s'entendent comme suit:

— œuvres artistiques: œuvres qui ont été créées à des fins artistiques, y compris la musique, les œuvres cinématographiques et les arts visuels;

— œuvres créatives: œuvres qui supposent des choix créatifs. Les œuvres dont la création est principalement motivée par des considérations fonctionnelles ou techniques ne peuvent être qualifiées de créatives;

— œuvres satiriques: œuvres qui visent à critiquer la société, la politique, les entreprises ou des personnalités publiques au moyen de techniques humoristiques (y compris l'ironie, le sarcasme, la moquerie, le pastiche, etc.);

— œuvres fictives: œuvres culturelles présentant des personnes, des objets, des lieux, des entités ou des événements, dans un cadre imaginaire mais vraisemblable;

— œuvres analogues: œuvres partageant des caractéristiques fondamentales ou des éléments expressifs ou fonctionnels semblables à ceux des catégories susmentionnées, mais qui ne relèvent pas clairement d'une seule catégorie.

(121) Les catégories de contenu susmentionnées peuvent également s'appliquer à un ou plusieurs programmes qui devraient s'entendre comme un seul élément dans le cadre d'une grille ou d'un catalogue établi par un fournisseur de services de médias, dont la forme et le contenu sont comparables à ceux de la radiodiffusion télévisuelle ([33]). L'emploi du terme «œuvres» dans le contexte de l'article 50, paragraphe 4, du règlement sur l'IA n'a aucune incidence sur le fait que le contenu puisse bénéficier ou non d'une protection au titre de la législation de l'Union relative au droit d'auteur et aux droits voisins.

([33]) Voir, dans le même esprit, l'article 1, point b), de la directive SMA.

(122) **Manifestement:** le fait qu'un hypertrucage relève d'une (ou de plusieurs) des catégories de contenu susmentionnées devrait être manifeste pour les personnes physiques qui y sont exposées. Cela signifie que les catégories devraient être interprétées de manière stricte, cette exigence étant justifiée par l'application d'un régime d'information plus souple en raison de l'incidence potentielle de l'obligation sur la liberté d'expression et la liberté des arts et des sciences, tout en veillant néanmoins à traiter de manière adéquate les risques de désinformation, de manipulation et de tromperie. Par conséquent, les contenus dont la nature est potentiellement floue ou ambiguë pour le public sont exclus du champ d'application de ce régime de transparence assoupli. Les facteurs pertinents pour évaluer la nature manifeste des contenus artistiques, créatifs, satiriques ou fictifs sont les suivants: i) si le contenu affiche des formats ou des styles caractéristiques des catégories de contenu (par exemple, ironie ou exagération pour les œuvres satiriques, certains styles d'art, affichage de thèmes, techniques ou esthétiques distinctifs, etc.); ii) le contexte dans lequel le contenu est présenté (par exemple, si la plateforme, le support ou le lieu où le contenu est présenté, mis à disposition ou distribué est associé à un usage artistique, créatif, satirique ou fictif); et iii) les attentes du public (par exemple, un film, un environnement de jeu ou une scène de réalité virtuelle). En outre, la condition selon laquelle le contenu doit «manifestement» relever de l'une des catégories de contenu susmentionnées exclut le contenu du champ d'application de ces catégories si sa nature est exclusivement informative ou commerciale et est reconnaissable en tant que telle (par exemple, les reportages d'actualité). À cet égard, certains types de contenus (par exemple, les publicités ou les documentaires) contenant des hypertrucages pourraient être considérés comme des œuvres manifestement créatives ou fictives dans certaines situations spécifiques, mais pas dans d'autres, étant donné que l'évaluation se fait au cas par cas. Lorsque l'hypertrucage combine plusieurs caractères (informatifs et créatifs, par exemple), le caractère informatif devrait toujours prévaloir et exiger le respect des exigences d'étiquetage standard.

(123) **Divulgation appropriée n'entravant pas l'affichage ou la jouissance de l'œuvre**: les hypertrucages qui font partie d'œuvres ou de programmes manifestement artistiques, créatifs, satiriques ou fictifs ne sont pas exclus de l'obligation de transparence prévue à l'article 50, paragraphe 4, premier alinéa, du règlement sur l'IA. Les déployeurs doivent quand même indiquer l'origine artificielle du contenu ou sa manipulation par une IA, mais ils peuvent le faire d'une manière appropriée qui n'entrave pas l'affichage ou la jouissance de l'œuvre. Le considérant 134 du règlement sur l'IA précise que cette divulgation appropriée ne devrait pas entraver l'exploitation et l'utilisation normales de l'œuvre, tout en permettant de maintenir son utilité et sa qualité. Le caractère approprié des mesures de divulgation fait l'objet d'une évaluation au cas par cas, dans le cadre de laquelle les déployeurs peuvent prendre en considération tous les facteurs pertinents (y compris la nature de l'œuvre, le public, le contexte, etc.). En tout état de cause, les déployeurs doivent se conformer à l'article 50, paragraphe 5, du règlement sur l'IA (voir la section 7). À la lumière de l'article 50, paragraphe 7, du règlement sur l'IA, les déployeurs peuvent adhérer à un code de bonne pratique, jugé approprié en vertu de cette disposition, afin de démontrer le respect de cette obligation de divulgation. Pour de plus amples informations concernant les effets de l'application d'un tel code et la surveillance des autres déployeurs qui n'adhèrent pas à un code, voir la section 8.1 ci-dessous.

(124) **Garanties appropriées pour les droits et libertés de tiers:** le considérant 134 du règlement sur l'IA précise que le respect de l'obligation de transparence limitée et l'utilisation d'hypertrucages artistiques, créatifs, satiriques ou fictifs doivent se faire «sous réserve de garanties appropriées pour les droits et libertés de tiers». Par conséquent, les déployeurs doivent veiller à ce que les droits et libertés de tiers (y compris, par exemple, leur droit à la protection des données à caractère personnel ou leurs droits de propriété intellectuelle) soient protégés et respectés de manière adéquate lors de la création ou de la publication d'hypertrucages, même si ceux-ci sont considérés comme artistiques, créatifs, satiriques ou fictifs. L'invocation de l'obligation de transparence limitée ne peut pas justifier le non-respect des droits fondamentaux des personnes ou des titulaires de droits en vertu du droit de l'Union en matière de propriété intellectuelle ou de protection des données ([34]).

Exemples d'hypertrucages qui constituent des œuvres créatives, satiriques, fictives ou analogues:

— œuvre artistique/fictive: films (ou bandes-annonces de film) contenant des hypertrucages générés ou manipulés par une IA présentant des acteurs existants rajeunis ou des répliques numériques d'acteurs ou d'autres personnes décédés, disponibles dans un cinéma ou sur une plateforme de diffusion en continu;

— œuvre artistique/créative: musique générée par une IA, quel que soit son genre, ressemblant au style particulier d'artistes existants;

([34]) Cela vaut également pour les textes publiés informant le public sur des questions d'intérêt public.

— œuvre satirique/fictive: image d'une personnalité politique existante manipulée par une IA la mettant en scène dans une situation manifestement destinée à critiquer de manière humoristique certaines décisions politiques prises ou soutenues par cette personne;

— œuvre créative/fictive analogue: images de jeux vidéo générées par une IA comportant des simulations de personnes réelles et existantes sous forme d'hypertrucages.

Exemples d'hypertrucages qui ne constituent pas une œuvre artistique, créative, satirique, fictive ou analogue:

— vidéo manipulée par une IA dans le style d'une chaîne de téléachat comportant des hypertrucages d'êtres humains faisant la publicité d'un produit dans une scène générée par une IA présentant l'utilisation fonctionnelle du produit par des consommateurs simulés dans le but de persuader les téléspectateurs d'acheter le produit;

— image de célébrités générée par une IA suggérant la participation de ces personnalités à des activités qui ne se sont jamais produites, sans aucune visée fictive, satirique ou analogue;

— vidéo manipulée par une IA montrant un influenceur virtuel réaliste testant un produit réel sponsorisé, axé uniquement sur la présentation des fonctionnalités du produit;

— vidéo générée par une IA représentant des scènes d'holocauste réalistes partagées sur des plateformes de médias sociaux accessibles au public.

6.1.4. *Exception à des fins répressives*

(125) Si l'utilisation d'un hypertrucage (par exemple, une vidéo générée par une IA simulant des événements qui ne se sont jamais produits) est autorisée par la loi à des fins de prévention ou de détection des infractions pénales, d'enquêtes ou de poursuites en la matière, ses déployeurs sont exemptés de l'obligation de transparence prévue à l'article 50, paragraphe 4, premier alinéa, du règlement sur l'IA. Les précisions fournies aux points 46 à 48 ci-dessus concernant une exception similaire prévue à l'article 50, paragraphe 1, du règlement sur l'IA sont également pertinentes pour cette disposition.

6.1.5. *Interaction avec d'autres actes juridiques de l'Union*

(126) L'obligation de transparence prévue à l'article 50, paragraphe 4, premier alinéa, du règlement sur l'IA interagira avec les obligations incombant aux fournisseurs de TGPL et de TGMRL de recenser et d'atténuer les risques systémiques conformément aux articles 34 et 35 du DSA, en particulier son article 35, paragraphe 1, point k) ([35]). Cette dernière disposition énumère, parmi les éventuelles mesures d'atténuation des risques que les fournisseurs de TGPL et de TGMRL peuvent adopter, «le recours à un marquage bien visible pour garantir qu'un élément d'information, qu'il s'agisse d'une image, d'un contenu audio ou vidéo généré ou manipulé, qui ressemble nettement à des personnes, à des objets, à des lieux ou à d'autres entités ou événements réels et apparaît à tort aux yeux d'une personne comme authentique ou digne de foi, est reconnaissable lorsqu'il est présenté sur leurs interfaces en ligne». En complément, les fournisseurs de TGPL et de TGMRL peuvent également mettre à disposition une fonctionnalité facile d'utilisation permettant aux destinataires du service de signaler ce type d'information. Bien que cette disposition du DSA paraisse semblable à l'article 50, paragraphe 4, du règlement sur l'IA, elle la complète et contient deux distinctions importantes:

— **Champ d'application matériel différent:** par rapport à l'article 50, paragraphe 4, du règlement sur l'IA, l'article 35, paragraphe 1, point k), du DSA s'applique à un éventail plus large de contenus qui peuvent apparaître à tort comme authentiques ou véridiques, quelles que soient les technologies utilisées pour créer le contenu (c'est-à-dire qu'il s'applique aux contenus qui ont été générés ou manipulés à l'aide de systèmes d'IA ou de tout autre outil).

([35]) Voir également la section 4.4 pour de plus amples informations concernant l'interaction entre l'article 50 du règlement sur l'IA et le DSA.

— **Champ d'application personnel différent:** l'article 50, paragraphe 4, du règlement sur l'IA impose aux déployeurs de systèmes d'IA générant ou manipulant des hypertrucages d'étiqueter ces contenus, tandis que l'article 35, paragraphe 1, point k), du DSA s'applique aux fournisseurs de TGPL et de TGMRL diffusant ces contenus ou permettant de les trouver au moyen d'une recherche. Par exemple, les déployeurs de systèmes d'IA intégrés dans une TGPL au titre du règlement sur l'IA peuvent être considérés comme des «destinataires du service» au sens de l'article 3, point b), du DSA ([36]) s'ils utilisent la TGPL pour diffuser du contenu généré avec le système d'IA concerné, tandis que la TGPL peut être considérée comme le fournisseur du système d'IA si les conditions de l'article 3, point 3), du règlement sur l'IA sont remplies. Lorsque les fournisseurs de TGPL ou de TGMRL mettent à la disposition de ces déployeurs des outils d'étiquetage leur permettant d'étiqueter leurs hypertrucages conformément à l'article 50, paragraphe 4, du règlement sur l'IA (c'est-à-dire une indication claire et reconnaissable de leur origine artificielle), les déployeurs peuvent s'appuyer sur ces outils pour remplir leur obligation de transparence au titre de cette disposition dans le contexte des TGPL ou TGMRL utilisés. La mise à disposition d'une telle fonctionnalité est sans préjudice de la responsabilité des déployeurs en vertu du règlement sur l'IA de s'acquitter de leurs obligations en matière d'étiquetage au titre de l'article 50, paragraphe 4, dudit règlement. En outre, les fournisseurs de TGPL et de TGMRL seront également considérés comme des déployeurs d'un système d'IA au titre du règlement sur l'IA s'ils utilisent un système d'IA sous leur autorité à leurs propres fins professionnelles (par exemple, en créant des visuels à utiliser dans le cadre de supports de communication marketing).

(127) Si l'hypertrucage comprend des informations relatives à une personne physique (vivante) directement ou indirectement identifiable, les déployeurs, en tant que responsables du traitement au titre du droit de l'Union en matière de protection des données, devront également se conformer aux obligations correspondantes découlant du droit de l'Union en la matière ([37]). Ces obligations comprennent, par exemple, le recours à une base juridique appropriée pour le traitement des données à caractère personnel, la communication d'informations concernant le traitement des données relatives aux personnes concernées représentées, etc.

(128) Les déployeurs de systèmes d'IA générant des hypertrucages devront en outre veiller à se conformer au droit pertinent de l'Union en matière de propriété intellectuelle (tel que le droit des marques de l'Union et la législation de l'Union relative au droit d'auteur ou aux droits voisins) si leurs hypertrucages intègrent des éléments protégés au titre de ces actes législatifs. L'apposition d'une étiquette ou d'une icône sur un hypertrucage n'a, en tant que telle, aucune incidence sur son éligibilité à la protection du droit d'auteur, qui devrait être évaluée conformément à la législation applicable de l'Union en matière de droit d'auteur.

(129) Lorsque des personnes existantes sont représentées dans des hypertrucages, les déployeurs devront également respecter les régimes juridiques pertinents protégeant les droits de la personnalité sur les images ou les voix personnelles. L'obligation de transparence prévue à l'article 50, paragraphe 4, premier alinéa, du règlement sur l'IA ne signifie pas que les hypertrucages générés ou manipulés par une IA qui sont préjudiciables et illégaux en vertu du droit de l'Union ou du droit national applicable (par exemple, en matière de publicité trompeuse ou de droit pénal, tel que le matériel relatif à des abus sexuels sur enfants et les images intimes non consenties) puissent être générés et diffusés (voir également la section 4.4).

6.2. Principaux composants, concepts et obligations de transparence s'y rapportant pour les textes générés ou manipulés par une IA au titre de l'article 50, paragraphe 4, du règlement sur l'IA

(130) L'article 50, paragraphe 4, deuxième alinéa, du règlement sur l'IA s'applique si plusieurs conditions sont remplies:

i) le système doit être un système d'IA;

ii) utilisé pour générer ou manipuler des textes (voir la section 4.2.1.) publiés dans le but d'informer le public sur des questions d'intérêt public;

iii) l'utilisation du système d'IA ne relève d'aucune des exceptions prévues par le règlement, à savoir 1) lorsque le contenu généré par l'IA a fait l'objet d'un processus d'examen humain ou de contrôle éditorial et lorsqu'une personne physique ou morale assume la responsabilité éditoriale de la publication du contenu ou 2) lorsque l'utilisation est autorisée par la loi à des fins de prévention ou de détection des infractions pénales, d'enquêtes ou de poursuites en la matière.

([36]) Considérant 2 du DSA.

([37]) Pour des explications plus précises, voir les lignes directrices conjointes de la Commission et du comité européen de la protection des données sur l'interaction entre le règlement sur l'IA et la législation de l'UE en matière de protection des données (en cours d'élaboration).

6.2.1. *Texte publié dans le but d'informer le public sur des questions d'intérêt public*

(131) Afin de clarifier le champ d'application de l'article 50, paragraphe 4, deuxième alinéa, du règlement sur l'IA, trois éléments sont examinés ci-dessous: i) un texte publié ii) dans le but d'informer le public iii) sur des questions d'intérêt public.

i) **Texte publié:** le texte généré ou manipulé par une IA devrait être publié. Cela signifie que le texte devrait être accessible simultanément et/ou successivement à un nombre indéterminé et assez important de lecteurs potentiels sans lien particulier entre eux, que ce soit ou non contre paiement (par exemple, abonnements) ([38]). En revanche, le texte n'est pas considéré publié si l'accès est limité à des personnes spécifiques appartenant à un groupe privé fermé (par exemple, un petit groupe fermé sur une application de messagerie instantanée, ou si le groupe est trop petit ou insignifiant). Parmi les exemples de textes non considérés comme publiés figurent notamment la correspondance privée et interpersonnelle (à des fins professionnelles) ou les textes ou communications internes à une organisation (par exemple, les publications sur les réseaux internes d'une entreprise).

ii) **Dans le but d'informer le public:** le texte publié généré ou manipulé par une IA devrait viser à informer le public. Cela signifie que le texte devrait avoir pour objectif de communiquer des connaissances, des opinions ou des faits. En revanche, les textes courts qui ne communiquent pas des connaissances, des opinions ou des faits de manière substantielle ne sauraient être considérés comme informant le public.

iii) **Sur des questions d'intérêt public:** le texte devrait informer le public «sur des questions d'intérêt public». En général, ces questions devraient s'entendre comme celles qui présentent un intérêt pour la société dans son ensemble, que ce soit au niveau local, national, de l'Union ou international, et qui justifient un débat public ou un contrôle public. À cet égard, les textes devraient être considérés comme traitant de questions d'intérêt public s'ils portent sur des sujets relatifs à la politique et aux processus démocratiques, à l'administration et aux services publics, à l'administration de la justice et à l'application des lois, à la protection des droits fondamentaux, à la sécurité publique, à la santé publique, à la protection de l'environnement, à la sécurité des consommateurs et à tout développement économique, financier, politique, scientifique ou culturel susceptible de constituer un sujet pertinent de débat public. Les questions qui peuvent être considérées d'intérêt public peuvent évoluer au fil du temps et d'un contexte à l'autre.

Exemples de textes publiés dans le but d'informer le public sur des questions d'intérêt public au sens de l'article 50, paragraphe 4, du règlement sur l'IA, qui doivent être étiquetés comme étant générés ou manipulés par une IA, sauf s'ils font l'objet d'un examen humain/d'un contrôle éditorial (voir la section 6.2.3):

— résumé généré par une IA d'un article rédigé par un auteur humain, publié sur le site web d'un journal et portant sur une décision récemment prise par un conseil municipal;

— parties manipulées par une IA d'un article de site web consacré à l'art de vivre comparant les effets de différents régimes alimentaires sur une maladie donnée chez les femmes d'âge moyen;

— rapports d'entreprise manipulés par une IA publiés sur le site web d'une société cotée en bourse contenant des informations destinées aux investisseurs;

— message généré par une IA publié sur le profil d'un institut météorologique sur les réseaux sociaux alertant les citoyens de conditions météorologiques sévères et les informant des mesures de précaution correspondantes.

Exemples de textes qui ne sont pas publiés dans le but d'informer le public sur des questions d'intérêt public ne relevant pas du champ d'application de l'article 50, paragraphe 4, du règlement sur l'IA:

— romans fantasy générés par une IA;

— texte manipulé par une IA faisant partie de la publicité d'une entreprise ou des descriptions de ses produits (à l'exclusion de toute allégation liée, par exemple, à la santé, à la sécurité des consommateurs ou à la durabilité);

— résumé de l'actualité généré par un dialogueur qui n'est accessible qu'à l'utilisateur l'ayant interrogé;

— texte manipulé par une IA sur l'invite d'un consultant pour conseiller un client sur les mesures à prendre pour assurer le respect de la législation applicable.

([38]) Voir, par analogie, l'arrêt de la Cour (troisième chambre) du 2 juin 2005, *Mediakabel*, C-89/04, ECLI:EU:C:2005:348, point 33; l'arrêt de la Cour (troisième chambre) du 14 juillet 2005, *Lagardère Active Broadcast*, C-192/04, ECLI:EU:C:2005:475, point 31; l'arrêt de la Cour (troisième chambre) du 7 décembre 2006, *SGAE*, C-306/05, ECLI:EU:C:2006:764, points 37 à 40; l'arrêt de la Cour (deuxième chambre) du 26 avril 2017, *Stichting Brein*, C-527/15, ECLI:EU:C:2017:300, points 43 à 44 et jurisprudence citée; l'arrêt de la Cour (première chambre) du 20 juin 2024, *GEMA*, C-135/23, ECLI:EU:C:2024:526, point 38 et jurisprudence citée.

6.2.2. *L'obligation d'information prévue à l'article 50, paragraphe 4, deuxième alinéa, du règlement sur l'IA*

(132) L'article 50, paragraphe 4, deuxième alinéa, du règlement sur l'IA impose aux déployeurs d'un système d'IA qui génère ou manipule des textes publiés dans le but d'informer le public sur des questions d'intérêt public d'indiquer que le texte a été généré ou manipulé par une IA. Le considérant 134 du règlement sur l'IA précise que, à l'instar des déployeurs d'hypertrucages, les déployeurs de publications textuelles relevant du champ d'application de la disposition devraient déclarer de manière claire et reconnaissable que ce contenu a été créé ou manipulé par une IA en étiquetant les sorties d'IA en conséquence et en mentionnant son origine artificielle. Comme cela est requis pour les hypertrucages, les méthodes d'étiquetage ou d'information (y compris les avertissements) appliquées conformément à l'article 50, paragraphe 4, deuxième alinéa, du règlement sur l'IA devraient également être claires et perceptibles par les personnes physiques (par exemple, mesures visibles ou audibles) sans qu'elles aient besoin d'utiliser des outils techniques particuliers ou d'effectuer des actions spécifiques. À la lumière de l'article 50, paragraphe 7, du règlement sur l'IA, les déployeurs peuvent adhérer à un code de bonne pratique, jugé approprié en vertu de cette disposition, afin de démontrer le respect de cette obligation d'information. Pour de plus amples informations concernant les effets de l'application d'un tel code et la surveillance des déployeurs qui n'adhèrent pas à un tel code, voir la section 8.1 ci-dessous.

6.2.3. *Exception à l'obligation de transparence pour les textes faisant l'objet d'un examen humain ou d'un contrôle éditorial et dont la publication relève de la responsabilité éditoriale d'une personne physique ou morale*

(133) L'article 50, paragraphe 4, deuxième alinéa, du règlement sur l'IA prévoit une exception à l'obligation de transparence énoncée dans ladite disposition lorsque deux conditions cumulatives sont remplies: i) le texte généré ou manipulé par une IA doit avoir fait l'objet d'un examen humain ou d'un contrôle éditorial et ii) une personne physique ou morale doit assumer la responsabilité éditoriale de la publication. Le cas échéant, les déployeurs peuvent s'appuyer sur des normes professionnelles ou déontologiques applicables pertinentes pour démontrer le respect de ces exigences, comme expliqué ci-dessous.

i. Texte faisant l'objet d'un examen humain ou d'un contrôle éditorial

(134) La première condition permettant d'invoquer l'exception à l'article 50, paragraphe 4, deuxième alinéa, du règlement sur l'IA exige que le texte généré ou manipulé par une IA ait fait l'objet d'un examen humain ou d'un contrôle éditorial. L'examen humain désigne l'examen délibéré du fond du contenu par une ou plusieurs personnes physiques possédant les connaissances pertinentes et portant un jugement professionnel relatif au sujet examiné (par exemple, évaluation par des pairs dans le milieu universitaire ou chaînes de validation par des professionnels). La vérification de l'exactitude du contenu est une exigence minimale qui devrait faire partie de cet examen. Le contrôle éditorial désigne le contrôle sur le contenu exercé, dans la pratique, par une entité éditoriale responsable (par exemple, un rédacteur en chef) ayant le pouvoir d'approuver, de modifier ou de rejeter le contenu du texte pour des raisons de fond (y compris la vérification des informations et la garantie de la fiabilité des sources). Cette condition est sans préjudice des procédures éditoriales et d'examen existantes ainsi que des normes professionnelles applicables aux fournisseurs de services de médias.

(135) Les contrôles superficiels, uniquement formels ou procéduraux (par exemple, vérification orthographique ou correction grammaticale), la simple existence d'une politique éditoriale, de processus d'examen automatisés ou d'une approbation éditoriale sommaire sans intervention approfondie de la part de l'examinateur humain ou de l'entité éditoriale, ne peuvent remplir les conditions d'un examen humain ou d'un contrôle éditorial aux fins de la présente exception.

(136) Lorsque des systèmes d'IA sont utilisés pour modifier, compléter ou reformuler du contenu à la suite d'une validation éditoriale, le contenu qui en résulte doit être considéré comme généré ou manipulé par une IA aux fins de l'article 50, paragraphe 4, du règlement sur l'IA. Toute intervention substantielle de l'IA après le processus d'examen humain ou de contrôle éditorial rendra donc l'exception caduque.

(137) À la lumière de l'article 50, paragraphe 7, du règlement sur l'IA, les déployeurs peuvent adhérer à un code de bonne pratique, jugé approprié en vertu de cette disposition, afin de démontrer le respect de la condition relative à l'examen humain ou au contrôle éditorial (en particulier par les déployeurs qui ne sont pas des fournisseurs de services de médias soumis aux normes éditoriales professionnelles ou déontologiques existantes) prévue dans ladite disposition. Pour de plus amples informations concernant les effets de l'application d'un tel code et la surveillance des déployeurs qui n'adhèrent pas à un tel code, voir la section 8.1.

ii. *Responsabilité éditoriale*

(138) La deuxième condition permettant d'invoquer l'exception prévue à l'article 50, paragraphe 4, deuxième alinéa, du règlement sur l'IA requiert qu'une personne physique ou morale assume la responsabilité éditoriale de la publication du contenu. Cela signifie que ladite personne doit être juridiquement responsable en dernier ressort de la publication du contenu, y compris de l'examen humain ou du contrôle éditorial (par exemple, une personne, un comité de rédaction ou la société d'édition). Afin de garantir le respect de l'obligation de rendre des comptes au public et la confiance de ce dernier, et conformément aux normes professionnelles existantes en matière de médias, l'identité et les coordonnées de la personne morale, de la personne physique ou de la fonction assumant la responsabilité éditoriale devraient être mises à la disposition du public dans un endroit facilement repérable (si ces informations ne sont pas encore disponibles par d'autres moyens). Ces renseignements peuvent figurer en ligne, par exemple dans les conditions générales d'un site web ou dans d'autres informations juridiques destinées aux utilisateurs. Hors ligne, ils peuvent être inclus dans le colophon ou les mentions d'édition d'une publication.

Exemples de textes faisant l'objet d'un examen humain ou d'un contrôle éditorial avec une personne physique ou morale assumant la responsabilité éditoriale:

— article manipulé par une IA dans un journal ou résumé généré par une IA d'un article qui a fait l'objet du contrôle éditorial du rédacteur en chef concerné, sous la responsabilité éditoriale de la personne morale qui publie le journal;

— blog universitaire manipulé par une IA qui a fait l'objet d'un examen interne par des pairs et dont le centre de recherche concerné assurant la gestion du blog assume la responsabilité éditoriale;

— avertissements de sécurité publique générés par une IA approuvés par un agent public avant d'être distribués aux citoyens, sous la responsabilité de l'agence publique compétente en matière de protection civile;

— rapports sur la durabilité générés par une IA et publiés sur le site web d'une société cotée en bourse, après avoir fait l'objet d'un examen humain par des professionnels exerçant des fonctions pertinentes (par exemple, dans le domaine de la conformité);

— traduction assistée par une IA d'un article écrit par un être humain dans lequel la traduction a fait l'objet d'un examen humain.

Exemples de textes qui ne remplissent pas la condition de l'examen humain ou du contrôle éditorial indispensable pour bénéficier de l'exception:

— site web sur lequel sont publiés des articles générés par une IA sur la politique de l'Union, sans examen humain ni contrôle éditorial délibérés;

— articles générés par une IA qui sont examinés et révisés par un autre système d'IA et dans lesquels un éditeur humain effectue une simple vérification grammaticale superficielle avant publication;

— livre autopublié généré par une IA sur le changement climatique, mis à disposition sur une plateforme de commerce électronique, qui n'a fait l'objet d'aucun examen par une personne physique ou morale compétente (ni par la plateforme).

6.2.4. *Exception à des fins répressives*

(139) Si l'utilisation de systèmes d'IA pour la production de textes publiés dans le but d'informer le public sur des questions d'intérêt public sans examen humain (par exemple, un avertissement de police généré par une IA concernant une menace terroriste publié sur les médias sociaux) est autorisée par la loi à des fins de détection et de prévention des infractions pénales, d'enquêtes et de poursuites en la matière, les déployeurs sont exemptés de l'obligation de transparence prévue à l'article 50, paragraphe 4, du règlement sur l'IA. Les précisions fournies aux points 46 à 48 ci-dessus concernant une exception similaire prévue à l'article 50, paragraphe 1, du règlement sur l'IA sont également pertinentes pour cette disposition.

6.2.5. *Interaction avec d'autres actes juridiques de l'Union*

(140) Pour le droit de l'Union en matière de médias, la responsabilité éditoriale est une notion essentielle définie à l'article 2, point 8), du règlement (UE) 2024/1083 (règlement européen sur la liberté des médias) ([39]). Dans ce contexte, on entend par «responsabilité éditoriale» «l'exercice d'un contrôle effectif tant sur la sélection des programmes ou des publications de presse que sur leur organisation, aux fins de la fourniture d'un service de médias, indépendamment de l'existence d'une responsabilité en vertu du droit national à l'égard du service fourni». Dans certaines situations, cette définition peut se recouper avec celle de la notion de responsabilité éditoriale employée à l'article 50, paragraphe 4, deuxième alinéa, du règlement sur l'IA, en particulier lorsque les déployeurs du système d'IA sont également considérés comme des fournisseurs de services de médias au sens du règlement européen sur la liberté des médias. La responsabilité éditoriale au sens de l'article 50, paragraphe 4, du règlement sur l'IA devrait donc être interprétée conformément à l'acquis existant en matière de médias, nonobstant le fait qu'il s'agit toujours d'une notion distincte qui peut également s'appliquer dans des contextes plus larges et à d'autres déployeurs, pour autant que ceux-ci assument la responsabilité éditoriale de la publication du texte. Les fournisseurs de services de médias peuvent s'appuyer sur leurs processus et normes éditoriaux existants pour bénéficier de l'exception prévue à l'article 50, paragraphe 4, deuxième alinéa, du règlement sur l'IA.

7. **EXIGENCES HORIZONTALES APPLICABLES AUX INFORMATIONS FOURNIES EN VERTU DE L'ARTICLE 50, PARAGRAPHE 5, DU RÈGLEMENT SUR L'IA**

(141) Conformément à l'article 50, paragraphe 5, du règlement sur l'IA, les informations visées à l'article 50, paragraphes 1 à 4, dudit règlement sont fournies aux personnes physiques concernées de manière claire et reconnaissable au plus tard au moment de la première interaction ou de la première exposition. Les informations sont conformes aux exigences applicables en matière d'accessibilité. Les personnes physiques concernées dans ce contexte comprennent les personnes qui interagissent directement avec le système d'IA (article 50, paragraphe 1, du règlement sur l'IA), les utilisateurs actifs ou passifs et les autres personnes exposées au contenu de synthèse généré ou manipulé par une IA [avec ou sans interaction directe avec le système d'IA (article 50, paragraphes 2 et 4, du règlement sur l'IA)], ainsi que les personnes soumises à l'utilisation de systèmes de reconnaissance des émotions ou de catégorisation biométrique (article 50, paragraphe 3, du règlement sur l'IA).

7.1. **Fourniture d'informations de manière claire et reconnaissable**

(142) Les informations seront considérées comme fournies de manière claire lorsqu'elles sont visibles, faciles à comprendre par les personnes physiques concernées (y compris les personnes handicapées) et accessibles à celles-ci. Les informations seront considérées comme fournies de manière reconnaissable lorsqu'il est facile de les distinguer des autres informations et de les repérer dans l'environnement dans lequel le contenu est présenté. Elles doivent également être facilement comprises par le public, y compris par des groupes spécifiques, tels que les enfants ou les personnes handicapées lorsqu'ils font partie du public prévisible susceptible d'être exposé au contenu ou d'interagir avec le système ([40]). Pour qu'elles soient fournies de manière claire et reconnaissable, les informations pertinentes peuvent notamment être présentées dans le cadre de l'interaction (au titre de l'article 50, paragraphe 1, du règlement sur l'IA) ou en accompagnement du contenu concerné (au titre de l'article 50, paragraphes 2 et 4, dudit règlement). Les informations ne seront pas considérées comme fournies de manière claire et reconnaissable lorsqu'elles peuvent facilement passer inaperçues ou ne pas être remarquées par des personnes physiques dans des conditions normales d'exposition ou d'interaction (par exemple, lorsqu'elles figurent uniquement dans un manuel, sont masquées sous plusieurs niveaux d'options de menu sur une interface en ligne, ou font partie des conditions d'utilisation qui ne sont souvent pas lues par les utilisateurs).

([39]) Règlement (UE) 2024/1083 du Parlement européen et du Conseil du 11 avril 2024 établissant un cadre commun pour les services de médias dans le marché intérieur et modifiant la directive 2010/13/UE (règlement européen sur la liberté des médias) (JO L, 2024/1083, 17.4.2024, ELI: http://data.europa.eu/eli/reg/2024/1083/oj).

([40]) Lorsque des enfants font partie du public cible du contenu, ces notifications doivent être: i) adaptées aux enfants, adaptés à l'âge, faciles à comprendre et facilement accessibles à tous les enfants, y compris ceux présentant un handicap et/ou ayant des besoins supplémentaires en matière d'accessibilité; ii) présentées clairement d'une manière facile à comprendre, aussi simples et succinctes que possible; iii) présentées d'une manière facile à examiner et qui offre un accès immédiat et intuitif, dans les circonstances où elles sont pertinentes; iv) présentées dans la ou les langues officielles de l'État membre dans lequel le service est fourni; v) attrayantes pour les enfants, ce qui peut également inclure l'utilisation de graphiques, de vidéos et/ou de personnages ou d'autres techniques; vi) fournies progressivement aux enfants et sur une longue période afin de leur permettre de les retenir le mieux possible.

7.2. **Première interaction ou première exposition**

(143) L'article 50, paragraphe 5, du règlement sur l'IA exige que les informations soient fournies aux personnes physiques concernées au plus tard au moment de la première interaction ou de la première exposition. La première interaction ou la première exposition désigne non seulement la première personne physique qui interagit avec un système d'IA ou qui y est exposée (ou qui est exposée à sa sortie), mais aussi toute première interaction ultérieure avec le système d'IA ou toute première exposition ultérieure à celui-ci (ou à sa sortie) par toute autre personne physique. En ce qui concerne la première interaction avec un système d'IA, les informations devraient être fournies au moins une fois au début d'une session interactive, tout en tenant compte des contextes spécifiques mentionnés aux points 36 et 40 ci-dessus. En ce qui concerne les contenus générés ou manipulés par une IA relevant du champ d'application de l'article 50, paragraphes 2 et 4, du règlement sur l'IA, l'obligation d'information s'applique à chaque sortie d'un système d'IA à l'égard de toute personne physique exposée au contenu. Cela doit s'entendre comme tout moment où une personne physique est raisonnablement susceptible d'être exposée à la sortie du système d'IA et de prendre connaissance de l'information, en tenant compte des spécificités de la communication de cette information dans le cas d'œuvres manifestement créatives, artistiques ou autres (voir la section 6.1.3.). En outre, les déployeurs sont autorisés à informer les personnes physiques avant la première interaction ou la première exposition réelles (par exemple, au début d'un contenu présentant des hypertrucages). Toutefois, s'il est raisonnablement prévisible que des personnes puissent ne pas percevoir le contenu depuis le début, une indication au début du contenu ne suffit pas à informer correctement ces personnes et devrait être complétée par une communication de l'information à des moments ultérieurs, dans la mesure du possible. Cela peut être mis en œuvre le cas échéant, en tenant compte de l'objectif des obligations de transparence consistant à garantir une communication efficace des informations et une prise de conscience réelle des utilisateurs dans des environnements d'IA toujours plus immersifs et réalistes où les utilisateurs éprouvent de plus en plus de difficultés à déterminer l'origine et la véracité du contenu et de l'interaction.

Exemples de fourniture d'informations lors de la première interaction ou de la première exposition au titre de l'article 50, paragraphe 5, du règlement sur l'IA:

— première interaction: lors du lancement d'une conversation avec un dialogueur ou du début d'une interaction physique avec un système d'IA;

— première exposition: au début d'une vidéo comportant un hypertrucage, au début d'une publication textuelle générée ou manipulée par une IA sur des questions d'intérêt public, ou lorsque la personne tombe sur des hypertrucages manipulés par une IA en faisant défiler des contenus sur les médias sociaux.

7.3. **Conformité avec les exigences applicables en matière d'accessibilité**

(144) Les fournisseurs et les déployeurs de systèmes d'IA relevant du champ d'application des différentes obligations de transparence énumérées à l'article 50, paragraphes 1 à 4, du règlement sur l'IA doivent avoir connaissance des exigences applicables en matière d'accessibilité [par exemple, en vertu de la directive (UE) 2016/2102 ([41]) et de la directive (UE) 2019/882 ([42])] et évaluer si leur produit, leur service ou leur contenu doit être conforme à ces règles. Le cas échéant, ils doivent garantir l'accessibilité des informations transmises au titre de l'article 50 du règlement sur l'IA, conformément aux exigences applicables en matière d'accessibilité. L'article 50 du règlement sur l'IA n'impose pas d'exigences distinctes ou supplémentaires en matière d'accessibilité.

8. **APPLICATION DE L'ARTICLE 50 DU RÈGLEMENT SUR L'IA**

(145) Cette section vise à clarifier: i) les effets de l'adhésion à un code de bonne pratique en ce qui concerne le respect des obligations de transparence pour les contenus générés par une IA; ii) le rôle des autorités de contrôle compétentes dans l'application de l'article 50 du règlement sur l'IA; iii) les sanctions applicables en cas de violation de ces obligations; et iv) le moment où les obligations prévues à l'article 50 du règlement sur l'IA entrent en application.

([41]) Directive (UE) 2016/2102 du Parlement européen et du Conseil du 26 octobre 2016 relative à l'accessibilité des sites internet et des applications mobiles des organisations du secteur public (Texte présentant de l'intérêt pour l'EEE) (JO L 327 du 2.12.2016, p. 1).

([42]) Directive (UE) 2019/882 du Parlement européen et du Conseil du 17 avril 2019 relative aux exigences en matière d'accessibilité applicables aux produits et services (Texte présentant de l'intérêt pour l'EEE), PE/81/2018/REV/1 (JO L 151 du 7.6.2019, p. 70).

8.1. **Effets de l'adhésion à un code de bonne pratique jugé approprié**

(146) Les fournisseurs et les déployeurs de systèmes d'IA relevant du champ d'application de l'article 50, paragraphes 2 et 4, du règlement sur l'IA peuvent démontrer qu'ils respectent leurs obligations de transparence pour les contenus générés par une IA au titre de ces dispositions en adhérant à un code de bonne pratique qui est jugé approprié conformément à l'article 50, paragraphe 7, du règlement sur l'IA ([43]). Ce code ne remplace pas le règlement sur l'IA ou les présentes lignes directrices. Il les complète en fournissant le seul cadre pratique reconnu à l'échelle de l'Union permettant aux fournisseurs et aux déployeurs de systèmes d'IA générative de démontrer le respect de ces obligations, indépendamment de leur lieu d'établissement, de leur lieu d'exploitation et des autorités de surveillance du marché compétentes.

(147) Si les fournisseurs et les déployeurs peuvent également démontrer le respect de ces obligations par d'autres moyens adéquats, l'adhésion à un code de bonne pratique jugé approprié est un moyen simple, prévisible et juridiquement sûr de démontrer la conformité à ces obligations. Pour les signataires d'un code de bonne pratique jugé approprié conformément à l'article 50, paragraphe 7, du règlement sur l'IA ([44]), la Commission et les autorités de surveillance du marché compétentes axeront leurs activités de surveillance sur l'évaluation du respect, par ces signataires, du code de bonne pratique et de la mise en œuvre des mesures qu'il contient. Étant donné que les fournisseurs et les déployeurs adhérant à un code seront transparents quant aux mesures qu'ils mettent en œuvre pour se conformer à l'article 50, paragraphes 2, 4 et 5, du règlement sur l'IA, ils bénéficieront d'une confiance accrue de la Commission, des autres autorités de surveillance du marché compétentes et d'autres parties prenantes (y compris du grand public). Toute décision, par les fournisseurs et les déployeurs signataires d'un code de bonne pratique jugé approprié, de ne pas appliquer certaines sections du code aura pour conséquence qu'ils perdront le bénéfice de pouvoir démontrer plus facilement leur conformité à cet égard.

(148) Les fournisseurs et les déployeurs qui ne sont pas signataires d'un code de bonne pratique jugé approprié conformément à l'article 50, paragraphe 7, du règlement sur l'IA sont censés démontrer comment ils ont respecté leurs obligations au titre de l'article 50, paragraphes 2, 4 et 5, dudit règlement par d'autres moyens adéquats. En outre, ces fournisseurs et ces déployeurs doivent expliquer comment les mesures qu'ils mettent en œuvre garantissent le respect des obligations qui leur incombent en vertu du règlement sur l'IA. Par exemple, ils devraient procéder à une analyse des lacunes qui compare les mesures qu'ils ont mises en œuvre avec les mesures prévues par un code de bonne pratique jugé approprié. Étant donné que les autorités compétentes (y compris le Bureau de l'IA) auront une moins bonne compréhension de la manière dont les fournisseurs et les déployeurs qui ne sont pas signataires d'un code de bonne pratique garantissent le respect de leurs obligations au titre de l'article 50, paragraphes 2, 4 et 5, du règlement sur l'IA, elles auront probablement besoin d'informations plus détaillées lors du contrôle de la conformité. Par exemple, les fournisseurs qui n'adhèrent pas à un code de bonne pratique peuvent également faire l'objet d'un plus grand nombre de demandes de renseignements et d'accès afin d'évaluer l'efficacité, l'interopérabilité, la solidité et la fiabilité des solutions techniques mises en œuvre pour se conformer à l'article 50, paragraphe 2, du règlement sur l'IA. Les déployeurs peuvent également faire l'objet de telles demandes en ce qui concerne leurs pratiques d'étiquetage au titre de l'article 50, paragraphe 4, dudit règlement.

(149) Les autorités compétentes peuvent tenir compte des engagements mis en œuvre conformément à un code de bonne pratique jugé approprié en tant que circonstance atténuante lors de la fixation du montant des amendes, en fonction des circonstances spécifiques ([45]).

(150) Si un code de bonne pratique n'est pas jugé approprié conformément à l'article 50, paragraphe 7, du règlement sur l'IA, la Commission peut adopter un acte d'exécution précisant des règles communes pour la mise en œuvre des obligations prévues à l'article 50, paragraphes 2, 4 et 5, dudit règlement, qui seraient applicables à tous les fournisseurs et déployeurs concernés de systèmes d'IA relevant du champ d'application de ces dispositions ([46]).

([43]) Voir les informations pertinentes sur le code de bonnes pratiques sur la transparence des contenus générés par l'IA |Bâtir l'avenir numérique de l'Europe publié le 10 juin 2026.

([44]) Voir l'avis de la Commission sur l'évaluation du code de bonnes pratiques pour la transparence des contenus générés par l'IA |Bâtir l'avenir numérique de l'Europe.

([45]) Article 99, paragraphe 7, point e), du règlement sur l'IA.

([46]) Article 50, paragraphe 7, du règlement sur l'IA.

8.2. **Autorités de surveillance du marché**

(151) Les autorités de surveillance du marché désignées par les États membres, le Bureau de l'IA ([47]) et le Contrôleur européen de la protection des données ([48]) sont chargées de superviser et de faire respecter les règles applicables aux systèmes d'IA relevant de leur compétence, y compris les obligations de transparence énoncées à l'article 50 du règlement sur l'IA. Le contrôle du respect de ces règles a lieu dans le cadre du système de surveillance du marché et de conformité des produits établi par le règlement (UE) 2019/1020 et le règlement sur l'IA. Les pouvoirs de contrôle des autorités de surveillance du marché en ce qui concerne les systèmes d'IA sont définis dans lesdits règlements. Ces autorités peuvent prendre des mesures d'exécution concernant les obligations énumérées à l'article 50 du règlement sur l'IA de leur propre initiative ou à la suite d'une plainte que toute personne concernée ou toute autre personne physique ou morale ayant des motifs de considérer qu'il y a eu violation a le droit d'introduire ([49]). Les responsabilités des autorités de surveillance du marché au titre du règlement sur l'IA n'ont pas d'incidence sur les pouvoirs et les missions des autres autorités de contrôle ou autorités chargées des droits fondamentaux en vertu de leur propre cadre juridique dans les situations où l'article 50 du règlement sur l'IA et ledit cadre sont tous deux applicables (par exemple, en matière de protection des données ou de protection des consommateurs).

8.3. **Sanctions**

(152) Les fournisseurs et les déployeurs qui ne respectent pas les obligations de transparence applicables prévues à l'article 50 du règlement sur l'IA peuvent se voir infliger une amende pouvant aller jusqu'à 15 000 000 EUR ou, si l'auteur de l'infraction est une entreprise, jusqu'à 3 % de son chiffre d'affaires annuel mondial total réalisé au cours de l'exercice précédent, le montant le plus élevé étant retenu. Les institutions, organes et agences de l'Union qui n'observent pas les obligations de transparence peuvent faire l'objet d'amendes administratives pouvant aller jusqu'à 750 000 EUR. Dans le cas des petites et moyennes entreprises (PME), y compris les jeunes pousses, chaque amende s'élève au maximum aux pourcentages ou montant susmentionnés, le chiffre le plus faible étant retenu. Les amendes doivent être effectives, proportionnées et dissuasives et tiennent compte des intérêts des PME et des petites entreprises à moyenne capitalisation, y compris les jeunes pousses, ainsi que de leur viabilité économique. Pour décider s'il y a lieu d'imposer une amende administrative et pour décider du montant de l'amende administrative dans chaque cas d'espèce, les autorités compétentes devraient évaluer toutes les caractéristiques propres à chaque cas et, le cas échéant, une série de facteurs qui influent sur la gravité de la violation (par exemple, la nature, la gravité et la durée de la violation et de ses conséquences, le fait que la violation a été commise délibérément ou par négligence, toute autre circonstance aggravante ou atténuante, le degré de coopération établi avec les autorités de surveillance du marché, etc.).

8.4. **Entrée en application**

(153) Conformément à l'article 113 du règlement sur l'IA, l'article 50 dudit règlement s'applique à partir du 2 août 2026. Cela impose que tous les systèmes d'IA relevant du champ d'application mis sur le marché ou mis en service dans l'Union soient conformes à cette disposition à cette date, quelle que soit leur date de mise sur le marché ou de mise en service ([50]). Le règlement modifiant le règlement sur l'IA (train de mesures omnibus numérique sur l'IA), qui a été récemment adopté par le législateur de l'Union, prévoit une règle ciblée de maintien des droits acquis uniquement en ce qui concerne les obligations de marquage et de détection prévues à l'article 50, paragraphe 2, du règlement sur l'IA pour les systèmes d'IA générative mis sur le marché ou mis en service avant le 2 août 2026. Cette règle accorde aux fournisseurs de ces systèmes existants une période transitoire pour mettre leurs systèmes en conformité au plus tard le 2 décembre 2026. Les systèmes d'IA partiellement interactive et partiellement générative ne peuvent bénéficier de cette période transitoire qu'en ce qui concerne l'obligation de marquage prévue à l'article 50, paragraphe 2, du règlement sur l'IA, tandis que le respect de l'obligation d'information pour les systèmes d'IA qui interagissent directement avec des personnes physiques doit être assuré à partir du 2 août 2026.

([47]) En tant qu'autorité de surveillance du marché pour les systèmes d'IA fondés sur des modèles d'IA à usage général fournis par le même fournisseur (article 75, paragraphe 1, du règlement sur l'IA). Cela couvre l'obligation applicable aux fournisseurs en vertu de l'article 50, paragraphes 1 et 2, du règlement sur l'IA. En ce qui concerne l'article 50, paragraphes 3 et 4, dudit règlement, le Bureau de l'IA ne sera compétent que si le fournisseur de modèle d'IA à usage général et de système d'IA concerné est également un déployeur du système.

([48]) En tant qu'autorité de surveillance du marché pour les institutions, agences et organes de l'Union.

([49]) Article 85 du règlement sur l'IA.

([50]) La règle spéciale de maintien des droits acquis prévue à l'article 111, paragraphe 2, du règlement sur l'IA s'applique aux systèmes d'IA à haut risque mis sur le marché avant la date d'application des règles relatives à ces systèmes uniquement en ce qui concerne leur conformité aux exigences et obligations applicables aux systèmes d'IA à haut risque. Si le système d'IA à haut risque est également soumis à une ou plusieurs obligations de transparence au titre de l'article 50 du règlement sur l'IA, cette règle spéciale de maintien des droits acquis ne s'applique pas, étant donné que l'application des différentes obligations pour les systèmes à haut risque et les obligations en matière de transparence au titre de l'article 50 sont cumulatives et que le maintien des droits acquis est justifié et limité uniquement en ce qui concerne le respect des dispositions relatives aux systèmes à haut risque.

(154) Les sorties générées ou manipulées par une IA relevant du champ d'application de l'article 50, paragraphe 2, du règlement sur l'IA et les hypertrucages relevant du champ d'application de l'article 50, paragraphe 4, premier alinéa, dudit règlement, qui ont été générés ou manipulés avant le 2 août 2026, ne doivent pas faire l'objet d'un marquage ou d'un étiquetage rétroactif. Il en va de même pour les textes portant sur des questions d'intérêt public relevant du champ d'application de l'article 50, paragraphe 4, deuxième alinéa, du règlement sur l'IA qui ont été générés ou manipulés par une IA et publiés avant le 2 août 2026. À l'inverse, si des textes qui ont été générés ou manipulés avant le 2 août 2026 sont publiés à cette date ou après celle-ci, ils doivent être étiquetés. Compte tenu de l'objectif de transparence et d'amélioration de l'intégrité de l'écosystème informationnel et de la confiance en celui-ci poursuivi par le règlement sur l'IA, les déployeurs de systèmes d'IA et les autres acteurs qui détiennent ou diffusent des hypertrucages préexistants non étiquetés sont néanmoins encouragés à les étiqueter sans devoir déployer des efforts disproportionnés à cette fin (tels que l'audit de bases de données de contenu préexistantes ou la modification d'emballages de produits déjà imprimés).

9. **RÉVISION ET MISE À JOUR DES LIGNES DIRECTRICES**

(155) Les présentes lignes directrices constituent une première interprétation des obligations de transparence énoncées à l'article 50 du règlement sur l'IA, assortie d'exemples pratiques. La Commission réexaminera les présentes lignes directrices dès que nécessaire à la lumière de l'expérience pratique acquise au cours de la mise en œuvre des obligations de transparence et en fonction du rythme des évolutions technologiques, sociétales et réglementaires dans ce domaine. Cela inclut également toute expérience pertinente tirée des mesures d'exécution en matière de surveillance du marché et des interprétations de l'article 50 du règlement sur l'IA données par la CJUE. À l'issue de ce réexamen, la Commission pourrait décider de retirer ou de modifier les présentes lignes directrices. Elle encourage les fournisseurs et les déployeurs de systèmes d'IA, les autorités nationales de surveillance du marché par l'intermédiaire du Comité IA, le forum consultatif sur l'IA, la communauté des chercheurs et les organisations de la société civile à contribuer à ce processus en répondant aux futurs appels à consultation publique.