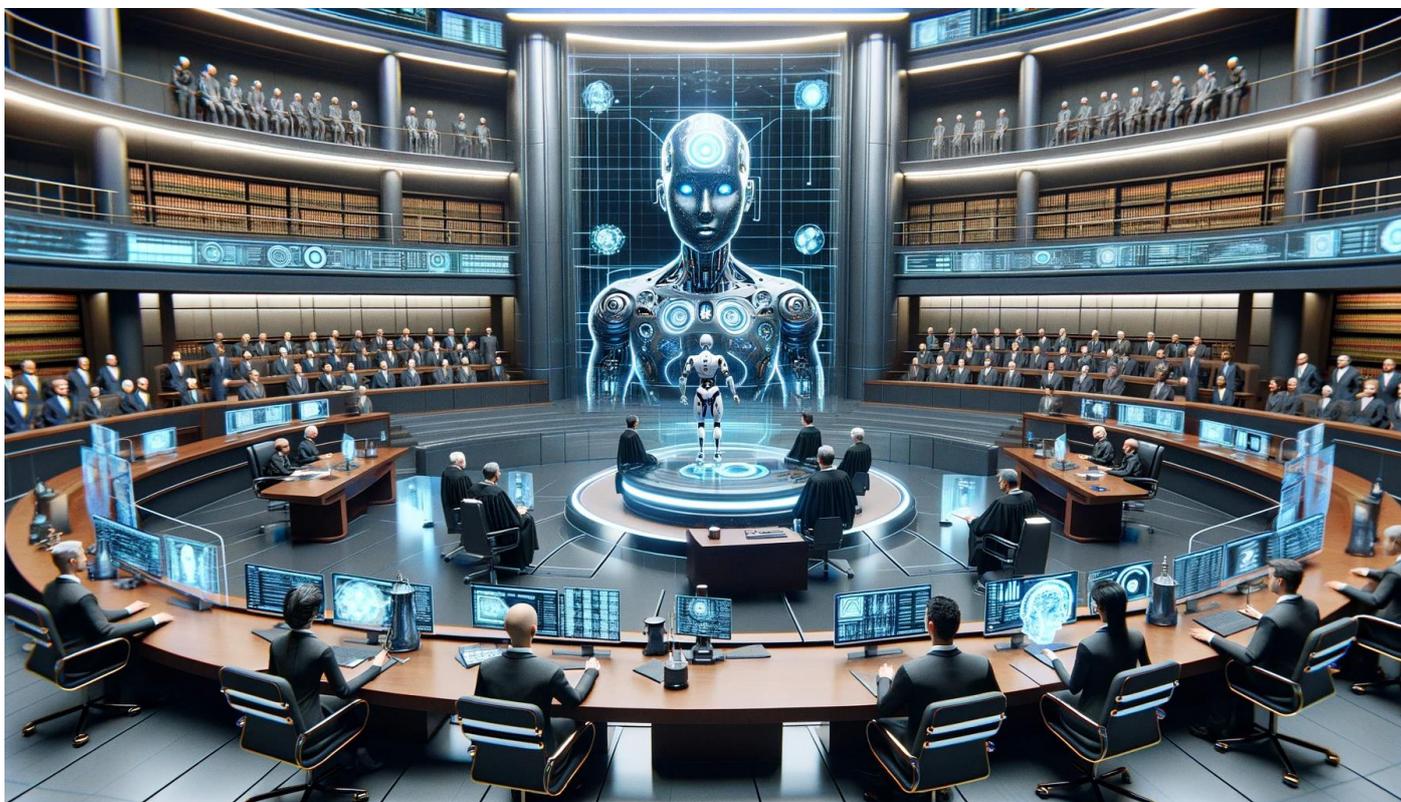


Article 4 du Règlement européen sur l'intelligence artificielle : le concept de « maîtrise de l'IA »

Par [Emmanuel Walle](#) | 05/11/2024



LE CONCEPT DE « MAITRISE DE L'INTELLIGENCE ARTIFICIELLE »

- Dernière mise à jour : **5 novembre 2024**

Mots-clés : [Actualités](#) / [Ethique intelligence artificielle](#) / [Obligation d'information](#) / [Règlement européen](#) / [Intelligence Artificielle](#) / [Responsabilité](#)

Version PDF

1. Introduction à la maîtrise de l'IA

Au cœur de l'innovation technologique, l'intelligence artificielle (IA) s'impose comme un vecteur incontournable de transformation à travers les secteurs et les industries. Avec cette avancée fulgurante, l'Union européenne a pris des mesures proactives pour encadrer son développement et son utilisation responsable. [L'article 4 du règlement UE 2024/1689](#) incarne cette démarche régulatrice en introduisant **la notion de « maîtrise de l'IA »**. Mais que signifie réellement ce concept et quels impacts aura-t-il sur les fournisseurs et déployeurs (utilisateurs) de systèmes d'IA ?

Dans cet article, nous plongeons au cœur des dispositions du règlement UE 2024/1689 pour démêler le fil d'une notion aussi révolutionnaire qu'essentielle : **la maîtrise de l'IA**. Nous explorerons en détail ce que le législateur européen entend par **compétences et connaissances adéquates requises** pour ceux qui sont aux commandes des technologies d'intelligence artificielle.

- Quelles responsabilités pèsent désormais sur les épaules des fournisseurs ?
- Quelles implications pour la formation des opérateurs ?
- Et quelles garanties cela offre-t-il quant à une IA sûre, éthique et sous contrôle humain ?

Le concept de « maîtrise de l'IA » introduit par l'article 4 du règlement UE 2024/1689 recouvre **une obligation légale spécifique**. Ce règlement impose à ceux qui fournissent et déploient des systèmes d'IA l'obligation de s'assurer que les personnes engagées dans l'exploitation et l'utilisation d'un système d'IA possèdent **un niveau de connaissances et de compétences adéquat**.

La maîtrise de l'IA est un pilier essentiel du cadre réglementaire européen, visant à promouvoir **une IA éthique et digne de confiance**. Elle établit les principes et exigences essentiels nécessaires pour garantir **une supervision et une intervention humaines efficaces** au sein des systèmes d'IA, en particulier ceux classés à haut risque.

L'objectif principal de la réglementation dans le cadre de cette définition de la maîtrise de l'IA est de **trouver un équilibre** entre l'exploitation des avantages de l'IA et l'atténuation des risques potentiels et des conséquences imprévues. En imposant une surveillance et un contrôle humains, la réglementation vise à **maintenir l'action et la responsabilité humaines** dans les processus de prise de décision basés sur l'IA, en particulier dans les domaines qui ont un impact significatif sur la vie des individus.

2. Supervision et contrôle humains

Le règlement IA ne fournit pas de liste exhaustive des **différents niveaux de supervision humaine requis**. Il est toutefois prévu que ce concept soit davantage développé dans les actes d'exécution du règlement. On peut anticiper l'existence de différents niveaux de supervision, comme **une surveillance continue, une surveillance périodique ou une surveillance déclenchée par des alertes**.

Les opérateurs en charge de la supervision doivent posséder des **compétences spécifiques** pour comprendre le fonctionnement des systèmes d'IA et identifier d'éventuelles anomalies. La mise en place de **programmes de formation adaptés** est donc essentielle.

Il est important de noter que la supervision humaine ne dégage pas **les fournisseurs et les déployeurs** de leur responsabilité. Ils **restent responsables du bon fonctionnement des systèmes d'IA**, même lorsqu'ils sont sous supervision humaine.

Les fournisseurs et les déployeurs doivent s'assurer que les opérateurs impliqués dans l'exploitation et l'utilisation des systèmes d'IA possèdent un niveau de connaissances et de compétences adéquat. Cela implique de **comprendre les capacités et les limites** des systèmes, **ainsi que les risques potentiels associés** à leur utilisation.

Les opérateurs humains doivent être formés pour **reconnaître et réagir aux biais, erreurs ou conséquences imprévues** qui peuvent résulter des systèmes d'IA. Ils doivent également être capables d'**interpréter les résultats** des systèmes et de **prendre des décisions éclairées** sur la base de ces informations.

Les fournisseurs et les déployeurs sont responsables de fournir **une documentation claire et complète, du matériel de formation et un soutien continu** pour garantir que les opérateurs peuvent superviser et intervenir efficacement dans les systèmes d'IA lorsque cela est nécessaire.

3. Transparence et explicabilité

L'exigence de transparence va au-delà de la simple description des algorithmes utilisés dans les systèmes d'IA. Elle nécessite également **que les décisions prises par ces systèmes soient intelligibles**, en particulier lorsqu'elles ont un impact significatif sur les individus.

Les fournisseurs doivent préparer une **documentation technique complète détaillant le fonctionnement interne de leurs systèmes d'IA**. Cette documentation doit faciliter une compréhension approfondie des systèmes par les opérateurs et les autorités de régulation. Pour parvenir à une véritable transparence, il faut **non seulement divulguer les algorithmes, mais aussi expliquer comment ils parviennent à des résultats ou à des décisions spécifiques**.

En outre, les décisions prises par les systèmes d'IA, en particulier celles qui ont des conséquences importantes pour les individus, doivent **être compréhensibles et explicables**. Cela signifie que le raisonnement qui sous-tend les décisions doit être accessible et compréhensible pour les personnes concernées. Les fournisseurs doivent s'efforcer de

développer des systèmes d'IA capables de **fournir des explications claires et significatives** de leurs résultats, permettant aux utilisateurs de saisir la logique et la justification des décisions.

La transparence et l'explicabilité sont essentielles pour renforcer la confiance dans les systèmes d'IA et garantir leur déploiement responsable. En favorisant la **transparence** ainsi qu'une meilleure **compréhension des processus décisionnels de l'IA**, le règlement vise à donner aux individus, aux organisations et aux organismes de réglementation les moyens de **prendre des décisions éclairées** et d'**exercer une surveillance efficace** sur ces puissantes technologies.

4. Surveillance post-commercialisation

Le règlement sur l'IA ne prescrit pas de méthodes spécifiques pour la surveillance post-commercialisation des systèmes d'IA. Il permet plutôt aux fournisseurs de **choisir la méthode la plus adaptée à leurs systèmes**, à condition qu'elle soit **efficace** et permette **une détection proactive des problèmes**.

Les méthodes potentielles de surveillance des systèmes d'IA incluent :

- Surveillance continue des résultats du système et des indicateurs de performance
- Audits et tests périodiques du système
- Mécanismes de retour d'information des utilisateurs et de signalement des incidents
- Intégration des capacités de surveillance dans la conception du système.

Quelle que soit la méthode choisie, les prestataires doivent veiller à ce qu'elle permette d'**identifier rapidement les risques potentiels, les biais ou les défaillances du système d'IA**. Une détection précoce est essentielle pour atténuer les dommages et résoudre rapidement les problèmes.

En cas d'incident grave ou de violation, les fournisseurs sont tenus d'**informer les autorités compétentes concernées et tous les utilisateurs ou parties intéressés**. Cette obligation de signalement des incidents vise à promouvoir la transparence et la responsabilité, permettant de prendre les mesures appropriées pour résoudre le problème et éviter qu'il ne se reproduise.

La détection proactive des problèmes est un aspect essentiel du suivi post-commercialisation. Les prestataires doivent mettre en œuvre des mécanismes pour identifier les problèmes potentiels avant qu'ils ne dégénèrent en incidents graves. Cela peut impliquer la définition de seuils pour les mesures de performance, la surveillance des anomalies ou des écarts par rapport au comportement attendu et l'examen régulier des résultats du système pour détecter d'éventuels biais ou erreurs.

Dans l'ensemble, les exigences de surveillance post-commercialisation soulignent l'importance d'une **vigilance continue** et d'une **surveillance responsable** des systèmes d'IA, même après leur déploiement initial. Une **surveillance efficace** est essentielle pour garantir la sécurité, la fiabilité et la fiabilité des systèmes d'IA tout au long de leur cycle de vie.

5. Responsabilité et responsabilité

La responsabilité des algorithmes est une question complexe au cœur de nombreux débats autour de la gouvernance de l'IA. Bien que l'article 4 n'apporte pas de solution définitive, il souligne l'importance de **préserver la responsabilité humaine** dans les systèmes d'IA.

L'un des principaux défis consiste à **déterminer les responsabilités** lorsqu'un système d'IA prend une décision qui entraîne des préjudices ou des conséquences néfastes. La nature opaque de nombreux algorithmes d'IA, en particulier ceux qui font appel à l'**apprentissage automatique**, peut rendre difficile l'attribution de la responsabilité à des personnes ou entités spécifiques impliquées dans le développement ou le déploiement du système.

Pour relever ce défi, le règlement souligne la nécessité de **disposer de mesures de documentation et de transparence solides** qui permettent de **comprendre clairement le fonctionnement des systèmes d'IA** et la manière dont ils

prennent des décisions. Cette transparence est essentielle pour identifier les points de défaillance ou les biais potentiels susceptibles d'entraîner des résultats néfastes.

Malgré les complexités techniques, le règlement renforce le principe selon lequel **les êtres humains doivent assumer en dernier ressort la responsabilité des actions et des décisions des systèmes d'IA**. Cette responsabilité s'étend à tous ceux qui participent au développement, au déploiement et à l'exploitation de ces systèmes, notamment les développeurs, les fabricants et les utilisateurs finaux.

Une approche potentielle pour atténuer les préoccupations en matière de responsabilité consiste à introduire des **exigences obligatoires en matière d'assurance responsabilité civile** pour les fournisseurs d'IA. Une telle assurance pourrait couvrir les dommages ou les pertes résultant de l'exploitation de leurs systèmes d'IA, offrant ainsi un filet de sécurité aux personnes ou aux entités concernées. Toutefois, les détails et la portée spécifiques de **ces exigences en matière d'assurance restent à déterminer**.

Il est important de noter que la question de la **responsabilité algorithmique** est un domaine en constante évolution et que le paysage réglementaire est susceptible de s'adapter à mesure que de nouveaux défis et considérations apparaissent. Un dialogue et une collaboration continus entre les décideurs politiques, les acteurs du secteur et les experts juridiques seront essentiels pour trouver le juste équilibre entre la promotion de l'innovation et la garantie d'un développement et d'un déploiement responsables de l'IA.

6. Conception centrée sur l'humain

La conception centrée sur l'humain est un aspect essentiel de la maîtrise de l'IA en vertu de l'article 4 du règlement de l'UE sur l'IA. Il souligne l'importance de prendre en compte les implications éthiques et sociales **dès le début du processus de conception du système d'IA**. Les fournisseurs doivent collaborer étroitement avec les utilisateurs pour s'assurer que les systèmes d'IA répondent à leurs besoins et à leurs attentes.

Adopter une approche centrée sur l'humain nécessite une compréhension approfondie des cas d'utilisation prévus, de l'impact potentiel sur les individus et la société, ainsi que des considérations éthiques à prendre en compte. **Cela implique d'intégrer des perspectives diverses**, notamment celles des utilisateurs finaux, des experts en la matière et des parties prenantes d'horizons divers.

En **impliquant les utilisateurs dans le processus de conception**, les fournisseurs peuvent obtenir des informations précieuses sur **les scénarios réels, les biais potentiels et les conséquences imprévues**. Cette approche collaborative permet d'identifier et d'atténuer les risques potentiels, garantissant ainsi que les systèmes d'IA sont conformes aux principes éthiques et aux valeurs sociétales.

De plus, **la conception centrée sur l'humain favorise la transparence et la confiance** dans les systèmes d'IA. Lorsque les utilisateurs participent activement au processus de conception, ils sont plus susceptibles de comprendre la logique qui sous-tend les décisions du système et d'avoir confiance dans ses capacités et ses limites.

En fin de compte, la conception centrée sur l'humain consiste à créer des systèmes d'IA qui non seulement fonctionnent efficacement, mais **qui respectent également les droits de l'homme**, favorisent l'équité et contribuent au bien-être de la société. En accordant la priorité aux considérations éthiques et sociales dès le départ, les fournisseurs peuvent développer des solutions d'IA dignes de confiance, responsables et véritablement bénéfiques pour l'humanité.

7. Impact sur l'innovation

Le règlement sur la maîtrise de l'IA pourrait avoir **des répercussions importantes sur l'innovation dans le secteur de l'IA**, en particulier pour les petites entreprises. Les **exigences de conformité**, telles que la garantie d'une supervision humaine adéquate, la transparence et la surveillance post-commercialisation, peuvent imposer **des coûts et des charges administratives substantiels** aux entreprises, en particulier celles qui disposent de ressources limitées.

Les petites entreprises et start-ups du secteur de l'IA pourraient avoir du mal à allouer suffisamment de fonds et de personnel pour répondre aux exigences de la réglementation. **L'embauche et la formation de personnel spécialisé** pour superviser les systèmes d'IA, **l'élaboration d'une documentation technique complète et la mise en œuvre de mécanismes de surveillance robustes** pourraient grever leurs budgets et détourner des ressources des activités de recherche et développement.

En outre, l'accent mis par la réglementation sur la conception centrée sur l'humain et les considérations éthiques pourrait ralentir le rythme de l'innovation pour certaines applications de l'IA. **Trouver le juste équilibre entre innovation et déploiement responsable de l'IA** pourrait s'avérer difficile, car les entreprises pourraient devoir consacrer davantage de temps et de ressources à répondre aux préoccupations éthiques et à atténuer les risques potentiels.

Toutefois, la réglementation présente également **des opportunités d'innovation**. Les entreprises qui adoptent de manière proactive les principes de maîtrise de l'IA et accordent la priorité au développement responsable de l'IA pourraient acquérir **un avantage concurrentiel** en instaurant la confiance des consommateurs et des organismes de réglementation. Cela pourrait ouvrir de nouveaux marchés et faciliter l'adoption de leurs solutions d'IA.

En outre, la réglementation pourrait stimuler **l'innovation dans des domaines tels que l'IA explicable, la collaboration homme-IA et les cadres de gouvernance de l'IA**. Les entreprises qui développent des solutions innovantes pour répondre aux exigences de la réglementation, telles que des processus décisionnels transparents ou des outils de surveillance efficaces, pourraient potentiellement créer de nouvelles sources de revenus ou obtenir un avantage de premier plan.

Dans l'ensemble, même si la réglementation sur la maîtrise de l'IA peut poser des problèmes aux petites entreprises à court terme, elle pourrait également **favoriser l'innovation à long terme** en favorisant un écosystème d'IA plus responsable et plus fiable. Les entreprises qui s'adaptent rapidement et tirent parti des principes de la réglementation pourraient se positionner comme des leaders du développement éthique de l'IA.

8. Harmonisation internationale

Il est essentiel de veiller à ce que la réglementation européenne sur l'IA soit conforme aux efforts et aux normes internationales pour assurer sa mise en œuvre efficace et son impact mondial. Le développement de la technologie de l'IA transcende les frontières et **une approche réglementaire harmonisée peut empêcher la fragmentation et promouvoir la collaboration entre les nations**.

Les efforts de coordination à l'échelle mondiale sont essentiels pour **établir des principes et des lignes directrices communs** en matière de gouvernance de l'IA. L'alignement sur des cadres internationaux largement acceptés peut favoriser la confiance, faciliter les flux de données transfrontaliers et permettre le déploiement transparent des systèmes d'IA dans les différentes juridictions. **Des réglementations divergentes pourraient créer des obstacles à l'innovation, entraver le commerce international et compromettre les avantages potentiels de l'IA**.

En outre, **l'harmonisation internationale** peut encourager **le partage des meilleures pratiques, des enseignements tirés et de l'expertise collective** en matière de gouvernance de l'IA. En collaborant avec d'autres pays et organisations, l'Union européenne peut contribuer à façonner **un paysage réglementaire mondial cohérent et homogène pour l'IA**, garantissant ainsi un développement et un déploiement éthiques et fiables de ces technologies dans le monde entier.

Il est essentiel que la réglementation européenne sur l'IA participe activement aux forums internationaux, dialogue avec les parties prenantes concernées et tienne compte des implications mondiales de ses dispositions. **L'alignement sur des normes et des principes** largement reconnus peut non seulement promouvoir une utilisation responsable de l'IA, mais aussi positionner l'Union européenne comme un **leader** dans la définition de l'avenir de la gouvernance de l'IA à l'échelle mondiale.

9. Evolution de la réglementation

Le règlement (UE) 2024/1689 sur l'intelligence artificielle est un texte législatif relativement récent. Comme tout règlement, il devrait faire l'objet de mises à jour et de clarifications au fur et à mesure de sa mise en œuvre et de sa mise en pratique. **Le domaine de l'intelligence artificielle évolue rapidement et la réglementation devra s'adapter pour suivre le rythme des avancées technologiques.**

Au fil du temps, certains aspects de la réglementation pourraient nécessiter d'être peaufinés ou de faire l'objet de directives supplémentaires. Par exemple, **le concept de « maîtrise de l'IA » et les niveaux spécifiques de supervision humaine requis pour différents types de systèmes d'IA pourraient devoir être définis plus clairement.** En outre, à mesure que de nouvelles technologies d'IA émergent, la réglementation pourrait devoir être élargie ou modifiée pour tenir compte de leurs caractéristiques uniques et des risques potentiels qu'elles présentent.

Il est essentiel **que la réglementation reste flexible et adaptée à l'évolution du paysage de l'IA.** Une collaboration continue entre les décideurs politiques, les experts du secteur et les autres parties prenantes sera nécessaire pour garantir que la réglementation reste pertinente et efficace dans la promotion d'un développement de l'IA éthique et digne de confiance.

Il convient de s'attendre à ce **que la réglementation soit régulièrement révisée et mise à jour,** en tenant compte des enseignements tirés de sa mise en œuvre et en comblant les éventuelles lacunes ou incohérences. Ce processus itératif contribuera à **renforcer le cadre réglementaire** et à garantir qu'il continue de remplir sa fonction première, à savoir **favoriser l'innovation** responsable en matière d'IA tout en **préservant les droits et les intérêts des individus et de la société** dans son ensemble.

10. Conclusion

L'exigence de maîtrise de l'IA introduite par l'article 4 du règlement UE 2024/1689 marque une étape importante dans la réglementation de l'intelligence artificielle. Cette disposition établit **un cadre solide pour garantir l'utilisation responsable et éthique des systèmes d'IA,** en particulier ceux classés à haut risque.

L'accent mis par la réglementation sur **la supervision humaine, la transparence, l'explicabilité, la surveillance post-commercialisation, la responsabilité et la conception centrée sur l'humain** sont des étapes cruciales pour **instaurer la confiance dans les technologies de l'IA.** Cependant, la mise en œuvre efficace de ces principes présente plusieurs défis.

L'un des principaux défis consiste à élaborer des **programmes de formation complets** pour doter les opérateurs humains des compétences et des connaissances nécessaires pour superviser et intervenir efficacement dans les systèmes d'IA. En outre, il faudra trouver un équilibre délicat pour **garantir la transparence et l'explicabilité** sans compromettre **les droits de propriété intellectuelle ou les secrets commerciaux.**

Il est également essentiel d'établir **des lignes directrices claires** pour la surveillance post-commercialisation et les mécanismes de signalement des incidents afin d'identifier et de traiter de manière proactive les problèmes potentiels liés aux systèmes d'IA. En outre, pour résoudre **la question complexe de la responsabilité** et de la responsabilité des algorithmes, il faudra une collaboration entre les décideurs politiques, les experts juridiques et les parties prenantes du secteur.

La mise en œuvre réussie du règlement sur la maîtrise de l'IA nécessitera **une étroite collaboration entre les secteurs public et privé.** Les régulateurs doivent travailler en étroite collaboration avec les fournisseurs et les déploieurs d'IA pour garantir une compréhension commune des exigences et favoriser un environnement propice à l'innovation et au développement éthique de l'IA.

En outre, à mesure que le paysage de l'IA continue d'évoluer rapidement, la réglementation elle-même devra peut-être s'adapter et subir des révisions pour répondre aux nouveaux défis et aux avancées technologiques. Un **dialogue continu et le partage des connaissances** entre les parties prenantes internationales seront essentiels pour harmoniser les réglementations de l'IA à l'échelle mondiale.

Le concept de « maîtrise de l'intelligence artificielle »

En conclusion, **le règlement sur la maîtrise de l'IA représente une étape importante** vers l'établissement d'un écosystème d'IA fiable et responsable. Cependant, sa mise en œuvre effective nécessitera **un effort concerté** de toutes les parties prenantes, **une adaptation continue** aux changements technologiques et un engagement à **favoriser l'innovation tout en donnant la priorité aux considérations éthiques et au bien-être de la société.**

Emmanuel Walle

Avocat, Directeur du
département Social
numérique



Iris Defaut

Avocate, collaboratrice du
département Travail
numérique

