

Législation sur l'intelligence artificielle: le Conseil et le Parlement parviennent à un accord sur les premières règles au monde en matière d'IA

À l'issue de trois jours de pourparlers-marathon, la présidence du Conseil et les négociateurs du Parlement européen sont parvenus à un accord provisoire sur la proposition relative à des règles harmonisées concernant l'intelligence artificielle (IA), dénommée la "**législation sur l'intelligence artificielle**". Le projet de règlement a pour objectif de veiller à ce que les systèmes d'IA mis sur le marché européen et utilisés dans l'UE soient **sûrs** et à ce qu'ils respectent les **droits fondamentaux** et les valeurs de l'UE. Cette proposition qui fait date vise également à stimuler l'investissement et l'innovation dans le domaine de l'IA en Europe.

Il s'agit là d'une réalisation historique et d'un pas de géant vers l'avenir! L'accord conclu aujourd'hui répond efficacement à un défi mondial dans un environnement technologique en évolution rapide concernant un domaine clé pour l'avenir de nos sociétés et de nos économies. Et dans cette entreprise, nous sommes parvenus à maintenir un équilibre extrêmement délicat: encourager l'innovation et l'adoption de l'intelligence artificielle dans toute l'Europe tout en respectant pleinement les droits fondamentaux de nos citoyens.

Carme Artigas Brugal, secrétaire d'État espagnole à la numérisation et à l'intelligence artificielle

La législation sur l'IA est une initiative législative **phare**, susceptible de favoriser le développement et l'adoption, par les acteurs tant privés que publics, d'une IA sûre et digne de confiance dans l'ensemble du marché unique de l'UE. L'IA ayant la capacité de porter préjudice à la société, l'idée est principalement de réglementer celle-ci selon une approche **fondée sur les risques**, à savoir: **plus le risque est élevé, plus les règles sont strictes**. Cette proposition législative étant une première en son genre dans le monde, elle peut établir une **norme mondiale** pour la réglementation de l'IA dans d'autres pays ou territoires, comme ce fut le cas pour le règlement général sur la protection des données (RGPD), promouvant ainsi l'approche européenne de la réglementation des technologies sur la scène internationale.

Les principaux éléments de l'accord provisoire

Par rapport à la proposition initiale de la Commission, les principaux éléments nouveaux de l'accord provisoire peuvent être résumés comme suit:

- des règles relatives aux **modèles d'IA à finalité générale et à fort impact** susceptibles de créer un risque systémique à l'avenir ainsi qu'aux **systèmes d'IA** à haut risque
- un système révisé de **gouvernance**, doté de certains pouvoirs d'exécution au niveau de l'UE
- une extension de la liste des **interdictions**, s'accompagnant toutefois de la possibilité, pour les autorités répressives, d'utiliser l'**identification biométrique à distance** dans les espaces publics, sous réserve de garanties
- une meilleure protection des droits grâce à l'obligation pour les déployeurs de systèmes d'IA à haut risque de procéder à une **analyse d'impact sur les droits fondamentaux** avant de mettre un système d'IA en service.

Plus concrètement, l'accord provisoire couvre les aspects suivants:

Définitions et champ d'application

Afin de veiller à ce que la **définition** d'un système d'IA englobe des critères suffisamment clairs pour distinguer l'IA des systèmes logiciels plus simples, l'accord de compromis aligne la définition sur l'approche proposée par l'OCDE.

L'accord provisoire précise également que le règlement ne s'applique pas aux domaines ne relevant pas du champ d'application du droit de l'UE et ne devrait en aucun cas affecter les compétences des États membres en matière de **sécurité nationale** ou toute entité chargée de tâches dans ce domaine. En outre, la législation sur l'IA ne s'appliquera pas aux systèmes utilisés exclusivement à des fins **militaires** ou de **défense**. De même, l'accord prévoit que le règlement ne devrait pas s'appliquer aux

systèmes d'IA utilisés aux seules fins **de la recherche et de l'innovation**, ni aux personnes utilisant l'IA pour des raisons non professionnelles.

Classification de systèmes d'IA comme pratiques d'IA à haut risque et interdites

L'accord de compromis prévoit une **couche horizontale de protection**, incluant une classification à haut risque, afin de ne pas englober les systèmes d'IA qui ne sont pas susceptibles de causer des violations graves des droits fondamentaux ou d'autres risques importants. Les systèmes d'IA ne présentant qu'un **risque limité** seraient soumis à des **obligations de transparence** très peu contraignantes, par exemple celle consistant à faire savoir que le contenu a été généré par l'IA, de façon à ce que les utilisateurs puissent prendre des décisions éclairées sur la poursuite de leur utilisation.

Un large éventail de systèmes d'IA à **haut risque** seraient autorisés, mais seraient soumis à un ensemble d'exigences et d'obligations pour accéder au marché de l'UE. Ces exigences ont été clarifiées et adaptées par les colégislateurs de sorte qu'il soit, pour les parties intéressées, **techniquement plus faisable** et **moins contraignant** de s'y conformer, par exemple en ce qui concerne la qualité des données ou la documentation technique que les PME devraient établir pour démontrer que leurs systèmes d'IA à haut risque sont conformes aux exigences.

Étant donné que les systèmes d'IA sont développés et distribués au moyen de chaînes de valeur complexes, l'accord de compromis comprend des modifications clarifiant la **répartition des responsabilités** et des **rôles des différents acteurs** de ces chaînes, en particulier les fournisseurs et les utilisateurs de systèmes d'IA. Il clarifie également la relation entre responsabilités au titre de la législation sur l'IA et responsabilités qui existent déjà en vertu d'autres actes législatifs, tels que la législation pertinente de l'UE en matière de protection des données ou la législation sectorielle.

Pour certaines utilisations de l'IA, le risque est jugé **inacceptable** et, par conséquent, ces systèmes seront interdits dans l'UE. L'accord provisoire interdit, par exemple, la **manipulation comportementale** cognitive, le **moissonnage** non ciblé d'images faciales provenant de l'internet ou de la vidéosurveillance, la **reconnaissance d'émotions** sur le lieu de travail et dans les établissements d'enseignement, la **notation sociale**, la **catégorisation biométrique** pour en déduire des données sensibles, telles que l'orientation sexuelle ou des croyances religieuses, et certains cas de **police prédictive** appliquée à la personne.

Exceptions en matière d'application de la loi

Compte tenu des spécificités des **autorités répressives** et de la nécessité de préserver leur capacité à utiliser l'IA dans leurs activités essentielles, plusieurs modifications de la proposition de la Commission ont été approuvées en ce qui concerne l'utilisation de systèmes d'IA à des fins répressives. Sous réserve de **garanties** appropriées, ces modifications visent à tenir compte de la nécessité de respecter la confidentialité des données opérationnelles sensibles dans le cadre de leurs activités. Par exemple, une procédure d'urgence a été mise en place pour permettre aux services répressifs compétents de déployer un outil d'IA à haut risque qui n'a pas passé avec succès la procédure d'**évaluation de la conformité** en cas d'urgence. Toutefois, un mécanisme spécifique a également été mis en place pour garantir que les **droits fondamentaux** seront suffisamment protégés contre toute utilisation abusive potentielle des systèmes d'IA.

En outre, en ce qui concerne l'utilisation de systèmes d'**identification biométrique à distance** "en temps réel" dans des espaces accessibles au public, l'accord provisoire clarifie les objectifs pour lesquels une telle utilisation est strictement nécessaire à des fins répressives et pour lesquels les autorités répressives devraient donc être exceptionnellement autorisées à utiliser de tels systèmes. L'accord de compromis prévoit des **garanties supplémentaires** et limite ces exceptions aux cas de victimes de certaines infractions, à la prévention de menaces réelles, actuelles ou prévisibles, telles que les attentats terroristes, et à la recherche de personnes soupçonnées des crimes les plus graves.

Systèmes d'IA à usage général et modèles de référence

De nouvelles dispositions ont été ajoutées pour tenir compte de situations dans lesquelles les systèmes d'IA peuvent être utilisés à de nombreuses fins différentes (**IA à usage général**), et des situations dans lesquelles une technologie d'IA à usage général est ensuite intégrée dans un autre système à haut risque. L'accord provisoire porte également sur les cas spécifiques des systèmes d'IA à usage général.

Des règles spécifiques ont également été adoptées pour les **modèles de référence**, ces grands systèmes capables d'exécuter de manière compétente un large éventail de tâches distinctes, telles que produire des vidéos, des textes, des images, converser dans un langage latéral, calculer ou créer des codes informatiques. L'accord provisoire prévoit que les modèles de référence doivent respecter des **obligations de transparence spécifiques** avant leur mise sur le marché. Un régime plus strict a été mis en place pour les modèles de référence **à fort impact**. Il s'agit des modèles de référence formés avec une grande quantité de données et présentant une complexité, des capacités et des performances avancées nettement supérieures à la moyenne, qui peuvent diffuser des risques systémiques le long de la chaîne de valeur.

Une nouvelle architecture de gouvernance

Compte tenu des nouvelles règles relatives aux modèles d'IA à usage général et de l'évidente nécessité de les appliquer au niveau de l'UE, un **bureau de l'IA** est créé au sein de la Commission pour superviser ces modèles d'IA parmi les plus avancés, contribuer à promouvoir normes et pratiques d'évaluation et faire appliquer les règles communes dans tous les États membres. Un **groupe scientifique d'experts indépendants** conseillera le Bureau de l'IA en ce qui concerne les modèles d'IA à usage général, en contribuant à l'élaboration de méthodologies pour évaluer les capacités des modèles de référence, en donnant des conseils concernant la désignation et l'émergence de modèles de référence à fort impact, et en surveillant les éventuels risques matériels pour la sécurité liés aux modèles de référence.

Le **Comité de l'IA**, qui compterait des représentants des États membres, restera en place pour servir de plateforme de coordination et d'organe consultatif de la Commission et confèrera un rôle important aux États membres pour la mise en œuvre du règlement, y compris l'élaboration de codes de bonnes pratiques pour les modèles de référence. Enfin, un **forum consultatif** destiné aux parties prenantes, telles que les représentants de l'industrie, les PME, les jeunes pousses, la société civile et le monde universitaire, sera mis en place afin de fournir une expertise technique au Comité de l'IA.

Sanctions

Les amendes pour violation de la législation sur l'IA ont été fixées en pourcentage du chiffre d'affaires annuel mondial de l'entreprise incriminée au cours de l'exercice précédent ou à un montant prédéterminé, le montant le plus élevé étant retenu. Elles s'élèveraient à 35 millions d'euros ou à 7 % pour les violations des applications d'IA interdites, à 15 millions d'euros ou à 3 % pour les violations des obligations découlant de la législation sur l'IA et à 7,5 millions d'euros ou à 1,5 % pour la communication d'informations inexactes. Toutefois, l'accord provisoire prévoit des **plafonds plus proportionnés** pour les amendes administratives infligées aux PME et aux jeunes pousses en cas d'infraction aux dispositions de la législation sur l'IA.

L'accord de compromis précise également qu'une personne physique ou morale peut déposer une réclamation pour non-respect de la législation sur l'IA auprès de l'**autorité de surveillance du marché** compétente et peut s'attendre à ce que cette réclamation soit traitée conformément aux procédures spécifiques de cette autorité.

Transparence et protection des droits fondamentaux

L'accord provisoire prévoit une **analyse d'impact sur les droits fondamentaux** avant qu'un système d'IA à haut risque ne soit mis sur le marché par ses déployeurs. L'accord provisoire prévoit également d'accroître la **transparence** en ce qui concerne l'utilisation des systèmes d'IA à haut risque. En particulier, certaines dispositions de la proposition de la Commission ont été modifiées afin de spécifier que certains utilisateurs d'un système d'IA à haut risque qui sont des entités publiques seront également tenus de s'enregistrer dans la **base de données de l'UE** pour les systèmes d'IA à haut risque. De plus, des dispositions nouvellement ajoutées mettent l'accent sur l'obligation faite aux utilisateurs d'un **système de reconnaissance des émotions** d'informer les personnes physiques lorsque celles-ci sont exposées à un tel système.

Mesures de soutien à l'innovation

En vue de créer un cadre juridique plus propice à l'innovation et de promouvoir un apprentissage réglementaire fondé sur des données probantes, les dispositions relatives aux **mesures de soutien à l'innovation** ont été substantiellement modifiées par rapport à la proposition de la Commission.

Ainsi, il a été notamment précisé que les **bacs à sable réglementaires** de l'IA, qui sont censés établir un environnement contrôlé pour le développement, l'essai et la validation de systèmes d'IA innovants, devraient également permettre de tester des systèmes d'IA innovants dans des conditions réelles. En outre, de nouvelles dispositions ont été ajoutées, qui permettent de **tester** des systèmes d'IA **en conditions réelles**, sous certaines conditions et garanties. Afin d'alléger la charge administrative pesant sur les petites entreprises, l'accord provisoire contient une liste d'actions devant être engagées pour soutenir ces opérateurs et prévoit certaines dérogations limitées et clairement définies.

Entrée en vigueur

L'accord provisoire prévoit que la législation sur l'IA devrait s'appliquer **deux ans** après son entrée en vigueur, à quelques exceptions près pour certaines dispositions spécifiques.

Prochaines étapes

À la suite de l'accord provisoire intervenu aujourd'hui, les travaux se poursuivront au niveau technique dans les semaines à venir, afin de mettre au point les détails du nouveau règlement. La présidence présentera le texte de compromis aux représentants des États membres (Coreper) pour approbation une fois ces travaux terminés.

Le texte intégral devra être confirmé par les deux institutions et faire l'objet d'une mise au point par les juristes-linguistes avant son adoption formelle par les colégislateurs.

Informations générales

La proposition de la Commission, présentée en avril 2021, est un élément fondamental de la politique de l'UE visant à favoriser le développement et l'adoption, dans l'ensemble du marché unique, d'une IA sûre et licite qui respecte les droits fondamentaux.

La proposition suit une approche fondée sur les risques et établit pour l'IA un cadre juridique uniforme et horizontal dont le but est de garantir la sécurité juridique. Le projet de règlement vise à promouvoir les investissements et l'innovation dans le domaine de l'IA, à renforcer la gouvernance et l'application effective de la législation existante en matière de droits fondamentaux et de sécurité et à faciliter le développement d'un marché unique pour les applications d'IA. Il va de pair avec d'autres initiatives, notamment le plan coordonné dans le domaine de l'intelligence artificielle, qui vise à accélérer les investissements dans l'IA en Europe. Le 6 décembre 2022, le Conseil est parvenu à un accord en vue d'une orientation générale (mandat de négociation) sur ce dossier et a entamé des pourparlers interinstitutionnels avec le Parlement européen ("trilogues") à la mi-juin 2023.

Press office - General Secretariat of the Council

Rue de la Loi 175 - B-1048 BRUSSELS - Tel.: +32 (0)2 281 6319

press@consilium.europa.eu - www.consilium.europa.eu/press