COMMISSION
EUROPÉENNE

Bruxelles, le 29.7.2025
C(2025) 5053 final

COMMUNICATION DE LA COMMISSION

Lignes directrices de la Commission sur la définition d'un système d'intelligence artificielle au sens du règlement (UE) 2024/1689 (règlement sur l'IA)

I. Objet des lignes directrices

(1) Le règlement (UE) 2024/1689 du Parlement européen et du Conseil (ci-après le «règlement sur l'IA»)[1] est entré en vigueur le 1er août 2024. Il établit des règles harmonisées pour le développement, la mise sur le marché, la mise en service et l'utilisation de l'intelligence artificielle (IA) dans l'Union[2]. Son objectif est de promouvoir l'innovation dans l'IA et son adoption, tout en garantissant un niveau élevé de protection de la santé, de la sécurité et des droits fondamentaux dans l'Union, y compris la démocratie et l'état de droit.

(2) Le règlement sur l'IA ne s'applique pas à tous les systèmes, mais uniquement à ceux qui répondent à la définition d'un «système d'IA» au sens de l'article 3, point 1), dudit règlement. La définition du terme «système d'IA» est donc déterminante pour comprendre le champ d'application du règlement sur l'IA.

(3) L'article 96, paragraphe 1, point f), du règlement sur l'IA impose à la Commission d'élaborer des lignes directrices sur l'application de la définition d'un système d'IA telle qu'elle figure à l'article 3, point 1, dudit règlement. En publiant les présentes lignes directrices, la Commission vise à aider les fournisseurs et les autres personnes concernées, notamment les acteurs du marché et les parties prenantes institutionnelles, à déterminer si un système constitue un système d'IA au sens du règlement sur l'IA, ce qui facilitera l'application effective et le contrôle du respect dudit règlement.

(4) La définition du terme «système d'IA» est entrée en application le 2 février 2025[3], à l'instar d'autres dispositions des chapitres I et II du règlement sur l'IA, notamment son article 5 sur les pratiques interdites en matière d'IA. Étant donné que la définition du terme «système d'IA» est déterminante pour comprendre le champ d'application du règlement sur l'IA, y compris les pratiques interdites, l'adoption des présentes lignes directrices se fait parallèlement à celle des lignes directrices de la Commission sur les pratiques interdites en matière d'intelligence artificielle.

(5) Les présentes lignes directrices tiennent compte des résultats d'une consultation des parties prenantes et de la consultation du Comité européen de l'intelligence artificielle.

(6) Eu égard à la grande diversité qui caractérise les systèmes d'IA, les présentes lignes directrices ne peuvent proposer une liste exhaustive de tous les systèmes d'IA potentiels. Cette orientation est conforme au considérant 12 du règlement sur l'IA, qui précise que la notion de «système d'IA» devrait être clairement définie tout en offrant «la souplesse nécessaire pour tenir compte des évolutions technologiques rapides dans ce domaine». La définition du terme «système d'IA» ne doit donc pas être appliquée mécaniquement; chaque système doit être évalué sur la base de ses caractéristiques propres.

[1] Règlement (UE) 2024/1689.
[2] Article 1er du règlement sur l'IA.
[3] Article 113, troisième alinéa, point a).

(7) Les lignes directrices ne sont pas contraignantes. En définitive, seule la Cour de justice de l'Union européenne (CJUE) peut donner une interprétation faisant autorité du règlement sur l'IA.

II. Objectif et principaux éléments de la définition du terme «système d'IA»

(8) L'article 3, point 1), du règlement sur l'IA donne la définition suivante d'un système d'IA:
«On entend par "système d'IA", un système automatisé qui est conçu pour fonctionner à différents niveaux d'autonomie et peut faire preuve d'une capacité d'adaptation après son déploiement, et qui, pour des objectifs explicites ou implicites, déduit, à partir des entrées qu'il reçoit, la manière de générer des sorties telles que des prédictions, du contenu, des recommandations ou des décisions qui peuvent influencer les environnements physiques ou virtuels.»

(9) Cette définition comprend sept éléments principaux: (1) un système automatisé; (2) qui est conçu pour fonctionner à différents niveaux d'autonomie; (3) qui peut faire preuve d'une capacité d'adaptation après son déploiement; (4) et qui, pour des objectifs explicites ou implicites; (5) déduit, à partir des entrées qu'il reçoit, la manière de générer des sorties (6) telles que des prédictions, du contenu, des recommandations ou des décisions (7) qui peuvent influencer les environnements physiques ou virtuels.

(10) La définition du terme «système d'IA» adopte une perspective fondée sur le cycle de vie, qui comprend deux phases principales: la phase préalable au déploiement ou phase de «construction» du système et la phase postérieure au déploiement ou phase d'«utilisation» du système[4]. Il n'est pas nécessaire que les sept éléments qui composent cette définition soient présents de manière continue tout au long des deux phases de ce cycle de vie. Au contraire, la définition admet la possibilité que des éléments spécifiques apparaissent dans une phase donnée sans être présents dans les deux phases. L'approche utilisée pour définir un système d'IA reflète la complexité et la diversité des systèmes d'IA et débouche sur une définition qui, en prenant en considération un large éventail de systèmes d'IA, reflète les objectifs du règlement sur l'IA.

1. Système automatisé

(11) Le terme «automatisé» (en anglais *machine-based*, fondé sur des machines) renvoie au fait que les systèmes d'IA sont développés et tournent sur des machines. Il peut être compris comme incluant à la fois les composants matériels et logiciels qui permettent au système d'IA de fonctionner. Les composants matériels sont les éléments physiques de la machine, tels que les unités de traitement, la mémoire, les dispositifs de stockage, les unités réseau et les interfaces d'entrée/sortie, qui fournissent l'infrastructure de calcul. Quant aux composants logiciels, ils englobent le code informatique, les instructions, les programmes, les systèmes d'exploitation et les applications qui déterminent la manière dont le matériel traite les données et exécute les tâches.

[4] Pour une vue d'ensemble des phases d'un système d'IA, voir OCDE (2024), «Explanatory memorandum on the updated OECD definition of an AI system», OECD Artificial Intelligence Papers, No. 8, OECD Publishing, Paris, https://doi.org/10.1787/623da898-en, p. 7.

(12) Tous les systèmes d'IA sont fondés sur des machines, celles-ci étant indispensables à leur fonctionnement, notamment pour l'entraînement de modèles, le traitement de données, la modélisation prédictive et la prise de décision automatisée à grande échelle. L'intégralité du cycle de vie des systèmes d'IA avancés repose sur des machines qui peuvent comprendre de nombreux composants matériels ou logiciels. L'élément «automatisé» (*machine-based*) dans la définition des systèmes d'IA met en évidence le fait que les systèmes d'IA doivent être fondés sur le calcul et le fonctionnement de machines.

(13) Le terme «automatisé» couvre un large éventail de systèmes informatiques. Par exemple, les systèmes d'informatique quantique émergents les plus avancés actuellement, qui se démarquent nettement des systèmes informatiques traditionnels, sont des systèmes fondés sur des machines malgré leurs principes de fonctionnement uniques et l'exploitation de phénomènes de mécanique quantique, tout comme les systèmes biologiques ou organiques, pour autant qu'ils fournissent une capacité de calcul.

2. Autonomie

(14) Le deuxième élément de la définition est le fait qu'un système d'IA est «conçu pour fonctionner à différents niveaux d'autonomie». Le considérant 12 du règlement sur l'IA précise que les termes «différents niveaux d'autonomie» signifient que les systèmes d'IA sont conçus pour fonctionner avec un certain degré d'indépendance dans leur action par rapport à une implication humaine (*human involvement* dans la version anglaise) et de capacités à fonctionner sans intervention humaine.

(15) Les notions d'autonomie et d'inférence vont de pair: la capacité d'inférence d'un système d'IA (c'est-à-dire sa capacité à générer des sorties telles que des prédictions, du contenu, des recommandations ou des décisions qui peuvent influencer les environnements physiques ou virtuels) est une condition essentielle pour qu'il accède à l'autonomie.

(16) La notion d'autonomie ne se conçoit pas sans référence à une «implication humaine» et à une «intervention humaine», donc à l'interaction homme-machine. À l'une des extrémités du spectre des interactions homme-machine possibles, on trouve des systèmes conçus pour exécuter toutes les tâches au moyen de fonctions actionnées manuellement. À l'autre extrémité, il s'agit de systèmes capables de fonctionner sans aucune implication ou intervention humaine, c'est-à-dire de manière entièrement autonome.

(17) La référence à «un certain degré d'indépendance dans leur action» au considérant 12 du règlement sur l'IA exclut les systèmes conçus pour fonctionner exclusivement avec une implication et une intervention humaines manuelles intégrales. L'implication et l'intervention humaines peuvent être soit directes, par exemple au moyen de commandes manuelles, soit indirectes, par exemple au moyen de commandes fondées sur des systèmes automatisés qui permettent aux êtres humains de déléguer ou de superviser le fonctionnement du système.

(18) Par exemple, un système qui a besoin d'éléments en entrée fournis manuellement pour générer une sortie par lui-même est un système qui présente «un certain degré d'indépendance dans son action», car il est conçu avec la capacité de générer une sortie sans recours à une commande manuelle, ou sans que cette sortie soit explicitement et précisément spécifiée par un être humain. Un système expert qui se voit déléguer l'automatisation d'un processus par des êtres humains et qui est capable, en s'appuyant sur des éléments fournis par un être humain, de produire seul une sortie telle qu'une recommandation, est lui aussi un système doté d'«un certain degré d'indépendance dans son action».

(19) La référence, dans la définition du système d'IA figurant à l'article 3, point 1), du règlement sur l'IA, à un «système automatisé qui est conçu pour fonctionner à différents niveaux d'autonomie» souligne la capacité du système à interagir avec son environnement extérieur, plutôt que le choix d'une technique spécifique comme l'apprentissage automatique, ou celui d'une architecture type pour le développement du système.

(20) Par conséquent, le niveau d'autonomie est une condition nécessaire pour déterminer si un système peut être considéré comme un système d'IA. Tout système conçu pour fonctionner avec un certain degré raisonnable d'indépendance dans son action remplit la condition d'autonomie figurant dans la définition d'un système d'IA.

(21) Un système qui a la capacité de fonctionner avec une intervention humaine limitée ou nulle dans des contextes d'utilisation spécifiques, tels que les domaines à haut risque mentionnés aux annexes I et III du règlement sur l'IA, peut, sous certaines conditions, générer des risques potentiels supplémentaires et justifier une réflexion sur des mesures de contrôle humain. Le niveau d'autonomie est un aspect important à prendre en compte par un fournisseur lorsqu'il conçoit, par exemple, les mesures de contrôle humain ou d'atténuation des risques du système, dans le contexte de sa destination.

3. Capacité d'adaptation

(22) Le troisième élément de la définition figurant à l'article 3, point 1), du règlement sur l'IA est que le système d'IA «peut faire preuve d'une capacité d'adaptation après son déploiement». Les concepts d'autonomie et de capacité d'adaptation sont distincts mais étroitement liés. On les associe souvent dans les discussions, mais ils représentent des dimensions différentes de la fonctionnalité d'un système d'IA. Le considérant 12 du règlement sur l'IA précise que la «capacité d'adaptation» fait référence aux capacités d'autoapprentissage, qui permettent un changement de comportement du système au cours de son utilisation. Le nouveau comportement du système, une fois adapté, peut produire, avec les mêmes éléments en entrée, des résultats différents de ceux du système précédent.

(23) L'utilisation du terme «peut» en rapport avec cet élément de la définition indique que pour être considéré comme un système d'IA, un système **peut**, mais **ne doit pas nécessairement**, posséder des capacités d'adaptation ou d'autoapprentissage. Par

conséquent, la capacité d'un système à apprendre de manière automatique, à découvrir de nouveaux modèles, ou à identifier des relations entre les données au-delà des aspects sur lesquels a porté son entraînement initial est une condition facultative, donc non déterminante, pour déterminer si ce système peut être considéré comme un système d'IA.

4. Objectifs d'un système d'IA

(24) Les objectifs d'un système d'IA constituent le quatrième élément de la définition. Les systèmes d'IA sont conçus pour fonctionner par référence à un ou plusieurs objectifs. Les objectifs du système peuvent être définis explicitement ou implicitement. Les **objectifs explicites** renvoient à des buts clairement énoncés que le développeur encode directement dans le système. Il peut s'agir par exemple de l'optimisation d'une fonction de coût, d'une probabilité ou d'une récompense cumulée. Les **objectifs implicites** renvoient à des buts qui ne sont pas explicitement énoncés mais qui peuvent être déduits du comportement ou d'hypothèses qui sous-tendent le système. Ces objectifs peuvent découler des données d'entraînement ou de l'interaction du système d'IA avec son environnement.

(25) Le considérant 12 du règlement sur l'IA précise que «[l]es objectifs du système d'IA peuvent être différents de la destination du système d'IA dans un contexte spécifique». Les objectifs d'un système d'IA sont définis de manière interne au système, par référence aux buts des tâches à exécuter et à leurs résultats. Par exemple, dans une entreprise, un assistant virtuel fondé sur l'IA peut avoir pour **objectif** de répondre aux questions des utilisateurs sur un ensemble de documents avec une grande précision et un faible taux d'erreur. La **destination**, en revanche, est orientée vers l'extérieur et inclut le contexte dans lequel sera déployé le système et la manière dont il doit être exploité. En effet, à l'article 3, point 12), du règlement sur l'IA, la destination d'un système d'IA est définie comme l'«utilisation à laquelle un système d'IA est destiné par le fournisseur». Par exemple, dans le cas d'un assistant virtuel fondé sur l'IA en entreprise, la destination pourrait être d'aider un département donné de l'entreprise à exécuter certaines tâches. Pour y parvenir, il faudra peut-être que les documents à utiliser par l'assistant virtuel respectent certaines conditions (de longueur, de formatage, etc.) et que les questions de l'utilisateur se limitent au domaine dans lequel le système est censé fonctionner. Cette destination est atteinte non seulement grâce au fonctionnement interne du système pour réaliser ses objectifs, mais aussi grâce à d'autres facteurs tels que l'intégration du système dans le flux plus large du service à la clientèle, les données utilisées par le système ou la notice d'utilisation.

5. Inférences sur la manière de générer des sorties en utilisant des techniques d'IA

(26) Le cinquième élément qui caractérise un système d'IA est qu'il doit être en mesure de déduire, à partir des entrées qu'il reçoit, la manière de générer des sorties. Le considérant 12 du règlement sur l'IA précise qu'«[u]ne caractéristique essentielle des systèmes d'IA est leur capacité d'inférence». Comme expliqué dans la suite de ce

considérant, les systèmes d'IA sont à distinguer «des systèmes logiciels ou des approches de programmation traditionnels plus simples, et ne [devraient] pas couvrir les systèmes fondés sur les règles définies uniquement par les personnes physiques pour exécuter automatiquement des opérations». Cette capacité d'inférence est par conséquent une condition clé indispensable pour distinguer les systèmes d'IA d'autres types de systèmes.

(27) Le considérant 12 explique aussi que «[c]ette capacité d'inférence concerne le processus consistant à générer des sorties telles que des prédictions, du contenu, des recommandations ou des décisions, qui peuvent influencer l'environnement physique ou virtuel, et la capacité des systèmes d'IA à inférer des modèles ou des algorithmes, ou les deux, à partir d'entrées ou de données.» Cette interprétation de la notion d'inférence ne contredit pas celle de la norme ISO/IEC 22989, qui définit l'inférence comme un «raisonnement qui consiste à dériver une conclusion d'une prémisse connue»; cette norme comprend une note spécifique à l'IA: «[e]n IA, une prémisse est un fait, une règle, un modèle, une caractéristique ou des données brutes»[5].

(28) Le «processus consistant à générer des sorties telles que des prédictions, du contenu, des recommandations ou des décisions, qui peuvent influencer l'environnement physique ou virtuel» fait référence à la capacité du système d'IA, principalement en «phase d'utilisation», à générer des sorties à partir d'entrées. La «capacité des systèmes d'IA à inférer des modèles ou des algorithmes, ou les deux, à partir d'entrées ou de données» vise principalement, mais pas exclusivement, la «phase de construction» du système et met en évidence la pertinence des techniques utilisées pour la construction d'un système.

(29) Les termes «déduire... la manière de», utilisés à l'article 3, point 1), et clarifiés au considérant 12 du règlement sur l'IA, sont plus larges et ne se limitent pas uniquement à une interprétation étroite de la notion d'inférence/déduction en tant que capacité d'un système à inférer des sorties à partir d'entrées données, et donc à déduire un résultat. Par conséquent, la formulation utilisée à l'article 3, point 1), du règlement sur l'IA, à savoir «déduire... la manière de générer des sorties» doit se comprendre comme faisant référence à la phase de construction, au cours de laquelle un système déduit des sorties en recourant à des techniques d'IA rendant l'inférence possible.

5.1. Techniques d'IA permettant l'inférence

(30) En se concentrant spécifiquement sur la phase de construction du système d'IA, le considérant 12 du règlement sur l'IA précise en outre que «[l]es techniques permettant l'inférence lors de la construction d'un système d'IA comprennent des approches d'apprentissage automatique qui apprennent à partir des données la manière d'atteindre certains objectifs, et des approches fondées sur la logique et les connaissances qui font

[5] ISO/IEC 22989:2022, Technologies de l'information – Intelligence artificielle – concepts et terminologie relatifs à l'intelligence artificielle.

des inférences à partir des connaissances encodées ou de la représentation symbolique de la tâche à résoudre». Ces techniques sont en fait des «techniques d'IA».

(31) Cette précision souligne explicitement que la notion d'«inférence» doit se comprendre dans une acception plus large, comme englobant la «phase de construction» du système d'IA. Le considérant 12 du règlement sur l'IA donne ensuite des orientations supplémentaires à propos des techniques qui rendent possible cette capacité d'un système d'IA à inférer la manière de générer des sorties. Par conséquent, les techniques qui peuvent être utilisées pour permettre l'inférence comprennent «des approches d'apprentissage automatique qui apprennent à partir des données la manière d'atteindre certains objectifs, et des approches fondées sur la logique et les connaissances qui font des inférences à partir des connaissances encodées ou de la représentation symbolique de la tâche à résoudre».

(32) La première catégorie de techniques d'IA mentionnée au considérant 12 du règlement sur l'IA est celle des «approches d'apprentissage automatique qui apprennent à partir des données la manière d'atteindre certains objectifs». Cette catégorie comprend un large éventail d'approches permettant à un système d'«apprendre», comme l'apprentissage supervisé, l'apprentissage non supervisé, l'apprentissage autosupervisé et l'apprentissage par renforcement.

(33) Dans le cas de l'**apprentissage supervisé**, le système d'IA apprend à partir d'annotations (données étiquetées), l'annotation consistant à associer les données d'entrée avec la sortie correcte. Le système utilise ces annotations pour établir une cartographie depuis les entrées jusqu'aux sorties, puis généralise celle-ci à de nouvelles données que le système n'a encore jamais vues. Un système de détection de spam fondé sur l'IA est un exemple de système d'apprentissage supervisé. Pendant sa phase de construction, le système est entraîné sur un ensemble de données contenant des courriels que des humains ont étiquetés comme «spam» ou «non-spam», afin d'apprendre des modèles à partir des caractéristiques des courriels étiquetés. Une fois entraîné et mis en service, le système peut analyser les nouveaux courriels et les classer comme spam ou non-spam sur la base des modèles qu'il a appris à partir des données étiquetées.

(34) D'autres exemples de systèmes d'IA fondés sur l'apprentissage supervisé sont notamment les systèmes de classification des images entraînés sur un ensemble de données d'images, chaque image étant associée à un ensemble d'étiquettes (objet – voiture, par exemple); les systèmes de diagnostic pour dispositifs médicaux entraînés sur des données d'imagerie médicale étiquetées par des experts humains; ou encore des systèmes de détection de fraude entraînés sur des données de transaction étiquetées.

(35) Dans le cas de l'**apprentissage non supervisé**, l'apprentissage du système d'IA se fait à partir de données qui n'ont pas été étiquetées. Le modèle est entraîné sur des données sans aucune étiquette ni sortie prédéfinies. À l'aide de différentes techniques, telles que le regroupement, la réduction de la dimensionnalité, l'apprentissage par règles d'association, la détection d'anomalies ou les modèles génératifs, le système est entraîné pour découvrir des modèles, des structures ou des relations parmi les données sans que

lui soient fournies de lignes directrices explicites sur le résultat à atteindre. Les systèmes d'IA utilisés par les entreprises pharmaceutiques dans la recherche de nouveaux médicaments sont un exemple d'apprentissage non supervisé. Ces systèmes pratiquent l'apprentissage non supervisé (par exemple les techniques de regroupement ou de détection d'anomalies) pour regrouper les composés chimiques et prédire de nouveaux traitements potentiels pour une maladie en fonction de leurs similitudes avec des médicaments existants.

(36) L'**apprentissage autosupervisé** est une sous-catégorie de l'apprentissage non supervisé, dans le cadre de laquelle le système d'IA apprend à partir de données qui n'ont pas été étiquetées sous supervision, en utilisant les données elles-mêmes pour créer ses propres étiquettes ou objectifs. Les systèmes d'IA fondés sur l'apprentissage autosupervisé utilisent diverses techniques, telles que les auto-encodeurs, les réseaux antagonistes génératifs ou l'apprentissage contrastif. Un système de reconnaissance d'images qui apprend à reconnaître des objets en prédisant les pixels manquants dans une image est un exemple de système d'IA fondé sur l'apprentissage autosupervisé. D'autres exemples sont notamment les modèles de langage qui apprennent à prédire quel sera le jeton textuel (ou *token*) suivant dans une phrase, ou les systèmes de reconnaissance vocale qui apprennent à reconnaître le langage parlé en prédisant le prochain élément acoustique dans un signal audio.

(37) Les systèmes d'IA fondés sur l'**apprentissage par renforcement** apprennent à partir de données issues de leur propre expérience grâce à une fonction de «récompense». Contrairement aux systèmes d'IA qui apprennent à partir de données étiquetées (apprentissage supervisé) ou à partir de modèles (apprentissage non supervisé), les systèmes d'IA fondés sur l'apprentissage par renforcement apprennent à partir de l'expérience. Le système ne reçoit pas d'étiquettes explicites, mais apprend par un processus itératif d'essais et d'erreurs, en affinant sa stratégie sur la base d'une rétroaction fournie par l'environnement. Un bras robotisé s'appuyant sur l'IA, capable de réaliser des tâches telles que saisir des objets, est un exemple de système d'IA fondé sur l'apprentissage par renforcement. Ce type d'apprentissage peut aussi servir, par exemple, à optimiser les recommandations de contenu personnalisées dans les moteurs de recherche, ou les performances des véhicules autonomes.

(38) L'**apprentissage profond** est un sous-ensemble de l'apprentissage automatique qui utilise une architecture comportant différentes couches (réseaux neuronaux) pour l'apprentissage des caractéristiques ou représentations. Les systèmes d'IA fondés sur l'apprentissage profond peuvent apprendre des caractéristiques automatiquement à partir de données brutes, ce qui supprime la nécessité d'une ingénierie manuelle des caractéristiques. En raison du nombre de couches de neurones et du nombre de paramètres, les systèmes d'IA fondés sur l'apprentissage profond nécessitent généralement de grandes quantités de données pour leur entraînement, mais ils peuvent apprendre à reconnaître des modèles et à effectuer des prédictions avec une grande précision lorsqu'ils disposent de données suffisantes. Les systèmes d'IA fondés sur l'apprentissage profond sont largement utilisés et leur technologie est à l'origine de nombreuses avancées récentes dans le domaine de l'IA.

(39) À côté des différentes approches d'apprentissage automatique examinées ci-dessus, la deuxième catégorie de techniques mentionnée au considérant 12 du règlement sur l'IA est celle «des **approches fondées sur la logique et les connaissances** qui font des inférences à partir des connaissances encodées ou de la représentation symbolique de la tâche à résoudre». Au lieu d'apprendre à partir de données, ces systèmes d'IA apprennent à partir de connaissances, y compris des règles, des faits et des relations encodés par des experts humains. Sur la base des connaissances encodées par des experts humains, ces systèmes peuvent «raisonner» au moyen de moteurs déductifs ou inductifs ou en recourant à des opérations telles que le tri, la recherche, la mise en correspondance, le chaînage. En utilisant l'inférence logique pour tirer des conclusions, ces systèmes appliquent une logique formelle, des règles prédéfinies ou des ontologies à de nouvelles situations. Les approches fondées sur la logique et les connaissances comprennent, par exemple, la représentation des connaissances, la programmation (logique) inductive, les bases de connaissances, les moteurs d'inférence et déductifs, le raisonnement (symbolique), les systèmes experts et les méthodes de recherche et d'optimisation. Par exemple, les modèles classiques de traitement du langage fondés sur les connaissances grammaticales et la sémantique logique s'appuient sur la structure du langage, en identifiant les composants syntaxiques et grammaticaux des phrases pour extraire le sens d'un texte donné. Un autre exemple notable de systèmes d'IA reposant sur des approches logiques et fondées sur les connaissances est celui des systèmes experts des premières générations destinés au diagnostic médical, développés en encodant les connaissances d'un panel d'experts médicaux et destinés à tirer des conclusions à partir d'un ensemble de symptômes que présente un patient donné.

5.2. Systèmes qui ne relèvent pas de la définition du «système d'IA»

(40) Le considérant 12 explique aussi que la définition des systèmes d'IA devrait distinguer ces derniers des «systèmes logiciels ou des approches de programmation traditionnels plus simples, et ne devrait pas couvrir les systèmes fondés sur les règles définies uniquement par les personnes physiques pour exécuter automatiquement des opérations».

(41) Certains systèmes ont une capacité d'inférence, selon une interprétation restrictive, mais peuvent néanmoins ne pas relever de la définition du système d'IA, en raison de leur capacité limitée à analyser des modèles et à ajuster leurs sorties de manière autonome. On peut classer dans cette catégorie les systèmes suivants:

Systèmes destinés à améliorer l'optimisation mathématique

(42) Les systèmes utilisés pour améliorer l'optimisation mathématique ou pour accélérer et approximer les méthodes d'optimisation traditionnelles et bien établies, telles que les méthodes de régression linéaire ou logistique, ne relèvent pas de la définition du système d'IA. En effet, si ces modèles possèdent une capacité d'inférence, celle-ci ne va pas au-delà du traitement de données de base. Un indice qu'un système ne va pas au-delà du traitement de données de base peut être le fait qu'il a été utilisé de manière consolidée

pendant de nombreuses années[6]. Il s'agit, par exemple, de modèles fondés sur l'apprentissage automatique qui approximent des fonctions ou des paramètres dans les problèmes d'optimisation tout en maintenant les performances. Ces systèmes visent à améliorer l'efficacité des algorithmes d'optimisation utilisés dans les problèmes de calcul. Ils contribuent par exemple à accélérer les tâches d'optimisation en fournissant des approximations, des méthodes heuristiques ou des stratégies de recherche.

(43) Par exemple, des systèmes fondés sur la physique peuvent recourir à des techniques d'apprentissage automatique pour améliorer les performances informatiques, accélérer les simulations traditionnelles fondées sur la physique ou estimer des paramètres, qui sont ensuite intégrés dans les modèles physiques établis. Ces systèmes ne relèveraient pas de la définition du système d'IA Dans cet exemple, les modèles d'apprentissage automatique approximent des processus atmosphériques complexes, tels que la microphysique des nuages ou les turbulences, permettant d'établir des prévisions plus rapidement et avec une efficience de calcul supérieure.

(44) Un autre exemple de système qui ne relève pas non plus de la définition du système d'IA serait celui d'un système destiné à optimiser l'attribution de la bande passante et la gestion des ressources dans le domaine des télécommunications par satellite. Dans ce domaine, la gestion des demandes de trafic réseau en temps réel peut se heurter à des difficultés lorsqu'on recourt aux méthodes traditionnelles d'optimisation, notamment pour s'adapter aux variations de la demande des utilisateurs d'une région à l'autre. On peut faire appel par exemple à des modèles d'apprentissage automatique pour prédire le trafic réseau et optimiser l'allocation de ressources telles que la puissance et la largeur de bande aux répéteurs satellites, avec des performances similaires à celles des méthodes établies sur le terrain.

(45) Bien que ces systèmes puissent comporter des mécanismes d'autoajustement automatique, ceux-ci visent à optimiser le fonctionnement des systèmes en améliorant leurs performances informatiques plutôt que, par exemple, à permettre d'ajuster de manière intelligente leurs modèles de prise de décision. Dans ces conditions, ces systèmes sont exclus de la définition du système d'IA.

Traitement de données de base

(46) La notion de système de traitement de données de base renvoie à un système qui suit des instructions ou effectue des opérations prédéfinies, explicites. De tels systèmes sont développés et déployés pour exécuter des tâches sur la base d'entrées manuelles ou de règles, sans aucun «apprentissage, raisonnement ou modélisation» à quelque stade du cycle de vie du système que ce soit. Ils fonctionnent sur la base de règles fixes programmées par l'homme, sans recourir à des techniques d'IA telles que l'apprentissage automatique ou l'inférence fondée sur la logique, pour générer des sorties. Ces systèmes de traitement de données de base comprennent, par exemple, les systèmes de gestion de bases de données utilisés pour trier ou filtrer des données sur la base de critères spécifiques (par exemple, «trouver tous les clients qui ont acheté un produit spécifique

[6] En tout état de cause, les systèmes qui ont été mis sur le marché ou mis en service avant le 2 août 2026 bénéficient de la clause d'antériorité prévue à l'article 111, paragraphe 2, du règlement sur l'IA.

au cours du dernier mois»), les tableurs standard qui n'intègrent pas de fonctionnalités fondées sur l'IA, et les logiciels qui calculent la moyenne d'une population à partir d'une enquête que l'on exploite par la suite dans un contexte général.

(47) De même, les systèmes uniquement destinés à l'analyse descriptive, à la vérification d'hypothèses et à la visualisation ne relèvent pas de la définition d'un système d'IA. Par exemple, dans un logiciel de visualisation de rapports de vente, l'emploi de méthodes statistiques permet de créer un tableau de bord des ventes qui présente les ventes totales, les ventes moyennes par région et l'évolution des ventes dans le temps. À l'aide de méthodes statistiques, ces données peuvent être résumées et visualisées dans des diagrammes et des graphiques. Toutefois, le système ne fournit pas de recommandations sur la manière d'améliorer les ventes ni sur les produits qu'il faudrait promouvoir. Un autre exemple est celui d'un système logiciel qui applique des techniques statistiques aux sondages d'opinion ou aux données d'enquête afin de déterminer leur validité, leur fiabilité, leur corrélation et leur signification statistique. Ces systèmes ne pratiquent pas «l'apprentissage, le raisonnement ou la modélisation», ils présentent simplement des données de manière informative.

Systèmes fondés sur l'heuristique classique

(48) L'heuristique classique offre des techniques de résolution de problèmes qui s'appuient sur des méthodes fondées sur l'expérience pour trouver des solutions approximatives de manière efficiente. Les techniques heuristiques sont couramment utilisées dans les situations de programmation où il n'est pas réalisable de trouver une solution exacte en raison de contraintes de temps ou de ressources. Les systèmes heuristiques classiques reposent généralement sur des approches fondées sur des règles, la reconnaissance de modèles ou des stratégies d'essai et d'erreur plutôt que sur un apprentissage fondé sur les données. Contrairement aux systèmes modernes d'apprentissage automatique, qui adaptent leurs modèles en fonction des relations entrées-sorties, les systèmes heuristiques classiques appliquent des règles ou des algorithmes prédéfinis pour en tirer des solutions. Par exemple, un programme d'échecs utilisant un algorithme MiniMax doté de fonctions d'évaluation heuristiques peut évaluer les positions sur l'échiquier sans nécessiter d'apprentissage préalable à partir de données. Bien qu'efficaces dans de nombreuses applications, les méthodes heuristiques peuvent manquer d'adaptabilité et de capacité de généralisation par rapport aux systèmes d'IA qui apprennent à partir de l'expérience.

Systèmes de prédiction simples

(49) Tous les systèmes automatisés dont les performances peuvent être obtenues au moyen d'une règle d'apprentissage statistique de base, qu'on pourrait techniquement considérer comme reposant sur une approche d'apprentissage automatique, ne relèvent pas du champ d'application de la définition du système d'IA, en raison de leurs performances.

(50) Dans le domaine des prévisions financières (benchmarking de base), de tels systèmes automatisés peuvent être utilisés pour prédire les cours futurs d'actions en utilisant un estimateur avec la stratégie «moyenne» pour établir une prédiction de référence (par

exemple, en prédisant toujours le cours moyen historique). Ces méthodes de benchmarking de base permettent d'évaluer si des modèles d'apprentissage automatique plus avancés peuvent apporter une valeur ajoutée. Un autre exemple est l'utilisation de la température moyenne de la semaine écoulée pour prédire la température du lendemain. Ce système de référence estime uniquement des moyennes mais n'atteint pas la performance de systèmes de prévision de séries chronologiques plus complexes, qui nécessiteraient des modèles plus sophistiqués.

(51) Les systèmes d'estimation statique (par exemple un système d'estimation du temps de réponse d'un service d'assistance à la clientèle, reposant sur une estimation statique pour prédire le temps de résolution moyen à partir des données antérieures et d'indicateurs triviaux tels que la prévision de la demande dans un magasin donné pour prédire le nombre d'exemplaires d'un produit qui seront vendus chaque jour dans le magasin) sont d'autres exemples qui contribuent à établir une référence ou un benchmark, par exemple en prédisant une moyenne.

6. Sorties susceptibles d'influencer les environnements physiques ou virtuels

(52) Le sixième élément de la définition du système d'IA figurant à l'article 3, point 1), du règlement sur l'IA est que le système déduit «la manière de générer des sorties telles que des prédictions, du contenu, des recommandations ou des décisions qui peuvent influencer les environnements physiques ou virtuels». L'aptitude d'un système à générer des sorties, telles que des prédictions, du contenu et des recommandations, à partir d'entrées qu'il reçoit et en utilisant l'apprentissage automatique et des approches logiques et fondées sur les connaissances, est fondamentale pour caractériser ce que font les systèmes d'IA et ce qui les distingue d'autres formes de logiciels. La capacité de générer des sorties et le type de sortie que le système peut générer sont essentiels pour comprendre la fonctionnalité et l'impact d'un système d'IA.

(53) Les sorties des systèmes d'IA peuvent être classées dans l'une des quatre grandes catégories énumérées à l'article 3, point 1), du règlement sur l'IA: prédictions, contenu, recommandations et décisions. Chaque catégorie diffère par le niveau d'implication humaine qui lui correspond.

(54) Les **prédictions** sont l'une des sorties les plus courantes produites par un système d'IA et qui nécessitent l'implication humaine la plus faible. Une prédiction est l'estimation d'une valeur inconnue (sortie) à partir de valeurs connues fournies au système (entrée). On utilise des systèmes logiciels depuis des décennies pour générer des prédictions. Les systèmes d'IA utilisant l'apprentissage automatique sont capables de générer des prédictions qui révèlent des modèles complexes au sein des données et font des prédictions précises dans des environnements très dynamiques et complexes.

(55) Par exemple, les systèmes d'IA déployés dans les voitures autonomes sont conçus pour faire des prédictions en temps réel dans un environnement extrêmement complexe et dynamique, caractérisé par de multiples types d'agents et d'interactions, ainsi que par un nombre pratiquement infini de situations possibles, et pour prendre des décisions afin d'adapter leur comportement en conséquence. Des systèmes non-IA, généralement

fondés sur des données historiques, des données scientifiques ou des règles prédéfinies, tels que certains systèmes experts non-IA équipant des dispositifs médicaux, sont incapables de faire face à un tel degré de complexité. Dans un autre domaine, les systèmes d'IA utilisés dans la gestion de la consommation d'énergie sont conçus pour estimer la consommation d'énergie en analysant les données provenant de compteurs intelligents, de prévisions météorologiques et de modèles comportementaux des consommateurs. En s'appuyant sur des approches d'apprentissage automatique, ces systèmes d'IA sont conçus pour découvrir des corrélations complexes entre les variables précitées afin d'augmenter la précision des prédictions relatives à la consommation d'énergie.

(56) La notion de **contenu** désigne la génération de matériau nouveau par un système d'IA. Il peut s'agir de textes, d'images, de vidéos, de musique et d'autres formes de sortie. Un nombre croissant de systèmes d'IA utilisent des modèles d'apprentissage automatique (fondés par exemple sur les technologies Generative Pre-trained Transformer ou GPT) pour générer du contenu. Bien que, d'un point de vue technique, on puisse assimiler le contenu, en tant que catégorie de sortie, à une séquence de «prédictions» ou de «décisions», en raison de la prévalence de ces sorties dans les systèmes d'IA générative, le contenu est présenté, au considérant 12 du règlement sur l'IA, en tant que catégorie de sortie distincte.

(57) Le terme **recommandation** fait référence à des suggestions d'actions, de produits ou de services spécifiques adressées aux utilisateurs sur la base de leurs préférences, de leurs comportements ou d'autres données en entrée. Comme c'est le cas pour les prédictions, les systèmes conçus pour générer des recommandations peuvent être fondés sur l'IA ou non fondés sur l'IA. Les systèmes de recommandation fondés sur l'IA, par exemple, peuvent exploiter des données à grande échelle, s'adapter au comportement des utilisateurs en temps réel, fournir des recommandations hautement personnalisées et se développer de manière efficiente à mesure que le volume de données augmente – des fonctionnalités que les systèmes non-IA qui reposent sur des mécanismes statiques et fondés sur des règles et sur des données limitées possèdent rarement. Dans certains cas, les recommandations correspondent à des décisions potentielles, par exemple dans un système de recrutement, où les candidats recommandés seront ensuite évalués par des humains. Si ces recommandations sont appliquées automatiquement, elles deviennent des décisions.

(58) Les **décisions** correspondent à des conclusions formulées ou à des choix effectués par un système. Un système d'IA qui génère des décisions automatise des processus traditionnellement soumis à un jugement humain. Un tel système implique un processus entièrement automatisé produisant un résultat donné dans l'environnement qui l'entoure sans intervention humaine.

(59) En résumé, les systèmes d'IA, notamment ceux qui reposent sur des approches d'apprentissage automatique et sur des approches fondées sur la logique et les connaissances, diffèrent des systèmes non-IA dans leur capacité à générer des sorties telles que des prédictions, du contenu, des recommandations et des décisions, par le fait qu'ils peuvent gérer des relations et des modèles complexes au sein des données. Les

systèmes d'IA peuvent généralement produire des résultats plus nuancés que d'autres systèmes, par exemple en tirant parti de modèles appris au cours de leur entraînement ou en utilisant des règles définies par des experts pour prendre des décisions, en offrant un raisonnement plus élaboré dans des environnements structurés.

7. Interaction avec l'environnement

(60) Le septième élément de la définition d'un système d'IA est le fait que les sorties de ce système «**peuvent influencer les environnements physiques ou virtuels**». Cet élément souligne le fait que les systèmes d'IA ne sont pas passifs mais influent concrètement sur les environnements dans lesquels ils sont déployés. La référence aux «environnements physiques ou virtuels» indique que l'influence d'un système d'IA peut s'exercer aussi bien sur des objets physiques tangibles (par exemple, un bras robotisé) que sur des environnements virtuels, notamment des espaces numériques, des flux de données et des écosystèmes logiciels.

III. Observations finales

(61) La définition d'un système d'IA englobe un large éventail de systèmes. Pour déterminer si un système logiciel donné est un système d'IA, il convient de se fonder sur son architecture et ses fonctionnalités spécifiques et de prendre en considération les sept éléments de la définition énoncée à l'article 3, point 1), du règlement sur l'IA.

(62) Déterminer automatiquement les systèmes qui relèvent ou non de la définition d'un système d'IA, ou en dresser des listes exhaustives, est impossible.

(63) Seuls certains systèmes d'IA sont soumis à des obligations réglementaires et à une surveillance au titre du règlement sur l'IA. L'approche fondée sur les risques adoptée dans le règlement sur l'IA implique que seuls les systèmes présentant les risques les plus élevés pour les droits et libertés fondamentaux seront soumis aux interdictions énoncées à l'article 5, au régime réglementaire applicable aux systèmes d'IA à haut risque couverts par l'article 6 et aux exigences de transparence (qui s'appliquent à un nombre limité de systèmes d'IA prédéfinis décrits à l'article 50 du règlement). La grande majorité des systèmes, même s'ils sont considérés comme des systèmes d'IA au sens de l'article 3, point 1), du règlement sur l'IA, ne seront soumis à aucune exigence réglementaire au titre du règlement.

(64) Le règlement sur l'IA s'applique également aux modèles d'IA à usage général, qui sont régis par le chapitre V du règlement. L'analyse des différences entre les systèmes d'IA et les modèles d'IA à usage général sort du champ des présentes lignes directrices.

AMPLIATION CERTIFIÉE CONFORME
Pour la Secrétaire générale

Martine DEPREZ
Directrice
Prise de décision & Collégialité
COMMISSION EUROPÉENNE