ARTICLE

ACCORD PROVISOIRE SUR LA RÉGLEMENTATION DE L'IA PAR L'UE : UNE APPROCHE FONDÉE SUR LE RISQUE

Droit européen | 18/12/23 | Mahasti Razavi

PROPRIÉTÉ INTELLECTUELLE

L'Union européenne a franchi une étape cruciale vers la régulation de l'intelligence artificielle (IA) avec un accord provisoire en trilogue sur la proposition de règlement sur l'IA (IA Act). Il reste toutefois de nombreuses réunions techniques à venir pour préciser le texte, dont l'adoption définitive est attendue avant les élections de juin 2024. Le principe reste l'approche fondée sur les risques associés à l'utilisation des systèmes d'IA : plus le risque est élevé, plus les règles sont strictes. Ce texte complexe tente de concilier innovation et droits fondamentaux.

Quatre niveaux de risque seraient retenus :

- Risques inacceptables : interdiction de certains usages.
- Risque élevé : exigences accrues en matière de conformité et de transparence.
- Risque faible : adhésion volontaire à des codes de conduite et aux obligations de transparence.
- Risque de transparence spécifique : exigences de transparence pour les systèmes d'IA susceptibles de manipuler les utilisateurs.

Les modèles d'IA en open-source bénéficieraient d'exemptions mais sous certaines conditions.

Définitions et champ d'application : L'accord adopte une définition de l'IA cohérente avec celle de l'OCDE pour mieux distinguer les systèmes d'IA des autres logiciels. Il précise les responsabilités des différents acteurs (fournisseurs et utilisateurs) et l'articulation avec d'autres législations, dont le RGPD.

Le règlement s'appliquera à tous les acteurs, publics et privés, opérant dans l'UE ou affectant des personnes situées dans l'UE, mais pas aux utilisations privées et non professionnelles.

Les systèmes d'IA à risque inacceptable seraient interdits dans l'UE. Cela comprend : la manipulation cognitive, l'extraction non ciblée d'images faciales, la reconnaissance des émotions sur le lieu de travail et dans l'enseignement, la notation sociale, et la catégorisation biométrique pour déduire des données sensibles.

Les autorités publiques pourraient utiliser l'IA dans un but de sécurité publique sous réserve de garanties appropriées (ex : utilisation de systèmes d'identification biométrique à distance en temps réel dans les espaces accessibles au public).

Évaluation de l'impact et conformité des IA à haut risque : Avant la mise sur le marché par ceux d'un système d'IA à haut risque par ceux qui le déploient, une évaluation de l'impact sur les droits fondamentaux serait requise. En tout état de cause, les fournisseurs devraient soumettre leur système à une évaluation de la conformité démontrant le respect des exigences pour une IA de confiance (qualité des données, documentation et traçabilité, transparence, supervision humaine, précision, cybersécurité et robustesse), et répéter cette évaluation en cas de modifications substantielles. Les importateurs devraient s'assurer que les fournisseurs étrangers ont respecté les procédures d'évaluation de la conformité et disposent du marquage CE.

Modèles d'IA à usage général et modèles de fondation Des dispositions spécifiques ont été ajoutées pour réguler les modèles d'IA à usage général et les modèles dits de fondation. Un "*modèle d'IA à usage général*" est un modèle entraîné à l'aide d'une grande quantité de données en utilisant l'auto-supervision à grande échelle, pour des usages d'une grande généralité et qui est capable d'exécuter avec efficacité un large éventail de tâches distinctes, quelle que soit 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.

A cette fin, les fournisseurs de tels systèmes devraient fournir aux fournisseurs de système en aval les informations nécessaires pour s'assurer que leur système est sûr et conforme à la règlementation.

Un régime plus exigeant a été conçu pour les modèles de fondation "à fort impact", c'est à dire susceptibles d'entraîner, en raison de leurs capacités et performances, des risques systémiques tout au long de la chaîne de valeur. Les fournisseurs de ces modèles seraient tenus d'évaluer ces risques et de les atténuer, de signaler les incidents graves, de réaliser des évaluations de modèles, de garantir la cybersécurité et de fournir des informations sur la consommation d'énergie de leurs modèles.

Les fournisseurs de ces modèles devraient en outre mettre en place des politiques visant à garantir le respect de la législation sur les droits d'auteur lors de la formation de leurs modèles.

En l'état, les modèles d'IA à usage général qui ont été formés en utilisant une puissance de calcul totale de plus de 10^25 FLOPs seraient considérés comme présentant des risques systémiques.

Innovation et régulation : L'IA Act permettrait la création de bacs à sable réglementaires et la possibilité d'essais en conditions réelles mais dans un cadre contrôlé, pour une durée limitée.

Sanctions : Les sanctions pour non-conformité pourraient s'élever jusqu'à 35 millions d'euros ou 7 % du chiffre d'affaires annuel mondial total pour les infractions les plus graves.