
La régulation de l'intelligence artificielle : approches internationales et britannique

Nathan Cambien et David Newton

Résumé : L'intelligence artificielle (IA) a connu récemment des évolutions rapides, notamment grâce à de nouveaux algorithmes capables de traiter de vastes quantités de données (*Big Data*) et à l'essor de l'apprentissage profond (*deep learning*) reposant sur les grands modèles de langage. Ces développements entraînent tant de nouvelles opportunités que des risques significatifs. Pour faire face à ce phénomène, diverses initiatives de réglementation ont émergé. À l'échelle internationale, des organismes comme l'Organisation de coopération et de développement économiques, le Conseil de l'Europe et l'Union européenne jouent un rôle de premier plan dans l'élaboration de cadres normatifs. Parallèlement, certains États, tels que le Royaume-Uni, développent leurs propres réglementations nationales. Le présent article examine plusieurs initiatives de réglementation de l'IA, en particulier celles du Royaume-Uni et de l'Union européenne, et compare leurs approches pour en dégager des enseignements dans une perspective de gouvernance mondiale de l'IA.

Mots-clés : intelligence artificielle, définition, régulation, organisation de coopération et de développement économiques (OCDE), principes, Union européenne, AI Act, autorités de surveillance du marché, initiative britannique, autorités de régulation sectorielles.

Référence électronique

Nathan Cambien et David Newton, « La régulation de l'intelligence artificielle : approches internationales et britannique », La régulation internationale de l'IA : enjeux et perspectives [Dossier], *Confluence des droits_La revue* [En ligne], 12 | 2024, mis en ligne le 18 décembre 2024. URL : <https://confluencedesdroits-larevue.com/?p=3678>.

Éditeur

Confluence des droits_La revue
DICE Éditions
UMR Droits International, Comparé et Européen (DICE)
3, avenue Robert Schuman
13628 Aix-en-Provence
France

La régulation de l'intelligence artificielle : approches internationales et britannique

Nathan Cambien¹

*Professeur, Université d'Anvers et membre du Service Juridique,
Commission européenne*

David Newton²

*Anciennement référendaire à la Cour de Justice de l'Union européenne,
actuellement membre du Service Juridique de la Cour des comptes européenne*

Il est incontestable que l'IA va avoir un impact considérable dans divers domaines de nos vies professionnelles et quotidiennes. Que ce soit pour l'assistance aux médecins dans l'établissement de diagnostics, l'identification des meilleurs candidats pour un emploi, ou la détermination des prix de vente de biens ou de services, le volume d'informations que l'IA propose de traiter de façon automatisée promet des gains de productivité dans de nombreux domaines. Dans le contexte de la profession juridique³, les systèmes d'IA peuvent être utilisés pour aider les avocats à rédiger des conclusions – si ce n'est pour les rédiger à leur place⁴ – et pour aider les cours et les tribunaux à trancher certains points, tels que la détermination, en droit des marques, de l'existence d'un risque de confusion⁵.

L'intelligence artificielle (IA) n'est pas un nouveau phénomène. Les recherches dans cette discipline ont débuté dès 1956, lors d'un colloque organisé à Dartmouth College aux États-Unis. Le domaine a ensuite connu des développements importants dans la seconde moitié du xx^e siècle. À titre d'exemple, rappelons la fameuse victoire du superordinateur Deep Blue contre le champion du monde du jeu d'échecs, Gary Kasparov, en 1997. Toutefois, ce phénomène a connu récemment des évolutions rapides, surtout au regard des nouveaux algorithmes capables d'analyser de vastes quantités de données (*Big Data*) et du développement des possibilités d'apprentissage profond (*deep learning*) fondé sur les grands modèles de langage⁶.

¹ Je tiens à remercier les organisateurs et les participants de la conférence « La régulation internationale de l'intelligence artificielle : enjeux et perspectives » (25 novembre 2022, Université de Lorraine). Les propos exprimés dans cet article sont strictement personnels.

² Les propos exprimés dans cet article sont strictement personnels.

³ Voir, de manière plus générale, N. Waisberg et A. Hudek, *AI for Lawyers: How Artificial Intelligence Is Adding Value, Amplifying Expertise, and Transforming Careers*, John Wiley & Sons, 2021.

⁴ Voir « Un avocat américain s'excuse après avoir utilisé ChatGPT : "Je pensais que c'était un moteur de recherche fiable" », *Le Soir*, 8 juin 2023.

⁵ Voir S. Dahan, R. Bhambhoria, S. Townsend et X. Zhu, « Analytics and EU Courts: The Case of Trademark Disputes », in T. Capeta, I. Goldner Lang et T. Perišin (dir.), *The Changing European Union: A Critical View on the Role of the Courts*, Hart Publishing, 2022.

⁶ Les grands modèles de langage sont des réseaux de neurones profonds entraînés sur de grandes quantités de texte non étiqueté utilisant l'apprentissage auto-supervisé ou l'apprentissage semi-supervisé afin de prédire une suite probable à une entrée donnée : voir « Grand modèle de langage ».

Il est à noter que certains systèmes d'IA générative, notamment les grands modèles de langage, parviennent, de manière autonome, à exécuter des tâches complexes, sans que leurs concepteurs ne comprennent eux-mêmes encore parfaitement le fonctionnement de la « boîte noire » des réseaux neuronaux de l'apprentissage profond⁷. L'utilisation des systèmes d'IA présente ainsi des risques significatifs⁸. En l'absence de mécanismes de surveillance appropriés, ces systèmes pourraient commencer à prendre des décisions autonomes aux conséquences graves⁹. Ainsi, récemment, certains acteurs du monde de l'IA ont déclaré qu'être en mesure de répondre aux risques posés par l'intelligence artificielle devrait être « une priorité mondiale, au même titre que d'autres risques pour nos sociétés, tels que les pandémies et la guerre nucléaire¹⁰ ». Afin d'éviter un développement incontrôlé des systèmes d'IA, des entrepreneurs et des chercheurs ont même demandé une pause dans la création de tout système d'IA qui serait plus puissant que le GPT-4 d'OpenAI¹¹.

En réponse, de nombreuses initiatives se sont développées afin de réguler l'IA, même si elles ne vont pas sans difficulté. D'abord, tout régime juridique doit définir clairement l'objet auquel il s'applique. Or, il n'est pas simple d'appréhender ce qu'est l'intelligence, et encore moins l'intelligence artificielle. Selon le mathématicien américain John McCarthy, le concept d'intelligence peut se définir, de manière générique, comme la partie computationnelle de la capacité à atteindre des objectifs dans le monde. Toutefois, du fait de la relation entre cette définition générale de l'intelligence et l'intelligence humaine elle-même, il nous est difficile de parvenir à une détermination des processus computationnels qu'il convient de qualifier d'intelligence¹². En effet, les définitions de

⁷ Voir, à cet égard, W. Douglas Haven, « Large language models can do jaw-dropping things. But nobody knows exactly why. », *MIT Tech Rev*, 4 mars 2024.

⁸ Nous pensons ainsi aux risques liés à l'usage d'algorithmes devant décider des demandes de carte de crédit, accusé de discrimination à l'égard des demandes introduites par les femmes, ou aux algorithmes de publicité en ligne, qui peuvent cibler les usagers en fonction de leur race, de leur religion ou de leur sexe, ou encore aux filtres de CV automatisés, qui éliminent les candidatures féminines. Une étude récente publiée dans *Science* a montré que les outils de prédiction des risques utilisés dans les soins de santé, qui touchent des millions de personnes aux États-Unis chaque année, présentent des biais raciaux importants: Z. Obermeyer, B. Powers, C. Vogeli, and S. Mullainathan, « Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations », *Science*, 25 octobre 2019, vol. 366, n° 6464, p. 447-453. Une autre étude a révélé que le logiciel utilisé par les principaux hôpitaux pour classer par ordre de priorité les bénéficiaires de greffes de rein était discriminatoire à l'égard des patients noirs: S. Ahmed, C.T. Nutt, N.D. Eneanya *et al.*, « Examining the Potential Impact of Race Multiplier Utilization in Estimated Glomerular Filtration Rate Calculation on African-American Care Outcomes », *J. Gen. Intern. Med.*, 2021, vol. 36, p. 464-471.

⁹ L'on pense notamment à l'intelligence artificielle générale, également appelée modèle de fondation. Ainsi, Sam Altman, PDG de la société OpenAI, dans une interview de 2023, entrevoit la possibilité d'une extinction de l'espèce humaine du fait de l'intelligence artificielle générale hors de contrôle.

¹⁰ Center for AI Safety, Statement on AI Risk.

¹¹ Future of Life Institute, Pause Giant AI Experiments: An Open Letter, 23 mars 2023.

¹² J. McCarthy, « What Is IA? Basic Questions ».

l'intelligence sont variables et reposent sur des caractéristiques humaines elles-mêmes difficiles à définir, telles que la conscience et l'usage du langage, ainsi que les capacités d'apprentissage, d'adaptation, d'abstraction et de raisonnement¹³.

Selon la société américaine IBM, l'intelligence artificielle, dans sa forme la plus simple, est un domaine qui combine l'informatique et des ensembles de données pour permettre la résolution de problèmes. Elle englobe les sous-domaines de l'apprentissage automatique et de l'apprentissage profond. Elle s'intéresse aux algorithmes d'IA qui cherchent à créer des systèmes experts qui exécutent des prévisions ou des classifications sur la base de données d'entrée¹⁴. De façon plus synthétique, l'IA peut encore se définir comme la capacité d'une machine à percevoir en continu un contexte, à en extraire un sens et à réagir à ce contexte¹⁵.

S'il est ainsi possible de définir les contours de la notion d'intelligence artificielle, il n'est pas pour autant aisé d'appréhender ses caractéristiques d'un point de vue réglementaire. En effet, une forme de réglementation visant à régir *ex ante* le développement de systèmes d'IA semble difficile à mettre en œuvre pour plusieurs raisons. D'une part, de tels systèmes peuvent faire l'objet d'une élaboration discrète, sans publicité particulière, et diffuse, en ce que plusieurs personnes peuvent travailler de façon séparée et non coordonnée avant que leurs travaux ne soient éventuellement mis en rapport par un tiers. D'autre part, le fonctionnement de l'IA peut être opaque, puisque son mode de fonctionnement peut demeurer mystérieux même pour les personnes chargées de sa mise en œuvre, en raison du fait que les résultats auxquels parviennent les grands modèles de langage varient en fonction notamment des données qui leur sont soumises lors de leur apprentissage¹⁶. À l'inverse, la nature autonome des systèmes d'IA pose des questions quant à la prévisibilité du dommage causé et à l'imputation de la responsabilité pour ce même dommage, suscitant alors des doutes quant à la pertinence des systèmes de réglementation *ex post*.

C'est dans ce contexte aux contours aussi mouvants qu'il convient d'étudier certains modèles de régulation des systèmes d'IA qui ont été mis en œuvre au niveau international, et notamment en droit de l'Union (I), avant d'examiner plus en détail l'initiative britannique en la matière, aux fins de mettre en parallèle deux approches réglementaires différentes (II). En conclusion, nous chercherons à tirer les leçons des divers modèles réglementaires examinés dans une perspective de réglementation globale de l'IA (III).

¹³ M. U. Scherer, « Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies », *Harv. J. L. & Tech.*, 2016, vol. 29, p. 353.

¹⁴ IBM, « Qu'est-ce que l'intelligence artificielle? ».

¹⁵ Twistthink, « Defining IA ».

¹⁶ M. U. Scherer, « Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies » art. cit., p. 356.

I. Les propositions internationales

S'agissant des initiatives prises au niveau international et supranational, nous nous focaliserons sur les trois organisations qui ont été particulièrement actives dans le domaine de la réglementation de l'IA : l'Organisation de coopération et de développement économiques (OCDE), le Conseil de l'Europe et l'Union européenne.

En premier lieu, s'agissant de l'OCDE, il convient de rappeler que celle-ci a, parmi ses différentes missions, pour objectif d'aider les gouvernements en mettant à leur disposition des mesures et analyses des incidences, économiques et sociales, des technologies et des applications liées à l'IA¹⁷. À cet effet, elle a mis en place l'Observatoire des politiques relatives à l'IA qui vise à faciliter le dialogue entre les pays et à fournir une analyse multidisciplinaire et factuelle des données sur les principaux domaines dans lesquels l'IA peut avoir des répercussions¹⁸.

L'OCDE a cherché à clarifier la définition des systèmes d'IA. Ainsi, un système d'IA est défini par l'OCDE comme un système basé sur une machine qui peut, pour des objectifs explicites ou implicites, déduire, à partir des entrées qu'il reçoit, comment générer des sorties telles que des prédictions, du contenu, des recommandations ou des décisions qui peuvent influencer les environnements physiques ou virtuels. Les différents systèmes d'IA varient dans leurs niveaux d'autonomie et d'adaptabilité après déploiement¹⁹. Cette définition aborde ainsi la question des objectifs explicites ou implicites (à savoir dérivé des règles auxquels le système est soumis) influant sur le fonctionnement du système, l'importance des entrées – provenant d'humains ou de la machine elle-même – conditionnant les sorties, l'influence mutuelle entre la machine et son environnement, ainsi que la faculté d'adaptation continue des systèmes d'IA.

En guise de règles éthiques entourant l'usage de l'IA, l'OCDE a adopté les *Principes sur l'intelligence artificielle* qui visent à établir que les utilisations de l'IA doivent être innovantes et dignes de confiance mais aussi qu'elles respectent les droits de l'homme et les valeurs démocratiques^{20 21}.

¹⁷ OCDE, « Comment faire en sorte que l'IA profite à l'ensemble de la société? ».

¹⁸ Voir le site web de l'Observatoire OCDE des politiques de l'IA.

¹⁹ [<https://oecd.ai/en/wonk/ai-system-definition-update>].

²⁰ OECD, Réunion du Conseil au niveau des ministres, 22-23 mai 2019, *Recommandation du Conseil sur l'Intelligence Artificielle, C/MIN(2019)3/FINAL*, tel qu'amendé le 3 mai 2024.

²¹ En outre, en 2022, l'OCDE a lancé le Cadre pour la classification des systèmes d'IA, qui vise à permettre de distinguer les applications d'IA en fonction de leur impact potentiel sur les individus, la société et la planète (« OECD Framework for the Classification of AI Systems », *Documents de travail de l'OCDE sur l'économie numérique*, 2022, n° 323, 80 p.). Le cadre est destiné à permettre aux utilisateurs de se concentrer sur des risques spécifiques typiques de l'IA, tels que les biais, l'explicabilité et la robustesse, et à aider à l'élaboration des politiques et des réglementations, dans la mesure où les caractéristiques des systèmes d'IA influencent les mesures techniques et procédurales nécessaires à leur mise en œuvre. Pour un exemple plus récent, voir OCDE, « Advancing Accountability in AI: Governing and Managing Risks Throughout the Lifecycle for Trustworthy AI », *Documents de travail de l'OCDE sur l'économie numérique*, 2023, n° 349, p. 69.

Les Principes de l'OCDE sur l'IA, 2019

Principes d'une approche responsable à l'appui d'une IA digne de confiance	Politiques nationales et coopération internationale à l'appui d'une IA digne de confiance
Croissance inclusive, développement durable et bien-être	Investir dans la recherche et le développement en matière d'IA
Valeurs centrées sur l'humain et équité	Favoriser l'instauration d'un écosystème numérique pour l'IA
Transparence et explicabilité	Façonner un cadre d'action favorable à l'IA
Robustesse, sûreté et sécurité	Renforcer les capacités humaines et préparer la transformation du marché du travail
Responsabilité	Favoriser la coopération internationale au service d'une IA digne de confiance

Bien que ces principes soient non contraignants, ils constituent, depuis leur élaboration en 2019, une référence mondiale et la première norme intergouvernementale servant à réguler l'utilisation de l'IA.

L'on retiendra donc du travail de l'OCDE, notamment, l'élaboration d'une définition des systèmes d'IA, ainsi que de normes tendant vers l'usage responsable de ceux-ci.

En deuxième lieu, s'agissant du Conseil de l'Europe, il convient de rappeler que les initiatives de cette organisation internationale visent à assurer la promotion des valeurs fondamentales, telles que la protection des droits de l'homme, l'état de droit et la démocratie²². Ainsi, en avril 2022, le Conseil de l'Europe a créé le Comité sur l'intelligence artificielle (CAI), chargé d'établir un processus de négociation internationale pour élaborer un cadre juridique sur le développement, la conception et l'application de l'IA, lequel se fonde sur les normes du Conseil de l'Europe en matière de droits de l'homme, de démocratie et d'État de droit, et qui se veut propice à l'innovation²³.

Les travaux du CAI ont débouché sur la Convention-cadre du Conseil de l'Europe sur l'intelligence artificielle et les droits de l'homme, la démocratie et l'État de droit, adoptée le 17 mai 2024. Ce texte couvre l'utilisation des systèmes d'IA par les pouvoirs publics – y compris les acteurs privés qui agissent pour leur compte – et les acteurs privés. La définition des systèmes d'IA²⁴ y est substantiellement similaire à celle mise en avant par l'OCDE. La convention pose en outre une série de principes

²² C'est dans ce contexte que s'inscrivent différentes recommandations et résolutions, à l'instar de la Recommandation n° 2102(2017) sur la convergence technologique, l'intelligence artificielle et les droits de l'homme, qui préconise une coopération étroite avec les institutions de l'Union européenne et l'Organisation des Nations Unies pour l'éducation, la science et la culture (Unesco) afin de garantir un cadre juridique cohérent et des mécanismes de supervision efficaces au niveau international.

²³ Voir le site web du Conseil de l'Europe.

²⁴ Article 2 de la convention.

fondamentaux auxquels doivent se conformer les activités menées dans le cadre du cycle de vie des systèmes d'IA²⁵. Elle exige que les parties à la convention veillent à ce que, lorsqu'un système d'IA a un impact significatif sur la jouissance des droits de l'homme, les personnes affectées par celui-ci disposent de garanties, de protections et de droits procéduraux effectifs²⁶. Cette convention instaure un mécanisme de suivi, la Conférence des Parties, composé de représentants officiels des Parties à la Convention et destiné à déterminer dans quelle mesure ses dispositions sont appliquées. Leurs analyses et suggestions aident à assurer le respect de la Convention-cadre par les États afin de garantir son efficacité à long terme²⁷.

Il est à noter toutefois que, bien que contraignante, la Convention-cadre ne prévoit pas de sanctions. L'examen de conformité aux dispositions de la convention provient de mécanismes de contrôle effectif²⁸, qui assurent un simple « monitoring » du respect des dispositions de celle-ci, sans pouvoir de fixer des amendes. Il apparaît donc que le caractère contraignant des mécanismes d'exécution de la convention est peu élevé.

En dernier lieu, l'organisation internationale ayant pris les initiatives les plus ambitieuses afin de réglementer l'IA est sans doute l'Union européenne. En 2018, elle a adopté la stratégie européenne en matière d'IA²⁹ qui vise à faire de l'UE une « plaque tournante de classe mondiale pour l'IA » et à veiller à ce que l'IA soit centrée sur l'humain et digne de confiance. En avril 2021, la Commission a présenté son train de mesures sur l'IA (*AI Package*), comprenant une communication sur la promotion d'une approche européenne de l'intelligence artificielle³⁰, une mise à jour du plan coordonné sur l'intelligence artificielle³¹ et une proposition de règlement établissant des règles harmonisées en matière d'IA (*AI Act*)³².

²⁵ Parmi lesquels dignité humaine et autonomie personnelle (article 7), égalité et non-discrimination (article 10), respect de la vie privée et protection des données à caractère personnel (article 11), transparence et contrôle (article 8), obligation de rendre des comptes et responsabilité (article 9), fiabilité (article 12), et innovation sûre (article 13).

²⁶ Voir l'article 15 de la Convention-cadre.

²⁷ Voir l'article 23 de la Convention-cadre.

²⁸ Décrits à l'article 26 de la Convention-cadre, qui prévoit, à son paragraphe 2, que les Parties à la Convention-cadre veillent à ce que les mécanismes de contrôle effectifs exercent leurs fonctions de manière indépendante et impartiale, et à ce qu'ils disposent des compétences, de l'expertise et des ressources nécessaires pour s'acquitter efficacement de leur mission de contrôle du respect des obligations nées de la Convention.

²⁹ Communication de la Commission, *L'intelligence artificielle pour l'Europe*, COM(2018) 237 final.

³⁰ Communication de la Commission au Parlement européen, au Conseil, au Comité économique et social européen et au Comité des Régions, *Favoriser une approche européenne en matière d'intelligence artificielle*, COM(2021) 205 final.

³¹ Communication de la Commission au Parlement européen, au Conseil, au Comité économique et social européen et au Comité des Régions, *Un plan coordonné dans le domaine de l'intelligence artificielle*, COM(2018) 795 final.

³² Proposition de Règlement du Parlement européen et du Conseil établissant des règles harmonisées concernant l'intelligence artificielle (Législation sur l'intelligence artificielle), COM(2021) 206 final.

Il est probable que l'AI Act³³, maintenant passé à l'état de loi à la suite de son adoption le 13 juin 2024, aura un effet considérable sur le développement de l'IA et sa réglementation. Comme le Règlement général sur la protection des données (RGPD)³⁴, l'AI Act pourrait devenir un standard mondial³⁵.

L'AI Act est basé sur les articles 16 et 114 TFUE, relatifs à la protection des données et au rapprochement des législations des États membres de l'Union respectivement. Sur le fond, il combine une réduction des obstacles au commerce avec des préoccupations se rapportant aux droits fondamentaux. Bien qu'en apparence novatrice, l'économie générale du règlement est tirée d'une décision de 2008 établissant un cadre commun de réglementation concernant la sécurité des produits³⁶. Les principales autorités chargées de l'exécution de l'AI Act sont des autorités de surveillance du marché, dont le fonctionnement est régi par le règlement 2019/1020 sur la surveillance du marché et la conformité des produits³⁷.

S'agissant de la définition des systèmes d'IA, l'AI Act les définit comme des systèmes automatisés conçus pour fonctionner à différents niveaux d'autonomie et pouvant faire preuve d'une capacité d'adaptation après leur déploiement, et qui, pour des objectifs explicites ou implicites, déduisent, à partir des entrées qu'ils reçoivent, la manière de générer des sorties telles que des prédictions, du contenu, des recommandations ou des décisions qui peuvent influencer les environnements physiques ou virtuels³⁸. L'on perçoit que cette définition des systèmes d'IA vise à prendre en compte les systèmes d'IA plus traditionnels, ou prédictifs³⁹, ainsi que les nouveaux systèmes d'IA génératives⁴⁰.

³³ Règlement (UE) 2024/1689 du Parlement européen et du Conseil du 13 juin 2024 établissant des règles harmonisées concernant l'intelligence artificielle et modifiant les règlements (CE) n° 300/2008, (UE) n° 167/2013, (UE) n° 168/2013, (UE) 2018/858, (UE) 2018/1139 et (UE) 2019/2144 et les directives 2014/90/UE, (UE) 2016/797 et (UE) 2020/1828 (règlement sur l'intelligence artificielle) (*JO L*, 2024/1689, 12.7.2024).

³⁴ Règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, et abrogeant la directive 95/46/CE (Règlement général sur la protection des données).

³⁵ M. Heikkilä, « A Quick Guide to the Most Important AI Law You've Never Heard Of », *MIT Technology Review*, 13 mai 2022.

³⁶ Décision n° 768/2008/CE du Parlement européen et du Conseil du 9 juillet 2008 relative à un cadre commun pour la commercialisation des produits et abrogeant la décision 93/465/CEE du Conseil.

³⁷ Règlement (UE) 2019/1020 du Parlement européen et du Conseil du 20 juin 2019 sur la surveillance du marché et la conformité des produits, et modifiant la directive 2004/42/CE et les règlements (CE) n° 765/2008 et (UE) n° 305/2011.

³⁸ L'on retrouve ainsi les mêmes caractéristiques de la définition de l'IA que celles apparaissant dans la définition de l'OCDE, notamment, l'existence d'objectifs tant explicites qu'implicites influant sur le fonctionnement du système, l'importance des entrées, provenant d'humains ou de la machine elle-même, conditionnant les sorties, l'influence mutuelle de la machine sur son environnement, et la capacité d'adaptation des systèmes d'IA.

³⁹ Tels que les algorithmes de recommandation de contenu.

⁴⁰ Et cela malgré la présence d'une définition des modèles d'IA à usage général également mise en avant dans le règlement (article 3, paragraphe 63, du règlement).

Le règlement suit une approche reposant sur une différenciation des types de risques associés aux systèmes d'IA et distingue ainsi les pratiques interdites en matière d'IA, les systèmes d'IA à haut risque et les risques faibles ou minimaux.

S'agissant des pratiques interdites en matière d'IA, cette interdiction se fonde sur les utilisations de l'IA considérées comme contraires aux valeurs de l'UE, en raison des violations des droits fondamentaux qu'elles entraînent. Il s'agit notamment de systèmes fondés sur la manipulation et l'exploitation des personnes, la reconnaissance des émotions, l'extraction non ciblée d'images faciales sur l'Internet, la notation sociale⁴¹, certaines applications de police prédictive, et les systèmes d'identification biométrique « en temps réel » et « à distance », interdits, dans ce dernier cas, sauf à des fins répressives spécifiques autorisées de façon indépendante.

La réglementation la plus détaillée concerne les systèmes d'IA considéré comme à haut risque, comprenant deux catégories de système (article 6). Tout d'abord, cette catégorie comprend tout système (autonome ou entrant dans la composition d'un produit) couvert par une des législations harmonisées de l'UE listées en annexe I et qui est soumis à une évaluation de conformité en vertu de ces mêmes textes⁴². Ensuite, cette catégorie comprend également des systèmes qualifiés automatiquement de haut risque en raison des domaines qu'ils touchent⁴³: sont ainsi considérés comme systèmes à haut risque les systèmes d'IA comprenant des secteurs tels que les infrastructures critiques, l'éducation, les services publics ou privés essentiels, l'emploi, l'administration de la justice⁴⁴.

Les systèmes d'IA à haut risque sont autorisés sur le marché européen, sous réserve que leurs fournisseurs veillent à ce qu'ils soient soumis, avant leur mise sur le marché ou leur mise en service, à une procédure d'évaluation de leur conformité⁴⁵, devant déboucher sur l'apposition d'une déclaration UE de conformité⁴⁶. Pour les systèmes ainsi mis sur le marché, le règlement prévoit notamment la mise en œuvre d'une

⁴¹ Voir la liste figurant à l'article 5 (Titre II) de l'*IA Act*. S'agissant de la question de notation sociale, voir plus spécifiquement la présentation du système de crédit social alloué aux entreprises et individus mis en œuvre en Chine.

⁴² Le champ est particulièrement large puisque sont visées les législations sur les machines, les jouets, les navires de plaisance, les ascenseurs, les équipements et systèmes de protection destinés à être utilisés en atmosphères explosibles, les installations à câbles, les dispositifs médicaux, la sûreté de l'aviation civile, la réception et la surveillance du marché des véhicules à deux ou trois roues et des quadricycles, ainsi que des véhicules agricoles et forestiers, les équipements marins, l'interopérabilité du système ferroviaire, etc.

⁴³ Et qui sont visés à l'annexe III du règlement.

⁴⁴ L'article 7 de l'*AI Act* permet également à la Commission de modifier le contenu de cette liste dès lors que les systèmes sont susceptibles de porter atteinte aux droits fondamentaux dans les domaines visés.

⁴⁵ Article 8 du règlement.

⁴⁶ Article 47.

procédure continue de gestion des risques (article 9), une politique contraignante de gouvernance des données utilisées pour l'entraînement afin d'assurer leur qualité et éviter les biais (article 10), l'établissement d'une documentation technique⁴⁷ (article 11), la tenue de journaux permettant de garantir un degré de traçabilité du fonctionnement d'un système d'IA (article 12), une obligation de transparence et d'information à l'attention des utilisateurs (article 13), un contrôle humain du système visant à prévenir ou minimiser les risques (article 14), ainsi que la mise en œuvre d'obligations générales d'exactitude, de robustesse et de cybersécurité (article 15).

Les autorités de surveillance du marché⁴⁸ auront le pouvoir de mener des inspections des systèmes d'IA à haut risque et auront accès à toutes les données utilisées par les fournisseurs dans le cadre de l'entraînement des systèmes d'IA⁴⁹. En cas de doutes quant aux risques pour la santé, la sécurité ou la protection des droits fondamentaux des personnes, l'autorité de surveillance du marché procèdera à une évaluation de la conformité du système d'IA concerné. Si elle constate que le système d'IA ne respecte pas les exigences et obligations énoncées dans le règlement, elle invitera sans tarder l'opérateur concerné à prendre toutes les mesures correctives appropriées pour mettre le système d'IA en conformité, le retirer du marché ou le rappeler dans un délai raisonnable et proportionné à la nature du risque⁵⁰.

Il est à noter ici que si l'AI Act prévoit la possibilité pour les personnes intéressées d'introduire une réclamation auprès d'une autorité de surveillance du marché⁵¹, il ne prévoit en revanche pas de mécanisme pour que les particuliers puissent contester les décisions de ces autorités. Or les mécanismes de plainte ont traditionnellement été les vecteurs les plus efficaces du développement de la jurisprudence de la Cour de justice de l'Union européenne, lorsque les autorités étaient réticentes à s'opposer aux pratiques des entreprises technologiques⁵².

Par ailleurs, pour les systèmes à faible risque, une obligation de transparence vis-à-vis des utilisateurs (article 50) est imposée, notamment lorsque les systèmes d'IA sont destinés à interagir avec une personne (*chatbots*), lorsqu'ils génèrent ou

⁴⁷ Dont le contenu est précisé à l'annexe IV.

⁴⁸ Pour une liste des autorités de surveillance du marché par pays et par secteur, voir le site de la Commission sur la mise en œuvre de la surveillance de marché au sein de l'UE, section intitulée « List of national market surveillance authorities ».

⁴⁹ Article 74.

⁵⁰ Article 79.

⁵¹ Article 85.

⁵² Voir, à cet égard, CJUE, 6 octobre 2015, *Schrems*, aff. C-362/14; M. Veale et F. Zuiderveen Borgesius, « Demystifying the Draft EU Artificial Intelligence Act », *Computer Law Review International*, 2021, vol. 22, n° 4, p. 97-112, p. 111.

manipulent des contenus (*deep fakes*) et lorsqu'ils ont pour finalité la reconnaissance d'émotion ou la catégorisation biométrique. Les utilisateurs doivent être informés de l'implication d'un système d'IA dans la communication, l'objectif étant de limiter les manipulations possibles.

En dernier lieu se trouve la catégorie des systèmes d'IA présentant un risque minime. Cette catégorie n'existe que par défaut et les usages en cause ne seront dès lors pas régulés par l'AI Act. Cet état de fait n'empêche naturellement pas ces utilisations d'être régulées par d'autres réglementations, telles que celles sur la sécurité des produits ou les données à caractère personnel⁵³.

S'agissant des sanctions, les États membres doivent établir, en vertu de l'article 99 de l'AI Act, les règles applicables pour tout manquement aux règles liées à l'IA contenues dans le règlement. Une variété de sanctions est prévue pour tout manquement persistant aux exigences du règlement, impliquant des amendes administratives d'un montant pouvant aller jusqu'à 35 000 000,00 € ou 7 % du chiffre d'affaires mondial de l'entreprise en cause.

Un certain nombre de critiques a déjà été émis à l'encontre de l'AI Act, non des moindres étant la question de savoir comment les autorités de surveillance du marché, habituées à statuer sur des questions de sécurité de produits physiques, tels que des dispositifs médicaux, sauront, conformément à l'exigence de l'article 79 de l'AI Act, trancher des questions ayant trait aux droits fondamentaux⁵⁴. Se pose également la question de l'interaction de l'AI Act avec les autres textes législatifs en vigueur dans le domaine du numérique⁵⁵, ainsi que le poids réglementaire issu de ce corpus législatif pesant sur les opérateurs économiques dans un domaine aux contours encore incertains.

⁵³ Voir, à cet égard, s'agissant de l'établissement automatisé d'une valeur de probabilité concernant la capacité d'une personne à honorer des engagements de paiement à l'avenir (« scoring ») et son acceptabilité au regard du règlement général sur la protection des données (Règlement (UE) 2016/679), arrêt du 7 décembre 2013, SCHUFA Holding AG, C-634/21, ECLI:EU:C:2023:957.

⁵⁴ Voir M. Veale et F. Zuiderveen Borgesius, « Demystifying the Draft EU Artificial Intelligence Act », art. cit. p. 105.

⁵⁵ Le groupe de réflexion Bruegel a publié à cet égard un recensement des textes législatifs en vigueur dans le domaine du numérique, qu'ils chiffrent au nombre de 116 (77 textes législatifs, 29 projets de textes, et 10 initiatives prévues). Les actes en cause concernent principalement (1) la recherche et l'innovation; (2) la politique industrielle; (3) la connectivité; (4) les données et la confidentialité; (5) les droits de propriété intellectuelle (DPI); (6) la cybersécurité; (7) le maintien de l'ordre et de la sécurité; (8) surveillance du marché; (9) le commerce électronique et la protection des consommateurs; (10) la concurrence, (11) les médias; et (12) les finances. Il est possible que tout ou partie de ces textes se recoupent avec les dispositions de l'AI Act, ce qu'il conviendra d'examiner au cas par cas.

En parallèle à l'aboutissement du modèle européen de réglementation de l'IA émerge un autre modèle, moins prescriptif que celui de l'Union européenne. Il s'agit de l'initiative britannique, décrite dans un *White Paper* sur une approche pro-innovation de la régulation⁵⁶.

II. L'initiative britannique

À la différence du cadre normatif détaillé caractérisant l'approche européenne, l'approche britannique se caractérisera, dans un premier temps à tout le moins, par des orientations, à l'attention des autorités nationales de régulation sectorielles, sur la mise en œuvre généralisée de principes transversaux. En effet, si le gouvernement britannique part du postulat que l'usage de l'IA va quasi inéluctablement engendrer des externalités négatives, qu'il sera nécessaire de surveiller et de contrôler, il souhaite ancrer toute réglementation contraignante dans une approche empirique de l'usage de l'IA et des besoins de réglementation y associés.

L'approche réglementaire proposée se dessine autour de quatre grands principes : une définition de l'IA fondée sur les caractéristiques du système en cause ; une approche qui se veut spécifique au contexte du déploiement de l'IA ; la fourniture d'un ensemble de principes transversaux devant être mis en œuvre par les diverses autorités de régulation dans le secteur dont elles ont la responsabilité de façon à contrer les risques liés à l'IA ; et la fourniture de nouvelles fonctions centrales pour aider les régulateurs à fournir le cadre réglementaire de l'IA⁵⁷.

En premier lieu, la définition de l'IA par le *White Paper* se fait par référence à deux caractéristiques qui génèrent, selon le gouvernement britannique, le besoin d'une réponse réglementaire sur mesure⁵⁸. La première de ces caractéristiques est l'adaptabilité des systèmes d'IA, du fait de leur alimentation permanente en données, ce qui leur permet de développer de nouvelles formes de raisonnement inductif non directement envisagées par leurs programmeurs. La deuxième caractéristique est l'autonomie des systèmes d'IA en raison de leur potentiel à prendre des décisions sans l'intention expresse ou le contrôle continu d'un humain. La souplesse dont témoigne cette approche doit permettre de pérenniser le cadre réglementaire mis en œuvre, et ce malgré la mise en service de nouvelles technologies non envisagées au moment de son entrée en vigueur.

⁵⁶ « A pro-innovation approach to AI regulation », en date 29 mars 2023.

⁵⁷ Les objectifs affichés par le gouvernement britannique par le biais de ce nouveau cadre réglementaire sont : la stimulation de la croissance en facilitant l'innovation et en réduisant l'incertitude grâce à l'application d'un cadre réglementaire ; l'accroissement de la confiance du public dans l'usage de l'IA ; la promotion des investissements au Royaume-Uni dans le domaine de l'IA par une approche souple et capable d'intégration avec d'autres cadres réglementaires au niveau mondial. Il est permis de se demander dans quelle mesure cette initiative représente une tentative du gouvernement britannique de tirer parti d'un cadre réglementaire plus souple que celui en développement au niveau de l'Union européenne, pour établir un avantage comparatif en sa faveur.

⁵⁸ Voir la section III.2.1 du *White Paper* intitulée « Defining Artificial Intelligence ».

En deuxième lieu, l'approche britannique envisage avant tout les usages d'une technologie, et non pas la technologie elle-même. Dans son *White Paper*, le gouvernement britannique met l'accent sur une approche spécifique au contexte, permettant aux autorités de régulation sectorielles d'évaluer les risques liés à l'utilisation ou non de l'IA. Cette évaluation des risques inclut dès lors le défaut d'exploitation des capacités des systèmes d'IA⁵⁹. Cette approche spécifique au contexte se veut une réponse proportionnée au risque et cherche à éviter d'étouffer l'innovation ou de manquer des occasions de tirer parti des bénéfices offerts par l'IA.

En troisième lieu, les autorités de régulation sectorielles devront mettre en œuvre les principes transversaux évoqués précédemment, qui s'appuient sur les principes de l'OCDE sur l'intelligence artificielle et qui sont au nombre de cinq. Premièrement, il convient d'assurer la *sûreté, la sécurité et la robustesse des systèmes d'IA*. Ainsi, ces systèmes doivent fonctionner de manière robuste, sécurisée et sûre tout au long du cycle de vie de l'IA et les risques liés à son utilisation doivent être continuellement identifiés, évalués et gérés. Deuxièmement, il convient d'assurer un degré de *transparence et explicabilité approprié*. La transparence fait référence à la communication des informations relatives à un système d'IA aux personnes concernées, par exemple aux fins de savoir comment, quand et pour quelles raisons un système d'IA est utilisé. L'explicabilité fait référence à la possibilité d'accéder, d'interpréter et de comprendre les processus décisionnels d'un système d'IA. Un niveau approprié de transparence et d'explicabilité signifie que les régulateurs disposent de suffisamment d'informations sur les systèmes d'IA pour mettre en œuvre les autres principes de l'OCDE. Troisièmement, les systèmes d'IA doivent *respecter le principe d'équité*. Ainsi, ils ne doivent pas porter atteinte aux droits des individus ou des organisations, ni discriminer de façon injustifiée entre les individus, ni créer des résultats de marché injustes. Quatrièmement, des mesures de *gouvernance* doivent être mises en place pour assurer une surveillance efficace de la fourniture et de l'utilisation des systèmes d'IA avec des chaînes de *responsabilité* clairement établies tout au long du cycle de vie de l'IA. Cinquièmement, il doit exister une *possibilité de recours et de réparation*. Ainsi, les utilisateurs, les tiers et les acteurs du cycle de vie de l'IA doivent pouvoir contester une décision ou une conséquence de l'utilisation d'un système d'IA qui leur est préjudiciable ou qui crée un risque matériel de préjudice.

⁵⁹ Par exemple, dans le domaine de la santé, de façon à établir un diagnostic plus précis. Il peut par ailleurs y avoir un coût d'opportunité important lié au fait de ne pas avoir accès à l'IA dans les opérations critiques pour la sécurité, depuis l'industrie lourde jusqu'aux logiciels de reconnaissance faciale utilisées par les forces de l'ordre dans des contextes tels que les grandes rencontres sportives, qui peuvent être aptes à dégénérer dans la violence.

Les principes transversaux qui structurent l'approche britannique seront mis en œuvre dans le cadre des régimes réglementaires déjà existants⁶⁰. Dans un premier temps, l'application de ces principes ne sera pas contraignante, permettant ainsi de veiller au cours de cette période à leur effectivité. Par la suite⁶¹, il est prévu de renforcer et de clarifier les obligations des autorités de régulation par la création d'un devoir de ces dernières de tenir dûment compte de ces principes. L'institution d'un devoir plutôt que d'une obligation vise à permettre à ces autorités d'exercer leur pouvoir discrétionnaire concernant la pertinence, à tout instant, des différents principes dans leurs domaines respectifs. Ainsi, le gouvernement britannique attend des autorités de régulation qu'elles évaluent les principes transversaux et les appliquent aux cas d'utilisation de l'IA qui relèvent de leur mandat, qu'elles publient des lignes directrices sur la manière dont les principes interagissent avec la législation existante pour aider l'industrie à appliquer ces principes, et qu'elles soutiennent les entreprises opérant dans le domaine d'activité de plusieurs régulateurs en collaborant et en produisant des orientations claires et cohérentes, y compris conjointes le cas échéant.

En dernier lieu, le gouvernement britannique propose de soutenir les autorités sectorielles dans leur mission de régulation des systèmes d'IA en participant à l'élaboration de lignes directrices qui les aideront à mettre en œuvre les principes transversaux⁶². Ainsi, le *White Paper* identifie certaines fonctions que le gouvernement pourra utilement remplir à cet égard. Premièrement, le gouvernement central devra mettre en œuvre la surveillance, l'évaluation et le retour d'information (*feedback*): il s'agit ici principalement d'élaborer et maintenir un cadre central de suivi et d'évaluation pour évaluer les impacts intersectoriels et sectoriels du nouveau régime, afin de permettre au gouvernement de s'assurer que la cadre réglementaire fonctionne comme prévu. Deuxièmement, il devra soutenir la mise en œuvre cohérente des

⁶⁰ À titre d'exemple, le Policy Paper de juillet 2022, intitulé « Establishing a pro-innovation approach to regulating AI », rappelle que certaines autorités de régulation britanniques sont déjà actives dans le domaine de l'IA. Ainsi, le Bureau du commissaire à l'information (Information Commissioner's Office) a publié plusieurs documents d'orientation, tels que des lignes directrices sur l'IA et la protection des données; la Commission pour l'égalité et les droits de l'homme (Equality and Human Rights Commission) s'est engagée à fournir des orientations sur la manière dont la loi sur l'égalité s'applique à l'utilisation des nouvelles technologies, telles que l'IA, dans la prise de décision automatisée; l'Agence de réglementation des médicaments et des produits de santé (Medicines and Healthcare products Regulatory Agency) a mené des consultations sur d'éventuelles modifications du cadre réglementaire pertinent pour garantir que les exigences de ce cadre fournissent un degré élevé d'assurance que les dispositifs médicaux présentent un degré de sûreté acceptable; et l'Agence responsable de la santé et de la sécurité (Health and Safety Executive) s'est engagée à entamer des recherches collaboratives avec l'industrie et le monde universitaire, afin d'arriver à une détermination claire des implications de l'IA sur la santé et la sécurité sur le lieu de travail. Il est à noter que les différentes autorités de régulation nationales britanniques sont recensées plus largement sur Wikipédia.

⁶¹ C'est-à-dire lorsque le calendrier législatif le permettra – voir la section III.2.4. du *White Paper*.

⁶² Cette approche n'est pas sans rappeler la création du Comité européen de l'intelligence artificielle prévue dans l'AI Act proposé par la Commission européenne.

principes, en élaborant des lignes directrices centrales pour aider les régulateurs à les mettre en œuvre. Troisièmement, il appartiendra au gouvernement de mener une évaluation transversale des risques, de façon à garantir que tout nouveau risque puisse être traité et ne tombe pas dans les interstices entre les domaines de compétence des autorités sectorielles. Quatrièmement, il conviendra qu'il prête une assistance aux innovateurs en identifiant les problèmes réglementaires transversaux qui ont des impacts néfastes sur l'innovation, notamment par la mise en place de bacs à sable réglementaires pour tester la pertinence des modalités réglementaires envisagées. Cinquièmement, il devra mener des actions de sensibilisation des entreprises et du grand public sur la réglementation et les risques liés à l'IA. Sixièmement, le gouvernement central devra opérer un balayage d'horizon, notamment par une évaluation des risques nouveaux et émergents liés à l'IA ou à sa non-utilisation, en collaboration avec les acteurs de l'industrie, les universités, les autorités de régulation et les partenaires au niveau mondial. Enfin, le *White Paper* envisage que le gouvernement central assure l'interopérabilité du cadre national avec les cadres réglementaires internationaux.

La comparaison entre l'approche en droit de l'Union et l'approche britannique de la réglementation de l'IA laisse poindre deux approches réglementaires contrastées : la gouvernance en amont par opposition à « l'innovation sans permission »⁶³. Ainsi, tandis que le cadre européen introduit des exigences, pour les systèmes d'IA jugées à haut risque, de se soumettre à une procédure d'évaluation de conformité (ex ante), le modèle britannique privilégie un redressement des torts éventuels dus à l'usage des systèmes d'IA (ex post). Pour résumer la différence d'approche avec un aphorisme, l'approche britannique vise à permettre de corriger toute erreur rapidement, tandis que l'approche en droit de l'Union vise à rendre l'erreur impossible. Il est légitime dès lors de s'interroger sur le poids réglementaire que fait peser sur les opérateurs économiques le modèle européen, et ce malgré la volonté de certitude juridique pour la protection des droits fondamentaux que ce modèle offre.

L'on retiendra par ailleurs la place importante réservée, dans les deux cas, aux autorités de régulation sectorielles. En effet, du fait de la flexibilité institutionnelle qui leur est accordée, ces autorités peuvent combiner des fonctions législatives, judiciaires et exécutives. En effet, elles peuvent, le cas échéant, à la fois définir les orientations réglementaires dans un domaine précis, trancher certains litiges à l'instar d'un tribunal et faire exécuter certaines décisions et orientations. En outre, elles peuvent exercer une influence plus subtile sur la conduite des acteurs du marché, en collectant et en publiant des informations sur les risques pour la sûreté

⁶³ Voir, à cet égard, A. McAfee, « European Competitiveness, and How Not to Fix It ».

et la sécurité provenant de la conduite des acteurs d'un marché ou de leurs produits. Les autorités de régulation et de surveillance ont donc un avantage institutionnel comparatif important par rapport au pouvoir judiciaire ou aux organes législatifs dans l'application de règles et de principes⁶⁴. Toutefois, dans le modèle britannique, cette combinaison de compétences est perçue comme permettant de faire émerger les informations nécessaires à l'élaboration, à titre ultérieur, d'un cadre législatif se rapportant à l'usage des systèmes d'IA.

La souplesse du modèle britannique, fondé sur l'émergence « organique » d'un cadre législatif contraignant tiré d'une accumulation d'expériences concrètes, allié à un régime d'innovation sans permission préalable, permettra sans doute aux entreprises de développement de systèmes d'IA installées au Royaume-Uni d'évoluer dans un contexte règlementaire plus favorable qu'au sein de l'Union⁶⁵. Alliée à la grande disponibilité de données en langue anglaise, cette situation pourrait permettre au Royaume-Uni de développer un avantage de nature à inciter l'implantation d'entreprises de développement de systèmes d'IA. En outre, de par son encouragement à faire usage des systèmes d'IA lorsque les circonstances le justifient, le système anglais encourage l'adoption de méthodes de travail ancrées dans la technologie, et pourrait également représenter un avantage compétitif pour les entreprises offrant leurs services au Royaume-Uni.

Il convient encore d'observer que, à la différence de l'AI Act, le *White Paper* ne prévoit pas explicitement de sanctions pour des torts qui pourraient être causés par les systèmes d'IA. Cela laisse penser que le gouvernement britannique estime que les autorités de régulation britanniques disposent actuellement de suffisamment de pouvoirs pour sanctionner les manquements aux principes transversaux explicités ci-dessus. À cet égard, il convient en outre de rappeler que, dans les deux systèmes, tout dommage dû à l'utilisation d'un système d'IA (tel que la détermination par le système d'IA d'un résultat discriminatoire faisant l'objet d'un contrôle humain inadéquat) pourra éventuellement faire l'objet d'une action en responsabilité délictuelle. Cette alternative présente l'avantage de responsabiliser les fournisseurs des systèmes d'IA en faisant planer sur eux, pareille à une épée de Damoclès, la menace de dommages

⁶⁴ Elles disposent en effet de compétences, généralement attribuées aux organes législatifs, puisqu'elles peuvent procéder à des missions d'enquête et recueillir des données, afin d'exercer ensuite leurs fonctions règlementaires, tout en ayant la possibilité, classiquement octroyée aux tribunaux, de régler de façon casuistique les difficultés à même de surgir dans la mise en œuvre concrète des systèmes d'IA. Voir J. M. Landis, *The Administrative Process*, Yale University Press, 1938, p. 22, cité in M. U. Scherer, « Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies », art. cit., p. 382.

⁶⁵ À cet égard, il importe en outre de garder en tête que le nouvel AI Act fait partie d'une panoplie de réglementations européennes affectant le domaine du numérique. L'on pensera ainsi au Digital Markets Act, au Digital Services Act, au Data Act et au Data Governance Act, parmi bien d'autres initiatives affectant ce domaine.

et intérêts devant être versés pour toute infraction aux droits individuels et collectifs dus à l'usage de systèmes d'IA qu'ils mettent sur le marché. Cela étant, il est possible d'estimer qu'un mécanisme de certification de conformité du système d'IA mis sur le marché, à tout le moins d'ordre facultatif, fait défaut dans le système britannique. Une telle certification offrirait une marge de sécurité juridique aux entreprises voulant faire preuve de bonne foi en obtenant une certification du système développé par elles. On pourrait ainsi imaginer, sur la base de cette certification facultative, le plafonnement des dommages et intérêts dus au titre d'une action en réparation pour des infractions aux droits individuels⁶⁶. Un tel mécanisme offrirait en même temps une assurance aux utilisateurs, tiers et acteurs du cycle de vie de l'IA que certains systèmes mis sur le marché ont fait l'objet d'un examen en aval par une autorité de régulation. L'apposition d'une telle marque de certification serait sans doute apte à encourager les entreprises à la rechercher spontanément.

III. Vers une réglementation globale de l'IA ?

Il ressort de ce qui précède que plusieurs organisations internationales – chacune à partir de ses propres compétences et objectifs – et différents pays, dont le Royaume-Uni, ont pris des initiatives importantes afin de soumettre l'IA à un cadre réglementaire. Quelles leçons peut-on tirer de cette brève analyse de certaines de ces initiatives ?

Premièrement, il est impératif d'encadrer l'usage de l'intelligence artificielle, vu les risques de dérapage. Les initiatives récentes à cet égard sont donc les bienvenues. Toutefois, force est de constater, d'une part, que les problèmes liés à l'usage de l'IA sont avérés depuis un moment⁶⁷. D'autre part, des normes juridiques contraignantes visant à adresser ces problèmes ont longtemps fait défaut. Sans vouloir douter de l'importance du *soft law* (recommandations, résolutions, etc.), la mise en œuvre de certaines règles s'avère utile. L'AI Act et la Convention-cadre du Conseil de l'Europe sur l'intelligence artificielle et les droits de l'homme, la démocratie et l'État de droit combrent en partie cette lacune.

Deuxièmement, le modèle de régulation britannique enseigne qu'il ne convient pas simplement de proscrire l'utilisation des systèmes d'IA dans certains domaines, tels que la notation sociale, à l'instar de la réglementation proposée en droit de l'Union. Il convient également de tirer parti des nouvelles fonctionnalités de l'IA, de façon à optimiser certains processus, notamment les processus médicaux. Or cela n'est pas sans évoquer de nouvelles questions, notamment celles de savoir comment

⁶⁶ M. U. Scherer, « Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies », art. cit., p. 393.

⁶⁷ Voir *supra* note 5.

tirer parti de ces systèmes tout en maintenant la qualification et l'autonomie des professionnels chargés de les mettre en œuvre, et pour lesquels la tentation de se fier entièrement au diagnostic ou à la solution proposée par le système d'IA sera forte, que ce soit dans le domaine médical, dans la décision d'allouer un crédit ou dans la détermination de la tarification commerciale à appliquer. Comment éviter, dans un tel contexte, une perte de compétences (*deskilling*) des professionnels concernés ?

Troisièmement, il ne suffit pas d'adopter des actes spécifiques destinés à réglementer de manière globale et exhaustive l'intelligence artificielle. Une telle approche ne pourrait pas confronter tous les enjeux de l'IA. En effet, il va falloir prendre l'IA en compte dans tous les domaines juridiques et tous les secteurs économiques et sociaux. Prenons l'exemple du droit de la concurrence. Traditionnellement, la notion d'entente couvre certains types de concertation entre différentes entreprises, ce qui présuppose un aspect « humain », à savoir une volonté de coordonner les prix de vente par exemple. La question qui se pose maintenant est de savoir comment il convient d'appliquer l'interdiction sur les ententes quand les prix sont déterminés par des algorithmes qui ont appris de manière autonome à fixer les prix à un niveau idéal pour les entreprises concernées. Si à travers des algorithmes, différentes entreprises parviennent à fixer les prix, convient-il de considérer que c'est le créateur de l'algorithme qui est responsable ? Un tel point de vue n'est pas évident lorsque le créateur n'a pas instruit l'algorithme sur la mise en œuvre d'ententes, mais que l'algorithme lui-même est arrivé au résultat que cela est la meilleure solution pour les entreprises concernées. Cet exemple illustre parfaitement qu'il convient d'adapter le droit ou l'interprétation des règles pour tenir dûment compte du phénomène de l'IA⁶⁸.

Quatrièmement, vu le caractère mondial de l'IA, une coordination internationale apparaît indispensable. En effet, il est évident que les conséquences de l'utilisation de l'IA ne s'arrêtent pas aux frontières des différents pays. Ainsi, la détermination par l'IA des prix de vente sur les plateformes de commerce international peut avoir pour conséquence que des commerçants locaux font faillite. Pareillement, un système d'IA développé dans un pays peut avoir des conséquences déterminantes sur le résultat des élections dans un autre pays, sans même mentionner la possibilité d'utiliser l'IA pour diriger des armes telles que des drones pilotés à distance.

Il s'ensuit qu'il est souhaitable et même impératif d'avoir une coordination internationale entre différents pays et différentes organisations internationales ayant un caractère régional telle que l'Union européenne. Cette nécessité est reconnue par

⁶⁸ Voir A. Ezrachi et M.E. Stucke, « Artificial Intelligence & Collusion: When Computers Inhibit Competition », *University of Illinois Law Review*, 2017, vol. 2017, p. 1775 ; A. Ezrachi et M.E. Stucke, *Virtual Competition: The Promises and Perils of the Algorithm-Driven Economy*, Harvard University Press, 2016.

diverses organisations internationales⁶⁹. Nous renvoyons également au principe 2.5 des principes sur l'intelligence artificielle de l'OCDE qui est intitulé « Favoriser la coopération internationale au service d'une IA digne de confiance⁷⁰ ».

La question essentielle qui se pose est de savoir comment parvenir à une telle coordination internationale⁷¹. Il existe plusieurs forums bilatéraux, tels que le TTC Joint Roadmap for Trustworthy AI and Risk Management⁷², et multilatéraux, tels que le Global Partnership on Artificial Intelligence⁷³, dans le cadre desquels la régulation de l'IA est abordée. Par ailleurs, la Convention-cadre du Conseil de l'Europe sur l'intelligence artificielle et les droits de l'homme, la démocratie et l'État de droit est ouverte à la signature à tous les pays ayant participé à son élaboration⁷⁴, notamment l'Argentine, Israël, le Japon et les États-Unis. Cependant, dès lors, notamment, que la Chine n'est pas impliquée, il est difficile de qualifier ces initiatives de véritablement globales.

Les lignes bougent toutefois à cet égard, ainsi qu'il ressort de la déclaration de Bletchley Park⁷⁵ ayant trait à la sécurité associée à l'usage de l'intelligence artificielle, et à laquelle la Chine est signataire. Il convient par ailleurs de noter la résolution votée par les Nations Unies intitulée « Saisir les possibilités offertes par des systèmes d'intelligence artificielle sûrs, sécurisés et dignes de confiance pour le développement durable »⁷⁶. Cette résolution vise à encourager les pays à protéger les droits de l'homme, à protéger les données personnelles et à surveiller les risques liés à l'IA sur une base, toutefois, non juridiquement contraignante.

⁶⁹ Ainsi le Comité des ministres du Conseil de l'Europe a pu invoquer « la nécessité de maintenir une collaboration étroite avec les autres organisations internationales et d'autres parties prenantes afin d'échanger des bonnes pratiques et de faire en sorte que le cadre juridique international soit aussi cohérent que possible » (Comité des ministres, *La convergence technologique, l'intelligence artificielle et les droits de l'homme*, adoptée à la 1297^e réunion des Délégués des ministres, 17 octobre 2017).

⁷⁰ En vertu de ce principe, « les pouvoirs publics devraient coopérer au niveau transnational et entre les différents secteurs d'activité afin de partager les connaissances, de développer des standards techniques et de progresser vers une approche responsable en matière d'IA ». OCDE, Réunion du Conseil au niveau des ministres, 22-23 mai 2019, *Recommandation du Conseil sur l'Intelligence Artificielle*, C/MIN(2019)3/FINAL.

⁷¹ Voir S. Chesterman, *We, the Robots? Regulating Artificial Intelligence and the Limits of the Law*, Cambridge University Press, 2022, spec. le chapitre 8, « New Institutions », p. 195-223.

⁷² *EU-US Trade and Technology Council Joint Roadmap on Evaluation and Measurement Tools for Trustworthy AI and Risk Management*, 1^{er} décembre 2022.

⁷³ Voir le site web du GPIA et le Joint Statement from Founding Members of the Global Partnership on Artificial Intelligence, 15 juin 2020.

⁷⁴ Ainsi qu'il ressort de son Article 30. L'article 31 prévoit encore la possibilité pour d'autres états d'adhérer ultérieurement à la convention.

⁷⁵ The Bletchley Declaration by Countries Attending the AI Safety Summit, 1^{er}-2 novembre 2023.

⁷⁶ Voir résolution A/78/L. 49 du 11 mars 2024, intitulée « Saisir les possibilités offertes par des systèmes d'intelligence artificielle sûrs, sécurisés et dignes de confiance pour le développement durable ».

Une solution éventuelle au problème de la coordination internationale consisterait en l'organisation d'une Conférence des Parties (COP) sur l'intelligence artificielle, inspirée par les COP sur le climat. En effet, certaines COP sur le climat ont abouti à la conclusion d'accords globaux contenant des normes contraignantes, tels que le Protocole de Kyoto⁷⁷ ou encore l'Accord de Paris⁷⁸. Il n'en reste pas moins que la mise en œuvre d'éventuels accords contraignants régissant l'IA dépendra en large partie de la volonté et de la détermination des autorités nationales compétentes, notamment les autorités de l'état où les systèmes d'IA sont localisés.

⁷⁷ [<https://unfccc.int/sites/default/files/resource/docs/french/cop3/kpfrench.pdf>].

⁷⁸ [https://unfccc.int/sites/default/files/french_paris_agreement.pdf].