NEWS FLASH

2024 / N°4

Le Règlement (UE) 2024/1689 sur l'intelligence artificielle (AI Act)[1] vient d'être publié :

Quels sont les aspects principaux à retenir ?

À l'issue d'un long processus législatif européen et d'intenses négociations entre les États membres, ***le Règlement européen sur l'intelligence artificielle*** *a été publié le 12 juillet 2024 au Journal Officiel de l'Union européenne (JOUE) et entrera prochainement en vigueur. Le Règlement vise à améliorer le fonctionnement du marché intérieur de l'Union européenne (UE) en établissant des règles harmonisées pour le développement, la mise sur le marché, la mise en service et l'utilisation des systèmes d'intelligence artificielle (IA) dans le respect des valeurs européennes, ainsi qu'à promouvoir l'adoption de l'IA axée sur le facteur humain et digne de confiance, tout en garantissant un niveau élevé de protection de la santé, de la sécurité et des droits fondamentaux.*

Pour rappel, le 21 avril 2021, la Commission européenne rendait publique sa proposition initiale de règlementation sur l'intelligence artificielle visant à encadrer le développement de l'IA tout en favorisant l'innovation technologique. Elle a été suivie par une orientation générale du Conseil de l'UE du 6 décembre 2022 et par des amendements importants du texte par le Parlement européen en date du 14 juin 2023, notamment pour intégrer des règles relatives aux systèmes d'intelligence artificielle générative, telles que ChatGPT. Les États membres sont parvenus à un accord politique sur le texte final du Règlement lors de trilogues interinstitutionnels du 8 décembre 2023. La version finale du texte a ensuite été votée par le Parlement européen lors de sa session plénière du 13 mars 2024 et approuvée par le Conseil de l'UE le 21 mai 2024.

1 *Règlement (UE) 2024/1689 du Parlement européen et du Conseil du 13 juin 2024 établissant des règles harmonisées concernant l'intelligence artificielle et modifiant les règlements (CE) n° 300/2008, (UE) n° 167/2013, (UE) n° 168/2013, (UE) 2018/858, (UE) 2018/1139 et (UE) 2019/2144 et les directives 2014/90/UE, (UE) 2016/797 et (UE) 2020/1828 (règlement sur l'intelligence artificielle)*

1. Qui est visé par le Règlement ?

Le Règlement adopte une approche horizontale en ce qu'il a vocation à s'appliquer indépendamment du secteur d'activité ou du type d'intelligence artificielle.

Ses dispositions sont applicables à tous les acteurs de la chaîne de valeur de l'IA, à savoir les fournisseurs, les déployeurs, les fabricants de produits, les importateurs, les distributeurs, les utilisateurs et autres personnes concernées[2].

Le Règlement s'applique en outre aux fournisseurs[3] et déployeurs[4] tant lorsqu'ils sont établis ou situés dans l'UE que si leur lieu d'établissement est situé dans un pays tiers, mais que les systèmes d'IA sont mis sur le marché dans l'UE ou les résultats générés par les systèmes d'IA sont utilisés dans l'UE.

2. Une définition large des systèmes d'IA

L'AI Act retient une définition large et technologiquement neutre du système d'intelligence artificielle qui ne fait référence ni à la notion de produit ni à celle de service numérique.

Le **système d'intelligence artificielle** est défini comme *«un système automatisé conçu pour fonctionner à différents niveaux d'autonomie et peut faire preuve d'une capacité d'adaptation après son déploiement, et qui, pour des objectifs explicites ou implicites, déduit, à partir des entrées qu'il reçoit, la manière de générer des sorties telles que des prédictions, du contenu, des recommandations ou des décisions qui peuvent influencer les environnements physiques ou virtuels»*[5].

Le choix d'une définition large s'explique par la volonté des institutions européennes de privilégier une approche de règlementation basée sur les risques que représentent les systèmes d'IA.

2 Article 2 du Règlement

3 Article 3 (3) du Règlement définit le fournisseur comme une personne physique ou morale, une autorité publique, une agence ou tout autre organisme qui développe ou fait développer un système d'IA ou un modèle d'IA à usage général et le met sur le marché ou met le système d'IA en service sous son propre nom ou sa propre marque, à titre onéreux ou gratuit.

4 Article 3 (4) du Règlement définit le déployeur comme une personne physique ou morale, une autorité publique, une agence ou un autre organisme utilisant sous sa propre autorité un système d'IA sauf lorsque ce système est utilisé dans le cadre d'une activité personnelle à caractère non professionnel.

5 Article 3 (1) du Règlement

3. La classification des systèmes d'IA sur base des risques

Le Règlement adopte une classification des systèmes d'IA en quatre catégories selon les risques que leurs résultats ou leur utilisation représentent pour les droits fondamentaux. L'intensité des obligations applicables à chaque catégorie d'IA dépend de leur niveau de risques. Il est recommandé aux opérateurs concernés de réaliser une cartographie précise des risques liés à leurs IA afin de pouvoir se conformer aux obligations adéquates.

- **Systèmes représentant des risques inacceptables :** pratiques interdites[6]

Les systèmes d'IA représentant des risques inacceptables sont ceux qui vont à l'encontre des droits fondamentaux et des valeurs de l'UE et ils sont de ce fait interdits.

Parmi les applications à risques inacceptables, on retrouve le *scoring* social, les pratiques manipulatoires sous le seuil de la conscience, l'exploitation de vulnérabilités, les pratiques tendant à inférer les émotions d'individus sur le lieu de travail ou d'éducation, les pratiques tendant à créer ou étendre des bases de données pour la reconnaissance faciale sur la base du moissonnage non ciblé d'images faciales, la reconnaissance faciale biométrique à distance dans les espaces accessibles au public (avec certaines exceptions) etc.

- **Systèmes à haut risque :** des règles strictes à respecter[7]

Dans cette catégorie sont classés les systèmes d'IA à haut risque qui touchent des domaines sensibles comme la santé, la justice, l'éducation, le maintien de l'ordre, les infrastructures critiques, les services publics et privés essentiels.

Un système d'IA est considéré à haut risque si :

(i) Les deux conditions suivantes sont remplies :

- Le système d'IA est destiné à être utilisé comme composant de sécurité d'un produit couvert par les actes législatifs d'harmonisation de l'Union énumérés à l'annexe I, ou le système d'IA constitue lui-même un produit ;
- Le produit dont le composant de sécurité mentionné au point précédent est le système d'IA, ou le système d'IA lui-même en tant que produit, est soumis à une évaluation de conformité par un tiers en vue de la mise sur le marché ou de la mise en service de ce produit conformément aux actes législatifs d'harmonisation de l'Union énumérés à l'annexe I,

OU

(ii) Fait partie des systèmes d'IA visés à l'annexe III du Règlement[8].

La Commission européenne peut adopter des actes délégués afin de modifier et/ou d'ajouter des cas d'utilisation de systèmes d'IA à haut risque à la liste figurant à l'annexe III.

Les systèmes d'IA à haut risque sont soumis à des exigences rigoureuses et doivent se conformer à des normes strictes de gestion des risques, de gouvernance des données et de contrôle humain, afin de garantir leur sécurité et leur fiabilité.

6 Article 5 du Règlement comprend une liste des pratiques d'IA interdites qui représentent des risques inacceptables de dommages pour les individus et pour la société.

7 Article 6 du Règlement

8 L'annexe III du Règlement répertorie des systèmes d'IA considérés comme étant à haut risque.

Ainsi, pour pouvoir être mis sur le marché ou en service, les systèmes d'IA à haut risque doivent satisfaire à des **exigences renforcées** telles que l'établissement d'un système de gestion des risques, d'une documentation technique, de gouvernance de données d'apprentissage (p.ex. la mise en place de mesure pour détecter, prévenir et atténuer d'éventuels biais susceptibles de porter atteinte à la santé et la sécurité des personnes, d'avoir une incidence négative sur les droits fondamentaux ou d'aboutir à une discrimination interdite), d'une transparence à l'égard des déployeurs, d'obligations en matières de robustesse et de cybersécurité, de même que la mise en place d'un contrôle humain effectif proportionné[9].

Ces exigences sont complétées par des **obligations à charge des différentes catégories d'opérateurs économiques** impliqués dans la chaîne de valeur de l'IA, telles que la mise en place d'un système de gestion de la qualité, la conservation de la documentation afférentes à l'IA pendant une période déterminée, la tenue de journaux générés automatiquement par les systèmes d'IA, une coopération avec les autorités compétentes, la prise de mesures techniques et organisationnelles appropriées, la réalisation d'une analyse d'impact des systèmes d'IA à haut risque sur les droits fondamentaux avant leur mise en service[10].

- **Systèmes à faible risque :** règles de transparence spécifiques[11]

Cette catégorie englobe des technologies telles que les chatbots, la reconnaissance d'émotions, les *deep fakes*, les IA Génératives comme ChatGPT, Bard, Dall-e, Midjourney etc.

Il s'agit de systèmes d'IA destinés à interagir directement avec des personnes physiques ou qui peuvent générer des contenus de synthèse (de type audio, image, vidéo ou texte), ou qui peuvent générer ou manipuler de tels contenus constituant des hypertrucages.

Les fournisseurs et les déployeurs de tels systèmes d'IA à faible risque sont soumis à des règles de transparence spécifiques envers les utilisateurs qui interagissent avec l'IA ou qui sont les destinataires des contenus générés par l'IA. L'accent est mis sur l'information des utilisateurs dans leur interaction avec les applications visées, avec l'objectif de prévenir toute manipulation involontaire.

Les contenus générés par un système d'IA (audio, image, vidéo ou texte) doivent être marqués dans un format lisible par machine et détectables comme générées ou manipulées artificiellement. Les IA génératives sont ainsi assorties d'obligations techniques de marquer d'un filigrane les contenus générés par l'IA (davantage destinées aux diffuseurs des contenus visés) et des obligations d'information sur la génération artificielle du contenu lorsque celui-ci s'analyse soit en un hyper-trucage ou *deep fake*, soit lorsqu'il s'analyse en un texte d'information sur des question d'intérêt public (à destination du public).

- **Systèmes à risque minimal :** ne sont pas spécifiquement règlementés

Les systèmes d'IA présentant des risques mineurs ne sont pas soumis à des règles spécifiques par l'AI Act. Toutefois, d'autres règlementations spécifiques peuvent leur être applicable (en matière de propriété intellectuelle, protection des données à caractère personnel etc.).

On retrouve dans cette catégorie des applications telles que les jeux vidéo ou les filtres à spam.

---

9 Chapitre III, Section 2 (articles 8 à 15) du Règlement
10 Chapitre III, Section 3 (articles 16 à 27) du Règlement
11 Article 50 du Règlement

4. Une régulation à part des modèles d'IA à usage général (*General purpose AI* ou GPAI)

Les modèles d'IA à usage général sont traités à part des systèmes d'IA dans le Règlement[12].

Le modèle d'IA à usage général est défini comme *«un modèle d'IA, y compris lorsque ce modèle d'IA est entraîné à l'aide d'un grand nombre de données utilisant l'auto-supervision à grande échelle, qui présente une généralité significative et est capable d'exécuter de manière compétente un large éventail de tâches distinctes, indépendamment de la manière dont le modèle est mis sur le marché, et qui peut être intégré dans une variété de systèmes ou d'applications en aval, à l'exception des modèles d'IA utilisés pour des activité de recherche, de développement ou de prototypage avant leur mise sur le marché»*[13].

En outre, **un modèle d'IA à usage général** pourra être classé comme **présentant un risque systémique**, et être soumis à des obligations spécifiques renforcées, s'il remplit l'un des critères suivants[14]:

(i) il dispose de capacités à fort impact évaluées sur la base de méthodologie et d'outils techniques appropriés, y compris des indicateurs et des critères de référence;

(ii) sur la base d'une décision de la Commission européenne, il possède des capacités ou un impact équivalents.

Les fournisseurs de modèle d'IA à usage général à risque systémique doivent informer la Commission européenne lorsque ces critères sont remplis dans un délai de 2 semaines après la date à laquelle ce critère est rempli ou après qu'il a été établi qu'il le sera. La Commission européenne doit veiller à ce qu'une liste des modèles à risque systémique soit publiée et tenue à jour[15].

Tous les fournisseurs de GPAI doivent rédiger une documentation technique, y compris le processus d'entraînement et d'essai et les résultats de l'évaluation du modèle, élaborer des informations et de la documentation à l'attention des fournisseurs qui ont l'intention d'intégrer le modèle GPAI dans leur propre système d'IA, établir une politique visant à se conformer au droit de l'UE en matière de droit d'auteur et droits voisins, publier un résumé suffisamment détaillé du contenu utilisé pour la formation du modèle GPAI[16].

Les modèles GPAI à licence libre et ouverte (dont les paramètres sont accessibles au public) ne doivent se conformer qu'aux deux dernières obligations susmentionnées, sauf s'il s'agit d'un GPAI à risque systémique[17].

Les fournisseurs de GPAI présentant des risques systémiques doivent **en plus** effectuer des évaluations de modèles, y compris mener et documenter des tests contradictoires afin d'identifier et d'atténuer le risque systémique; repérer, documenter et signaler les incidents graves et les éventuelles mesures correctives au Bureau de l'IA[18] et, le cas échéant, aux autorités nationales compétentes dans les meilleurs délais; et assurer un niveau adéquat de protection de cybersécurité[19].

LES CODES DE BONNES PRATIQUES[20]

Les fournisseurs de GPAI peuvent démontrer le respect de leurs obligations par l'adhésion volontaire à un code de bonne pratique dont le respect entraînera une présomption de conformité du modèle d'IA.

Les Codes de bonnes pratiques tiendront compte des approches internationales et couvriront notamment les obligations applicables aux fournisseurs de GPAI. Le Bureau de l'IA peut inviter les fournisseurs de GPAI et les autorités nationales compétentes à participer à l'élaboration des codes, tandis que la société civile, l'industrie et le monde universitaire, les fournisseurs en aval et les experts indépendants peuvent soutenir le processus.

Les Codes de bonnes pratiques doivent être élaborés au plus tard le 2 mai 2025.

---

12 Chapitre V (articles 51 à 56) du Règlement
13 Article 3 (63) du Règlement
14 Article 51 du Règlement
15 Article 52 du Règlement
16 Article 53 (1) du Règlement
17 Article 53 (2) du règlement
18 Le Bureau de l'IA est présenté au point 6 ci-dessous relatif aux instances de gouvernance
19 Article 55 du Règlement
20 Article 56 du Règlement

5. Les bacs à sable règlementaires ou l'environnement de test supervisé pour les entreprises[21]

Afin de soutenir l'innovation, le Règlement prévoit la mise en place de bacs à sable règlementaires de l'IA par les autorités nationales compétentes. Leur but sera d'offrir un environnement contrôlé pour favoriser l'innovation et faciliter le développement, l'entraînement, la mise à l'essai et la validation de systèmes d'IA innovants pendant une durée avant leur mise sur le marché ou mise en service ainsi que pour permettre des essais en conditions réelles supervisées.

6. Les instances de gouvernance

Le Règlement prévoit la création de plusieurs organismes de gouvernance de niveau européen.

- Le **Bureau de l'IA** est créé au sein de la Commission européenne[22],[23]. Le Bureau fonctionnera en tant que centre pour contrôler, superviser et appliquer les exigences du Règlement concernant les modèles et systèmes d'IA à usage général dans les États membres. Il jouera un rôle clé dans la mise en œuvre du Règlement en soutenant les organes de gouvernance étatiques dans leurs tâches. Il soutiendra la création de bacs à sable règlementaires où les entreprises pourront tester les systèmes d'IA dans un environnement contrôlé. Il fournira également des informations et des ressources aux PME pour les aider à se conformer aux règles.

- Le **Comité européen de l'intelligence artificielle (Comité IA)**[24] sera composé d'un représentant par État membre, doté de droit de vote, et de plusieurs observateurs sans droit de vote (ex. le Contrôleur européen de la protection des données[25], le Bureau de l'IA, autres autorités, organes ou experts nationaux). Son rôle sera d'assurer la cohérence et la coordination de la mise en œuvre du Règlement entre les autorités nationales compétentes.

- Le **Forum consultatif**[26] sera chargé de fournir une expertise technique et de conseiller le Comité IA et la Commission européenne. Il sera composé de représentants de divers parties prenantes (l'industrie, les jeunes pousses, les PME, la société civile et le monde universitaire). En outre, l'Agence des droits fondamentaux, l'ENISA, le Comité européen de normalisation (CEN), le Comité européen de normalisation électrotechnique (CENELEC) et l'Institut européen de normalisation des télécommunications (ETSI) seront des membres permanent du Forum.

- Le **Groupe scientifique d'experts indépendants**[27] doit être mis en place par la Commission européenne au moyen d'un acte d'exécution. Il aura pour rôle de soutenir les activités de contrôle de l'application du Règlement.

Les organismes de gouvernance de niveau européen sont complétés par des **autorités nationales compétentes**[28] (au moins une autorité notifiante et une autorité de surveillance du marché au sens du Règlement) qui vont également avoir vocation à mettre en œuvre le Règlement.

21 Articles 57 et 58 du Règlement

22 Article 64 du Règlement

23 Une décision de la Commission européenne en date du 24 janvier 2024, entrée en vigueur le 21 février 2024, a crée cette instance qui fera partie de la Direction générale des réseaux de communication, du contenu et de la technologie (DG CNECT) de la Commission européenne.

24 Articles 65 et 66 du Règlement

25 Lien vers le site internet du Contrôleur européen de la protection des données (CEPD)

26 Article 67 du Règlement

27 Article 68 du Règlement

28 Article 70 du Règlement

7. Les sanctions

Les États membres pourront déterminer le régime de sanctions et d'autres mesures d'exécution applicables aux violations du Règlement, en tenant compte des lignes directrices à publier par la Commission européenne.

Toutefois, certaines sanctions sont d'ores et déjà prévues par le Règlement[29]:

- le non-respect de l'interdiction des pratiques prohibées est susceptible d'amende administrative pouvant aller **jusqu'à 35 millions d'euros** ou **7% du chiffre d'affaires annuel mondial total** de l'entreprise réalisé au cours de l'exercice précédent;
- la non-conformité avec des dispositions spécifiquement énumérées[30] relatives aux opérateurs ou aux organismes notifiés peut être sanctionnée par une amende administrative pouvant aller jusqu'à **15 millions d'euros** ou **3% du chiffre d'affaires annuel mondial total**;
- la fourniture d'information inexactes, incomplètes ou trompeuses aux autorités nationales ou organismes notifiés en réponse à une demande peut faire l'objet d'une amende administrative pouvant aller jusqu'à **7,5 millions d'euros** ou **1% du chiffre d'affaires annuel mondial total**;
- les fournisseurs de modèles d'IA à usage général peuvent se voir imposer des amendes jusqu'à **15 millions d'euros** ou **3% du chiffre d'affaires mondial total**, en cas de violation des dispositions pertinentes du Règlement, s'ils n'ont pas donné suite à une demande de document ou d'information, ou s'ils ont fourni des informations inexactes, incomplètes ou trompeuses, s'ils ne se sont pas conformés à des mesures ordonnées par la Commission, ou s'ils n'ont pas donné accès à la Commission à un GPAI ou un GPAI présentant un risque systémique en vue de son évaluation.

8. L'entrée en vigueur

Le Règlement entrera en vigueur le 1er août 2024 (vingt jours après la publication officielle au JOUE). Il deviendra pleinement applicable à partir du 2 août 2026 (24 mois après la date d'entrée en vigueur), **à l'exception** de certaines dispositions qui deviendront applicables aux échéances suivantes:

- **2 février 2025** (6 mois après l'entrée en vigueur) – les dispositions relatives aux systèmes d'IA interdits;
- **2 mai 2025** – les Codes de bonnes pratiques;
- **2 août 2025** (12 mois après l'entrée en vigueur) – les règles concernant l'IA à usage général (GPAI), la nomination des autorités compétentes nationales;
- **2 août 2027** (36 mois après l'entrée en vigueur) – les dispositions relatives aux systèmes à haut risques.

Auteurs: Legal & Tax, Chambre de Commerce
juridique@cc.lu

29 Articles 99 et 101 du Règlement

30 L'article 99 (4) énumère les dispositions visées dont la violation peut faire l'objet de ces sanctions.