La Commission européenne définit ce qu'elle entend par système d'IA

Michael Montavon, le 17 mars 2025

La Commission européenne a publié en février 2025 ses lignes directrices sur la définition des systèmes d'IA. S'articulant autour de sept caractéristiques plus ou moins précises, elles visent à distinguer les systèmes d'IA des logiciels traditionnels.

Selon l'article 3, ch. 1 du Règlement (UE) 2024/1689 sur l'intelligence artificielle (Règlement sur l'IA), un système d'IA correspond à « 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 ».

Le 6 février 2025, la Commission européenne a publié ses lignes directrices sur la définition des systèmes d'IA afin de faciliter la compréhension et l'application du Règlement sur l'IA. En publiant ces lignes directrices, la Commission vise à aider les fournisseurs, mandataires, déployeurs, importateurs, distributeurs ou encore fabricants de systèmes d'IA, à déterminer si un logiciel entre ou non dans le champ d'application du Règlement européen.

Sept caractéristiques pour une définition

Selon la Commission européenne, la définition d'un système d'IA au sens du Règlement européen est structurée autour de sept caractéristiques, chacune contribuant à préciser la portée du texte et les implications réglementaires qui en découlent.

En vertu de la première caractéristique, un système d'IA doit être basé sur une machine, c'est-à-dire être fondé sur un ensemble de composants matériels (*hardware*) et logiciels (*software*) qui permettent son fonctionnement. Un tel système repose sur une infrastructure de calcul, incluant processeurs, mémoire et dispositifs de stockage, ainsi que sur des algorithmes ou modèles computationnels qui structurent le traitement de l'information. La Commission européenne précise que cette caractéristique inclut non seulement les systèmes émergents d'informatique quantique, mais aussi les systèmes biologiques ou organiques

pour autant qu'ils fournissent des capacités de calcul. Cette approche a pour ambition d'élargir la notion de système d'IA au-delà des dispositifs informatiques traditionnels en silicium, en intégrant des structures biologiques, telles que des réseaux de neurones cultivés en laboratoire (bio-ordinateurs).

La deuxième caractéristique mise en évidence est l'autonomie, c'est-à-dire la capacité d'un système d'IA de fonctionner avec un degré d'indépendance plus ou moins élevé (« *some degree of independance of action* »). Selon la Commission européenne, cette caractéristique n'implique pas nécessairement une absence totale de contrôle et/ou d'intervention humaine. En revanche, elle exclut les systèmes qui sont incapables d'opérer en l'absence d'interventions ou d'instructions humaines claires et constantes. La caractéristique de l'autonomie se manifeste par exemple dans des systèmes, tels que les assistants vocaux, qui analysent des requêtes et exécutent des actions sans intervention explicite de l'utilisateur à chaque étape du traitement. À l'inverse, un convertisseur de devises, qui applique un taux de change fixe à une valeur saisie manuellement, ne fait preuve d'aucune autonomie et ne correspond par conséquent pas à un système d'IA. Entre les deux extrémités, il existe une part d'ombre importante que les lignes directrices ne permettent pas de dissiper.

La troisième caractéristique est la capacité d'adaptation après déploiement des systèmes d'IA. Cette caractéristique fait référence aux capacités d'auto-apprentissage des systèmes d'IA, leur permettant de modifier leur comportement durant leur cycle de vie, en intégrant de nouvelles données et en affinant leurs prédictions ou en améliorant leurs performances. Se fondant sur le texte de l'article 3, ch. 1 du Règlement selon lequel un système d'IA « *peut faire preuve d'une capacité d'adaptation après son déploiement* », la Commission européenne relève qu'il s'agit certes d'une caractéristique fréquente des systèmes d'IA, mais qu'elle n'est pas indispensable. Là aussi, on peut regretter le manque d'apport tangible des lignes directrices de la Commission européenne. À titre d'exemple de capacité d'adaptation, on peut citer un moteur de recommandation exploitant l'historique de visionnage d'un utilisateur pour ajuster ses suggestions (p. ex. Netflix et Spotify). À l'inverse, un système de navigation GPS qui se contente d'indiquer un itinéraire sans réévaluation dynamique en fonction du trafic ne présente pas de capacité d'adaptation.

La quatrième caractéristique de la définition d'un système d'IA repose sur la présence d'objectifs explicites ou implicites. Certains systèmes d'IA sont développés avec des finalités précises, définies dès leur conception, telles que l'optimisation d'un indicateur de performance ou la minimisation d'une erreur. Une fois l'objectif fixé, le système d'IA va ensuite chercher le meilleur moyen de l'atteindre. D'autres système d'IA fonctionnent selon des

objectifs qui ne sont pas explicitement formulés mais qui émergent de l'analyse des données ou de l'interaction du système avec son environnement. À titre d'exemple, un système de recommandation de musique ou de films qui propose à l'utilisateur un contenu selon ses goûts peut poursuivre tant des objectifs explicites qu'implicites en fonction de son mode opérationnel. À l'inverse, un système qui propose un contenu de manière totalement aléatoire ne poursuit aucun objectif et ne correspond pas à la définition d'un système d'IA. Selon la Commission européenne, les objectifs intrinsèques d'un système d'IA doivent être distingués de l'utilisation pour laquelle il est destiné par son utilisateur.

La cinquième caractéristique réside dans la capacité d'inférence d'un système d'IA. Il s'agit d'une caractéristique essentielle qui distingue les systèmes d'IA des des logiciels traditionnels. Un système d'IA doit être capable de produire par lui-même des résultats (*outputs*) à partir des informations qu'il reçoit (*inputs*). Cette capacité repose sur différentes approches computationnelles, incluant l'apprentissage supervisé, l'apprentissage non supervisé, l'apprentissage auto-supervisé, l'apprentissage par renforcement, le recours aux réseaux de neurones ou encore les approches fondées sur la logique et la connaissance, qui réalisent des inférences à partir d'une connaissance encodée ou d'une représentation symbolique de la tâche à accomplir. Cette capacité d'inférence peut être illustrée par un détecteur de fraude bancaire qui analyse en temps réel des transactions et émet des alertes en cas d'activité suspecte. À l'inverse, un tableau de bord affichant des tendances de consommation sur la base de règles préétablies ne saurait être assimilé à un système d'IA, car il ne procède à aucune analyse autonome des données.

La sixième caractéristique mise en évidence est l'influence que les systèmes d'IA exercent sur l'environnement physique ou virtuel en modifiant un état de fait grâce à un certain *output*. On distingue quatre catégories d'*outputs* en fonction de l'influence qu'ils vont exercer :

- les prédictions qui constituent des valeurs inconnues générées à partir de données existantes pour prédire un évènement futur (p. ex. la météo ou la consommation énergétique);
- les contenus qui correspondent à la création originale de textes, d'images, de vidéos ou de musiques ;
- les recommandations qui correspondent à des actions qui sont suggérées à un être humain mais qui n'ont pas la capacité de s'auto-réaliser sans validation humaine. La Commission européenne précise que contrairement aux systèmes classiques basés sur

des règles fixes, les systèmes d'IA peuvent s'adapter en temps réel et offrir des recommandations personnalisées ;

- les décisions qui forment l'*output* le plus impactant qu'un système d'IA peut produire. Elles désignent les choix effectués de manière autonome par le système d'IA et qui vont modifier une situation de fait ou de droit via un processus entièrement automatisé, sans intervention humaine. La Commission européenne relève que si des recommandations sont appliquées sans validation humaine, elles peuvent devenir des décisions.

La septième et dernière caractéristique d'un système d'IA est sa capacité d'interagir avec son environnement, que ce soit dans le monde physique ou numérique. Un système d'IA n'est pas un outil passif qui se contente de traiter de l'information. Il s'agit d'un dispositif capable d'interagir dynamiquement avec son environnement. Par exemple, un programme de maintenance prédictive qui surveille des équipements industriels et déclenche des interventions avant qu'une panne ne survienne, est un cas caractéristique d'interaction dynamique. A l'inverse, un thermomètre numérique, qui affiche simplement une température mesurée sans déclencher d'actions correctives, n'entre pas dans le champ d'application de la définition.

Disons le d'emblée : le concept de système d'IA n'est pas devenu une notion juridique déterminée avec la publication de ces lignes directrices. En cherchant à embrasser un large spectre de technologies, la définition de ce qu'est un système d'IA souffre de chevauchements (par exemple entre autonomie, adaptabilité et capacité d'inférence), d'ambiguïtés et d'une application volontairement fluctuante de certains termes. En l'absence de critères véritablement distinctifs entre système d'IA et logiciels traditionnels, la qualification d'un système comme relevant du champ d'application du Règlement sur l'IA pourra sembler parfois opportuniste, dictée plus par les nécessités réglementaires que par une approche technique et conceptuelle cohérente.

Malgré tout, ces lignes directrices ont l'avantage de rappeler que tout algorithme quel qu'il soit n'est pas de l'IA et ne tombe pas automatiquement sous le coup des nombreuses obligations énoncées dans le Règlement européen. Une partie des lignes directrices est même consacrée à la fourniture d'une définition négative d'un système d'IA. Même si cette partie aussi n'est pas totalement exempte d'ambiguïtés, la Commission européenne semble exclure de la définition d'un système d'IA certains logiciels à propos desquels des doutes pouvaient exister. Il en va ainsi notamment des systèmes d'optimisation mathématique classiques, tels que la régression linéaire appliquée à la modélisation économique, des logiciels de gestion de bases de données, qui appliquent des filtres en fonction de critères déterminés par

l'utilisateur, des systèmes heuristiques qui appliquent des règles fixes sans adaptation, comme un programme d'échecs reposant exclusivement sur un algorithme statique, ou encore des systèmes de prédiction simples, qui se contentent d'extrapoler des valeurs futures sur la base d'une moyenne des données passées. Selon la Commission européenne, tous ces systèmes ne présentent pas le degré de sophistication requis pour être qualifiés de système d'IA.

Et en Suisse ?

Le législateur suisse n'ayant pas encore réglementé l'usage de l'IA, le droit suisse ne connaît logiquement pas de définition d'un système d'IA. Cependant, le Conseil fédéral a chargé en date du 12 février 2025 le Département fédéral de justice et police de rédiger un avant-projet de loi pour la fin de l'année 2026. Selon la volonté du Conseil fédéral, cet avant-projet devra suivre une approche sectorielle permettant, pour l'essentiel, à la Suisse de ratifier la Convention-cadre sur l'intelligence artificielle du Conseil de l'Europe (pour une présentation plus détaillée de la feuille de route du Conseil fédéral : swissprivacy.law/337). La Convention ayant adopté une définition qui correspond dans les grandes lignes à celle prévue par le Règlement européen sur l'intelligence artificielle, il semble probable que la Confédération n'adoptera pas une définition s'écartant fondamentalement des lignes directrices de la Commission européenne.

Dans l'attente d'une législation suisse en matière d'IA, l'Office fédéral de la justice relève dans son analyse juridique de base que, dans la mesure où le législateur suisse privilégie une approche technologiquement neutre, le cadre juridique actuel semble globalement capable d'appréhender les états de fait les plus importants impliquant des systèmes d'IA (p. 19). Bien qu'elle ne puisse être que temporaire, cette approche semble, peu ou prou, rejoindre celle du PFPDT. Dans un communiqué du 9 novembre 2023, il rappelait également que la LPD est directement applicable à l'IA, mettant en avant les obligations auxquelles fabricants, fournisseurs et exploitants de systèmes d'IA sont soumis en termes de finalité, de transparence, d'analyse des risques ou en présence d'une décision individuelle automatisée. En outre, l'usage de systèmes d'IA reste nécessairement encadré en droit public par la protection découlant des droits fondamentaux (interdiction des discriminations, autodétermination [informationnelle], usage abusif des données personnelles, droit à un procès équitable, etc.) et en droit privé par la protection de la personnalité garantie aux articles 28 ss du Code civil.

De son côté, le secteur financier semble avoir pris les devants sous l'impulsion de la FINMA. Dans une communication sur la surveillance 08/2024 publiée à la fin de l'année 2024, elle

a formulé différentes exigences concernant l'usage de l'IA. Elle réclame notamment la mise en place d'une gouvernance et l'établissement d'un inventaire des solutions d'IA utilisées et une classification des risques associés. Comme la FINMA ne fournit elle-même aucune définition d'un système d'IA, il revient pour le moment à chaque assujetti de fixer lui-même sa propre définition. Si cette approche flexible peut certes encourager les assujettis à adapter leur politique en matière d'IA à leur contexte spécifique, elle risque cependant d'engendrer des divergences dans l'interprétation et la mise en œuvre des règles futures. En parallèle, le Conseil fédéral a demandé au Secrétariat d'État aux questions financières internationales (SFI) de dresser durant l'année 2025 un état des lieux des conditions juridiques régissant les applications d'IA dans le domaine financier. Cet état des lieux devra présenter le cadre juridique en vigueur, mettre en avant le potentiel d'innovation en tenant compte des risques d'abus et identifier les éventuels besoins concrets de mesures juridiques ou réglementaires, que ce soit par l'adoption de nouvelles normes ou de standards.

Enfin, le règlement sur l'IA s'applique aux fournisseurs qui développent et aux déployeurs qui exploitent des systèmes d'IA, qu'ils soient établis dans l'UE ou dans un pays tiers, si leurs systèmes sont mis sur le marché, sont mis en service ou produisent des effets dans l'UE. Pour les entreprises suisses, cela signifie que si les systèmes d'IA qu'elles développent ou exploitent sont utilisés dans ou depuis l'UE, ils devront avec quelques nuances se conformer aux obligations du Règlement selon leur niveau de risque : absence d'obligations pour les systèmes d'IA à risque minimal, obligations de transparence pour les systèmes d'IA à risque limité, évaluation de conformité pour les systèmes d'IA à haut risque, et interdiction pour les systèmes d'IA à risque inacceptable. De plus, les fournisseurs suisses de systèmes d'IA à haut risque ou de modèles d'IA générative devront désigner un mandataire dans l'UE avant leur mise sur le marché européen (plus en détail sur le champ d'application extraterritorial du Règlement sur l'IA : https://swissprivacy.law/321/).

Proposition de citation : Michael Montavon, La Commission européenne définit ce qu'elle entend par système d'IA, 17 mars 2025 *in* www.swissprivacy.law/343

Les articles de swissprivacy.law sont publiés sous licence creative commons CC BY 4.0.