Loi européenne sur l’IA : une réglementation « digne de con?ance » ?

Description

Après d’âpres négociations, la loi sur l’intelligence artificielle (IA) de l’Union européenne (UE) est en passe d’être adoptée, sous réserve de l’accord final du Conseil de l’UE et du Parlement européen dans le courant du printemps 2024.

Cette adoption formelle est l’aboutissement d’un processus laborieux, entamé par la proposition législative de la Commission européenne en 2021, qui faisait elle-même suite aux recommandations d’un groupe européen d’experts de haut niveau sur l’IA créé en 2018.

En plus de définir le premier cadre régional explicitement dédié à l’IA, l’objectif avoué de l’exécutif européen en présentant cette loi était de garantir le développement en Europe d’une IA « *digne de confiance* », respectant les valeurs et règles de l’UE.

Si la question de la réglementation des IA génératives (on pense bien sûr à ChatGPT) a attiré toute l’attention du grand public à partir de 2022, ces négociations ont également donné lieu à des débats sur toute une série de mesures, certes moins visibles, mais tout aussi importantes, pour le futur de l’IA au niveau européen.

Avant de présenter brièvement le contenu de ce texte législatif, puis d’en préciser les implications, revenons sur la généalogie de la première loi européenne sur l’IA pour mieux en appréhender les déterminants et les forces motrices.

Une brève généalogie de la loi sur l’IA

Dès 2018, une série d’acteurs économiques et politiques, menée en premier lieu par l’industrie numérique, a appuyé la constitution d’un groupe européen d’experts de haut niveau, dans le but de proposer des lignes directrices pour le développement d’une « IA éthique ». Composé aux deux tiers de représentants de l’industrie, ce groupe a fait l’objet de nombreuses critiques dont celles de favoriser une forme d’« *éthique-washing* ».

Certaines propositions de ce groupe d’experts ont finalement été reprises dans la proposition législative de la Commission publiée en 2021. Ce texte de loi s’inscrivait alors dans une séquence au cours de laquelle la nouvelle présidente de la Commission Ursula von der Leyen tentait d’imposer un agenda politique visant le renforcement de la « *souveraineté numérique européenne* ». C’est ce *momentum* politique, relativement dirigiste vis-à-vis des politiques numériques de l’UE, qui a conduit à l’ouverture de négociations entre le

Parlement européen et le Conseil autour de ce nouveau cadre européen pour l'IA.

Ces négociations politiques ont été marquées par d'intenses controverses, reflétant à la fois les grands enjeux stratégiques et industriels que revêt le développement de l'IA dans les pays européens, mais également la divergence des visions politiques de ces derniers quant aux potentialités et aux dangers posés par ces systèmes.

Parmi ces controverses, l'une des plus fondamentales concernait la définition de ce que recouvre « IA » dans le texte de loi. Cette notion est connue pour son caractère nébuleux et controversé[1], et les colégislateurs européens sont loin d'avoir échappé à ces difficultés. Tandis que le Parlement européen comme la Commission souhaitaient une approche maximaliste, en définissant l'IA de façon à inclure le plus large spectre d'usages dans le champ de la loi, certains États membres au sein du Conseil de l'UE défendaient à l'inverse une approche plus minimaliste, réduisant l'IA à certaines formes avancées de *machine-learning*.

Plus généralement, le positionnement du Parlement européen[2] favorisait l'extension du champ des interdictions proposées par la Commission, notamment autour de la reconnaissance faciale dans l'espace public, tout en aménageant une certaine flexibilité réglementaire pour les développeurs et utilisateurs de l'IA, sous la pression de la droite européenne. De leur côté, les États membres au sein du Conseil de l'UE[3] ont fait largement bloc pour protéger leurs propres champions industriels nationaux de nouvelles règles qui pourraient freiner leur croissance, tout en s'opposant aussi aux obstacles réglementaires qui auraient pu contrevenir à leur usage de l'IA à des fins sécuritaires.

Alors que ces négociations menaient bon train en 2022, le succès et la fascination suscitée par l'arrivée soudaine de ChatGPT ont progressivement bouleversé ces équilibres politiques, et ajouté l'épineuse problématique de la réglementation des IA génératives à l'agenda des négociateurs.

La résolution de ces controverses aboutira prochaine- ment, avec l'adoption formelle de la loi prévue au printemps 2024. Le texte final, dont la version préliminaire a récemment fuité[4], donne à voir une nouvelle architecture pour la régulation européenne de l'IA, abordée dans la section suivante.

« Réguler les usages de l'IA par les risques » : analyse des fondements réglementaires de la loi sur l'IA

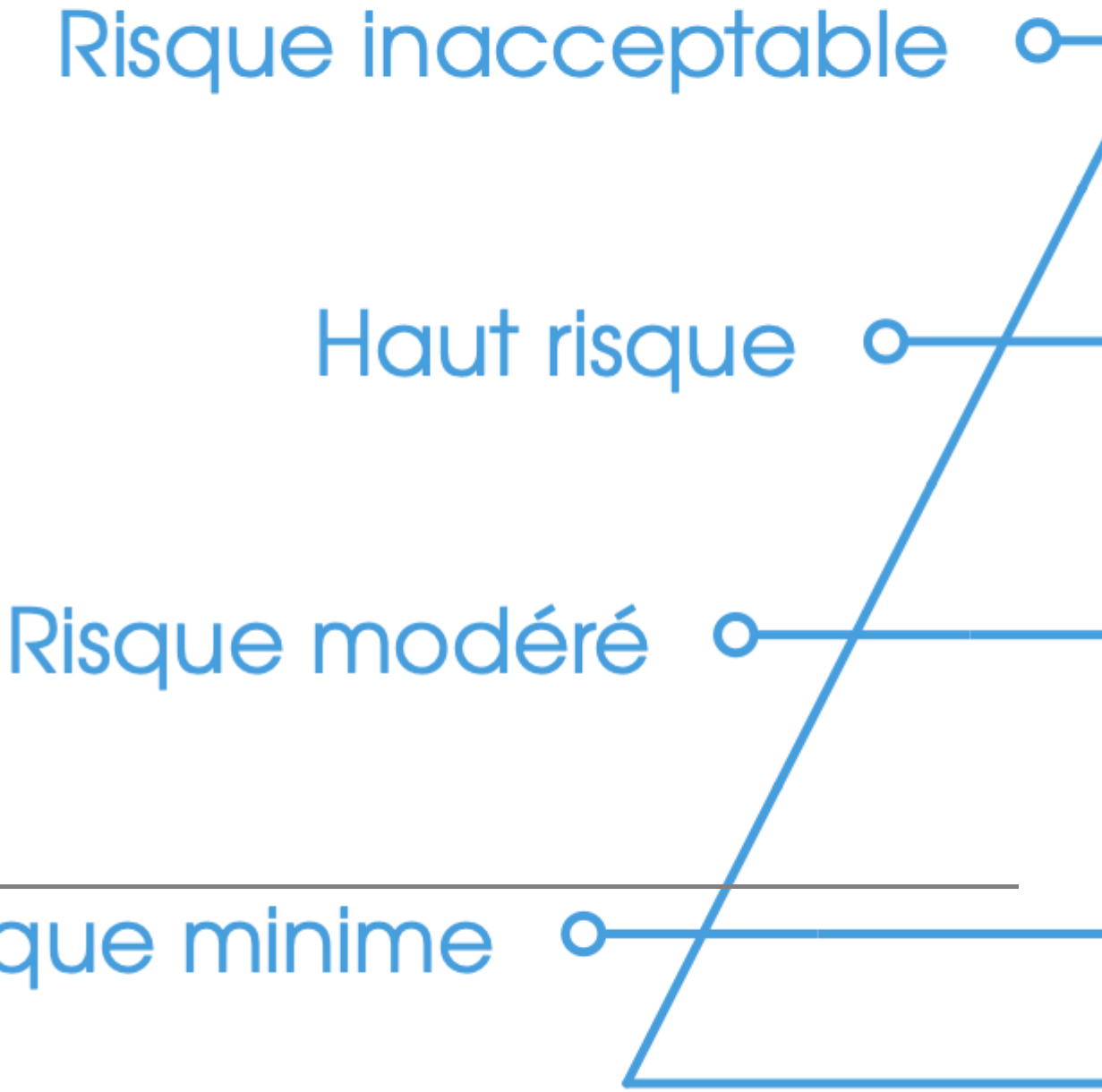
La loi sur l'IA vise à créer un cadre réglementaire permettant le déploiement de l'IA au niveau européen, tout en garantissant la maîtrise des risques et des problèmes que certains de ses usages génèrent. L'approche proposée par la Commission repose sur cette dualité, voire cette ambivalence, que l'on retrouve dans la version finale de la loi. En découle une régulation relativement « *parcimonieuse* »[5], venant concrétiser nombre des demandes de l'industrie (numérique) répétées en amont et tout au long du processus législatif.

Réguler « par les risques »

La première d'entre elles consiste à fonder l'approche réglementaire sur les risques. La régulation par les risques est notoirement plus flexible qu'une approche fondée sur les droits, sur laquelle repose par exemple le règlement général sur la protection des données (RGPD). En s'inspirant d'autres législations européennes visant la sécurité des produits (*product safety*), la loi minimise les obligations pour les développeurs, « déployeurs » et utilisateurs d'IA[6], lorsqu'ils procèdent à la mise sur le marché ou utilisent des systèmes ne représentant pas de risques apparents pour la sécurité des individus, tandis que ces obligations sont à l'inverse renforcées quand un système d'IA est perçu comme potentiellement risqué.

La loi distingue ainsi quatre types d'usages : les usages considérés comme inacceptables et donc interdits comme les systèmes de reconnaissance d'émotions déployés dans l'environnement de travail ou certaines formes de notation sociale[7] ; les usages à haut risque (soumis à des obligations renforcées) que sont les systèmes d'IA utilisés dans la gestion d'infrastructures critiques ; les usages à risque modéré (sujets à de faibles restrictions de transparence) et les usages dont les risques sont considérés comme minimes. Les fondements même de cette architecture proposée par la Commission n'ont pas été contestés lors des négociations. L'essentiel des débats au sein du Parlement et du Conseil a résidé dans la définition du contenu et du champ de chacune de ces catégories, présentées régulièrement sous la forme de la pyramide des risques suivante.

La pyramide des



Des secteurs et usages soumis à des obligations asymétriques

La loi sur l'IA ne réglemente pas l'IA en tant que telle, mais plutôt les usages de cette technologie. Les discussions autour de l'identification de ces usages, et de leur niveau de risque respectif, ont donné lieu à une multitude de controverses, témoignant du caractère relativement arbitraire de cette approche et des critères proposés par la Commission pour distinguer le niveau d'acceptabilité de ces systèmes. Chaque pan de l'industrie a ainsi tenté de minimiser les risques de ses propres pratiques, tandis que les organisations de la société civile dénonçaient, à l'inverse, l'absence de certains usages – comme la reconnaissance des émotions – dans des catégories de système d'IA amenés à être plus régulés.

Le résultat de ce « marchandage » est donc un texte législatif qui, bien que se voulant initialement horizontal (c'est-à-dire s'appliquant transversalement à l'ensemble des domaines de compétences de l'UE), introduit de nombreuses exemptions pour divers secteurs et usages (comme pour la santé et pour l'IA sécuritaire), limitant donc sa propre cohérence et sa « lisibilité ».

L'approche fondée sur les risques a également pour conséquence d'introduire des obligations minimales de mise en conformité pour les systèmes d'IA considérés comme sans risque. Plusieurs études académiques[8] soulignent que cette approche laisse, en effet, un large éventail de systèmes d'IA, pourtant susceptibles d'avoir un impact sérieux sur les droits fondamentaux, sans aucune réglementation pour ce qui est des risques spécifiquement liés à l'IA.

Cette observation est également confirmée pour les systèmes d'IA considérés comme à haut risque, dont la plupart peuvent être mis sur le marché par l'entremise d'un système d'autoévaluation à disposition des développeurs d'IA (avec plusieurs exceptions néanmoins). Cet état de fait a été dénoncé par des organisations comme la Quadrature du Net[9], pour qui cette approche fondée sur l'analyse des risques est destinée à rassurer le secteur privé et ne permet aucunement de garantir que les développeurs et « déployeurs » de systèmes d'IA respectent et protègent les droits humains.

La loi sur l'IA : un outil de dérégulation ?

Bien qu'il soit trop tôt pour évaluer les futurs impacts de cette législation sur le développement de l'IA au sein de l'UE, y compris en termes de protection des droits fondamentaux, on peut se demander si, paradoxalement, la loi sur l'IA ne pourrait pas contribuer à une forme de déréglementation et de nivellement par le bas des exigences au niveau européen.

Du fait de sa base légale (article 114 du TFUE – traité sur le fonctionnement de l'Union européenne), la loi, en effet, vise à « *empêcher les actions unilatérales des États membres qui risquent de fragmenter le marché et d'imposer des charges réglementaires encore plus lourdes aux opérateurs qui développent ou utilisent des systèmes d'IA* ». Certaines dispositions de la loi impliqueront une harmonisation maximale, susceptibles d'entraver les capacités des États membres à agir dans ce domaine. Ils devront, par là-même, laisser de côté

les règles nationales contradictoires et accepter les produits dits « conformes » sur leurs marchés.

De plus, si le régime visant les systèmes d'IA à haut risque semblait ambitieux dans la première mouture de la Commission, le Parlement et le Conseil ont relativement affaibli ces dispositions en introduisant un filtre, selon lequel seront véritablement régulés comme systèmes à haut risque, non pas les systèmes définis comme tels par les annexes techniques du texte, mais ceux posant en plus de cela un « *risque avéré significatif* ». Cette nouvelle disposition a donc créé une dérogation importante au projet de loi initial, vivement contestée par des organisations comme Access Now[10].

Enfin, cette loi repose sur un régime de mise en conformité, qui a pour corollaire de donner une importance significative aux standards techniques, dont la fonction sera de préciser la nature exacte des obligations pour les différents acteurs impactés par ce règlement. La formulation de ces standards est déléguée à CEN-CENELEC, deux organismes privés internationaux, rassemblant des comités de standardisation nationaux au niveau européen11. CEN-CENELEC travaille aujourd'hui sur une dizaine de standards qui permettra la mise en œuvre effective du texte de loi d'ici 2026. L'opacité de ces travaux de standardisation pour le public, qui contraste avec l'accès préférentiel réservé aux entreprises, a été vivement critiquée[12] et a mis en lumière, une fois de plus, les enjeux démocratiques de ces discussions techniques.

La (lointaine) mise en œuvre de la loi sur l'IA et la question de son « effet Bruxelles »

La nouvelle loi sur l'IA devrait être publiée au journal officiel de l'Union européenne dans le courant du printemps 2024, suite à son adoption formelle par le Parlement européen et par le Conseil de l'UE. Le calendrier de ce texte indique que la plupart de ses dispositions entreront en application deux ans après l'entrée en vigueur – à l'exception des interdictions visant les usages inacceptables.

Pendant cette période de transition, les discussions techniques et politiques ne devraient pas ralentir à l'échelle européenne, loin de là. D'une part, le processus d'adoption des standards de CEN-CENELEC, qui permettront la mise en œuvre de la loi, va ouvrir un nouvel espace techno-politique, dans lequel se jouera aussi le futur de la régulation de l'IA en Europe. D'autre part, à mesure que d'autres grandes puissances adoptent leur propre cadre réglementaire visant l'IA, la problématique de leur coordination ou de leur fragmentation devrait prendre encore plus d'importance, et renforcer les attentes autour de forums comme le Trade and Technology Council réunissant l'UE et les États-Unis, le G7, ou des forums globaux comme celui de Bletchley[13].

De ces tentatives d'alignement normatif entre blocs dépendra l'« effet Bruxelles »[14] de la loi sur l'IA au niveau global. Alors que l'UE avait bénéficié d'un réel avantage en étant le premier bloc à définir de nouvelles règles autour de la protection des données personnelles avec le RGPD, dont les principes se sont ensuite exportés à travers le monde, il semblerait que ce processus pourrait cette fois être limité par l'activisme normatif de la Chine et des États-Unis, soucieux de garder la mainmise sur ce secteur on ne peut plus stratégique.

Sources :

1. Benbouzid Bilel, Meneceur Yannick, Smuha Nathalie Alisa, « Quatre nuances de régulation de l'intelligence artificielle. Une cartographie des conflits de définition », *Réseaux*, n° 232-233, 2022/2-3, p. 29-64.

2. European Parliament, « MEPs ready to negotiate first-ever rules for safe and transparent AI », press releases, europarl.europa.eu, June 14, 2023.
3. Council of the European Union, « Artificial Intelligence Act: Council calls for promoting safe AI that respects fundamental rights », press release, consilium.europa.eu, December 6, 2022.
4. Bracy Jedidiah, « EU AI Act: Draft consolidated text leaked online », The International Association of Privacy Professionals (IAPP), iapp.org, January 22, 2024.
5. Bogucki Artur, Engler Alex, Perarnaud Clément, Renda Andrea, « The AI Act and emerging EU digital acquis », CEPS, ceps.eu, September 14, 2022.
6. Selon les termes consacrés par la loi.
7. En dépit des effets d'annonces, la loi sur l'IA n'interdit pas toutes les formes de systèmes de *social scoring* : Human Rights Watch, « EU: Artificial Intelligence Regulation Should Ban Social Scoring », hrw.org, October 9, 2023.
8. Stuurman Kees, Lachaud Éric, « Regulating AI. A label to complete the proposed Act on Artificial Intelligence », *Computer Law & Security Review*, vol. 44, sciencedirect.com, April 2022.
9. La Quadrature du net, « Règlement IA : l'Union européenne ne doit pas céder aux loggys sécuritaires », laquadrature.net, 5 octobre 2021.
10. Access Now, « Open letter: Council of the E.U. risks failing human rights in the AI Act », press releases, accessnow.org, November 28, 2023.
11. CEN-CENELEC fait référence au Comité européen de normalisation (CEN) et au Comité européen de normalisation en électronique et en électrotechnique (CENELEC). Ces deux organismes privés comptent trente- quatre pays membres et rassemblent des comités de standardisation nationaux à l'échelle européenne (au sens large, en incluant le Royaume-Uni par exemple).
12. Pouget Hadrien, « What will the role of standards be in AI governance? », Ada Lovelace Institute, April 5, 2023.
13. Le Trade and Technology Council vise à renforcer la coopération transatlantique, notamment en matière d'IA, des efforts poursuivis également à un niveau global au sein du G7 et du sommets de l'IA de Bletchley.
14. L'effet Bruxelles (plus communément appelé « *Brussels effect* » en anglais) est un concept d'Anu Bradford pour désigner la capacité de l'UE à développer des normes s'exportant en dehors de ses frontières.

Categorie

1. Droit

date créée

26 mars 2024
Auteur
perarnaudperarnaud