Réglementation sur l'Intelligence Artificielle et Justice

Avec l'appui scientifique de Monsieur Bruno Deffains
Professeur à l'Université Paris II Panthéon Assas

Note destinée à la Commission Européenne

Avec la proposition de règlement sur l'intelligence artificielle (IA) présentée le 21 avril 2021, la Commission européenne réaffirme l'importance stratégique de l'IA pour l'Europe et la nécessité d'encadrer son usage dans les différents secteurs d'application. Nous tenons à saluer cette initiative unique dans le monde, qui place l'intelligence artificielle au centre des réflexions juridiques, politiques et sociétales, desquelles elle est désormais indissociable. En investissant le terrain règlementaire, l'Union Européenne entend se positionner en chef de file, comme elle l'a déjà fait avec le RGPD et c'est une excellente nouvelle.

Comme avec le RGPD, l'objectif est d'instaurer un climat de confiance tout en permettant l'innovation. L'approche par les risques est équilibrée (I) pour permettre un juste déploiement des systèmes d'IA en Europe. Néanmoins, en raison de ses enjeux spécifiques et délicats, nous appelons à traiter le secteur de la Justice comme un secteur hautement sensible pour nos concitoyens (II). Nous regrettons que le texte soit trop restrictif et ne comprenne pas suffisamment les enjeux de la justice prédictive avec les garde-fous nécessaires au bon respect des droits fondamentaux des individus.

1. L'approche par les risques : une approche équilibrée

L'approche de la Commission européenne se fonde ainsi avant tout sur une identification et un encadrement des risques selon les domaines d'applications (publics ou privés) concernés. La proposition de la Commission européenne classifie les systèmes d'IA en quatre groupes : a) ceux créant un risque inacceptable (art.5), b) ceux à haut risque (art.6 et s.) et c) ceux avec un risque faible (art. 52), devant répondre à des exigences de transparence (comme les chatbots, où l'on devra savoir que l'on se trouve face à une « IA », ou les deepfakes) ou d) minimum (p.13). Les risques sont ainsi considérés comme inacceptables dans le champ de la sécurité, des moyens de subsistance et des droits des personnes. Les risques sont perçus comme élevés dans les domaines touchant les infrastructures critiques (énergie, transports…), l'éducation, la formation professionnelle, l'emploi, les ressources humaines, les publics essentiels à l'image du maintien de l'ordre, de la justice ou des processus de décision démocratiques. Ainsi, par exemple, les

systèmes d'identification biométrique à distance fondés sur l'IA**,** sont considérés à haut risque et la Commission rappelle l'interdiction de leur utilisation dans l'espace public et en temps réel "aux fins du maintien de l'ordre" en dehors de cas spécifiques encadrés judiciairement. Dans les secteurs où les risques sont limités, la Commission recommande une transparence d'information sur la présence de l'IA. Enfin, dans d'autres domaines d'application comme les jeux vidéo ou les filtres antispam, les risques liés à la présence de l'IA sont considérés comme minimes et ne nécessitent pas d'encadrement spécifique.

Il est important de souligner le tournant stratégique que constitue cette réglementation. Après une longue période de régulation prudente des outils et applications numériques, excepté sous l'angle de la protection des données ou de la lutte contre la cybercriminalité, la Commission entend accélérer le pas. Cette évolution touchant l'IA est l'une des composantes d'une stratégie clairement énoncée dans les lignes politiques 2019-2024 de la Commission et à laquelle vont contribuer d'autres textes parmi lesquels le Digital Market Act, le Digital Service Act, le Data Governance Act ou encore l'Open Data Directive. L'Europe compte à juste titre s'imposer en gardienne des valeurs dans le nouvel environnement numérique et jouer pleinement de son expérience de régulateur pour aligner les opérations de ses concurrents directs (américains et asiatiques) sur son rythme.

2. Le secteur de la Justice hautement sensible pour les citoyens européens : le texte mérite d'être approfondi avec davantage de garanties pour la *« justice predictive »*

Ces ambitions conduisent la Commission à proposer une approche originale en termes de risques en distinguant 4 types d'applications, avec des contraintes à l'intensité décroissant. Ce faisant, elle n'échappe toutefois pas à une certaine complexité pour la mise en conformité des applications présentant le plus de risque de dommage sur les individus. C'est particulièrement vrai dans le domaine de la justice. La classification des systèmes débouchera nécessairement sur des interprétations et des débats, des opérateurs pouvant être tentés d'éviter la régulation contraignante pour des applications à la limite du haut risque (est-ce qu'un potentiel système de gestion des audiences dans les tribunaux représente un « haut risque » par exemple ?). Le prononcé de sanctions au titre des éventuels manquement devra également prendre en compte la concurrence possible entre différents ordres de juridictions (pénales et administratives), problématique déjà connue entre les autorités de protection des données et les juridictions pénales notamment.

Par ailleurs, si l'on peut se réjouir de la mention de certains problèmes sous-estimés (comme les biais d'automatisation – art. 14, 4, b), on comprend que des questions de principe n'ont pas été approfondies, notamment celle de l'opportunité du recours à des algorithmes dans certaines

matières comme la police ou la justice. Même si, dans ce domaine à « haut risque », une procédure d'examen de conformité s'appliquera, l'annexe III autorise de fait le principe d'un appui de systèmes algorithmiques pour la prise de décision judiciaire, le profilage des individus ou le contrôle aux frontières. A titre d'exemple, rien n'est réellement dit sur l'atténuation concrète des biais ou le degré d'explicabilité attendu, notamment dans ces domaines. La conséquence de cette position sera de ne satisfaire personne en définitive : les défenseurs de droits jugeront comme une régression cette légitimation ; les legaltechs et autres opérateurs privés vont se voir imposer une très lourde mise en conformité sans nécessairement répondre au fond des critiques pouvant leur être adressées (comme le sens des prévisions des outils de « *justice prédictive* »). Nombre d'applications risquent donc d'être « blanchies » par un dispositif de mise en conformité. Les enjeux en la matière sont loin d'être négligeables tant on doit mesurer les conséquences que ces applications peuvent avoir sur le fonctionnement à long terme de la justice. La justice prédictive apparaît donc au cœur de nombreux débats qui renvoient à la fois à son potentiel de développement du fait de son appropriation croissante par les praticiens du droit et des inquiétudes suscitées sur un plan ontologique, notamment dans le contexte particulier des systèmes de droit codifiés où « la règle » prime en principe sur « le fait ». Si on peut s'accorder sur le fait que l'analyse quantitative des données jurisprudentielles permet de déterminer des « valeurs » qui doivent être interprétées et replacer dans le contexte particulier de la fabrique du droit qui ne saurait être détachée des caractéristiques inhérentes à l'organisation de la justice dans une démocratie moderne à la fois du point de vue du travail des juges et des conditions de collecte et de traitement des données. Ceci est d'autant plus important que dans la plupart des pays (de droit civil mais aussi de *« common law* »), le travail de traitement des données est largement confié au secteur privé, ce qui pose nécessairement la question de la contribution effective de la justice prédictive à l'amélioration du service public de la justice.

De manière générale, la proposition marque une étape importante mais elle semble encore hésiter entre le trop ou le trop peu de précisions, laissant craindre la mise en place d'un cadre juridique légitimant des applications tout à fait critiquables. Les garde-fous sont nécessaires. Créer de la confiance dans la justice dans une période de transformation numérique massive passe par des recours facilités pour les individus, par des mesures sociétales prenant mieux en compte la dimension collective des enjeux du numérique et en imposant une charge de mise en conformité, fondée sur des preuves scientifiques solides, qui soit une réelle valeur ajoutée qualitative pour les citoyens européens.

Pour toute information complémentaire, veuillez contacter Anne-Charlotte Gros, Directrice Générale de la Fondation pour le droit continental à l'adresse suivante : acgros@fondation-droitcontinental.org