

Bonn, Bucarest, Dublin, Lisbonne, Madrid, Milan, Paris, La Haye, Vienne, Varsovie

Comment définir un système d'IA au titre du règlement IA?

Série « Micro-Insights »

Groupe de travail CEDPO sur l'IA et les données

Octobre 2025

Auteurs:

Marc Bellon

Lionel Capel

Maria Moloney

Laura Morató Pascual

Alessandro Vasta

Informations de contact:

<https://cedpo.eu>

info@cedpo.eu

À propos de la série « Micro-Insights »

La série « Micro-Insights » est une initiative de publication du groupe de travail de CEDPO sur l'intelligence artificielle et les données. Elle propose des articles accessibles, clairs et concis, traitant des principaux enjeux à l'intersection des données et de l'intelligence artificielle. Avec une approche à la fois pratique et pédagogique, cette série vise à expliquer des sujets complexes tout en fournissant des recommandations concrètes pour leur mise en œuvre. Elle met en lumière les domaines clés, tout en conseillant les praticiens sur les impacts et les prochaines étapes à considérer. Avec l'entrée en vigueur du Règlement sur l'intelligence artificielle de l'UE (le « RIA ») en 2024, le cadre est désormais clairement défini pour tous les acteurs du secteur. Cela ouvre la voie à des discussions plus structurées et précises sur la réglementation des données et de l'IA.

La série "Micro-Insights" suivra l'évolution de l'IA et des données au cours des prochaines années. Alors que l'entrée en application du RIA se rapproche et que les technologies de l'IA évoluent de manière toujours plus novatrice et imprévisible, la série offrira des orientations actualisées et faisant autorité sur les principaux sujets de préoccupation.

La série inclut, entre autres, des documents sur :

- La réglementation des modèles d'IA à usage général en vertu du RIA
- Une vue d'ensemble du Pacte sur l'IA
- Les analyses d'impact sur les droits fondamentaux
- Le rôle des autorités de protection des données dans la réglementation de l'IA
- Une analyse de la question : le délégué à la protection des données (DPO) est-il la personne adéquate pour être également délégué à l'IA ?
- La base légale pour l'utilisation des données d'entraînement dans l'apprentissage automatique
- Une boîte à outils pour se préparer à la mise en œuvre du RIA

Table of Contents

1. Introduction	4
2. Critères de définition des systèmes d'IA.....	5
a) Un système automatisé.....	5
b) Différents niveaux d'autonomie	5
c) Capacité d'adaptation après son déploiement	5
d) Objectifs.....	6
e) Inférence	6
f) Génération de sorties	6
g) Influence sur les environnement.....	7
3. Conclusion.....	7

1. Introduction

L'un des objectifs principaux du règlement IA est de favoriser l'innovation en matière d'intelligence artificielle tout en garantissant un niveau élevé de protection de la santé, de la sécurité et des droits fondamentaux dans l'Union européenne. Il est essentiel de comprendre le champ d'application du règlement, car cela permet de déterminer précisément ce qui relève ou non de ses règles. Seuls les systèmes répondant à la définition d'un système d'IA figurant dans le règlement entrent dans son champ d'application. Compte tenu de la complexité du texte, la Commission européenne a été chargée de publier des lignes directrices afin d'aider les parties prenantes à comprendre et appliquer la définition d'un système d'IA établie par le règlement.

Pour cette raison, le 6 février 2025, la Commission européenne a publié des lignes directrices destinées à aider toutes les parties prenantes à mieux comprendre la définition des systèmes d'IA figurant dans le règlement IA et à soutenir la mise en application des dispositions, qui sont entrées en vigueur le même mois.

Ces lignes directrices sont un outil d'accompagnement mais ne sont pas juridiquement contraignantes. La Cour de justice de l'Union européenne demeure l'autorité compétente pour fournir des interprétations définitives du règlement IA. Compte tenu de l'évolution rapide des technologies d'intelligence artificielle, les lignes directrices mettent l'accent sur la nécessité d'une application flexible et non mécanique de la définition d'un système d'IA.

L'article 3, paragraphe 1, du règlement IA définit un système d'IA comme un système automatisé, conçu pour fonctionner à différents niveaux d'autonomie et pouvant faire preuve d'une capacité d'adaptation après son déploiement, 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.

En décomposant cette définition, on identifie sept critères principaux

1. Il doit s'agir d'un système automatisé
2. Il fonctionne à différents niveaux d'autonomie
3. Il peut faire preuve d'une capacité d'adaptation après son déploiement
4. Il doit y avoir au moins un ou plusieurs objectifs (explicites ou implicites)
5. Il doit y avoir inférence
6. Il doit y avoir une ou plusieurs sorties
7. Il doit y avoir une influence sur l'environnement

Le reste du document fournit des explications très brèves pour chacun de ces critères, afin d'aider le lecteur à déterminer de manière pratique ce qui constitue un système d'IA au sens du règlement IA.

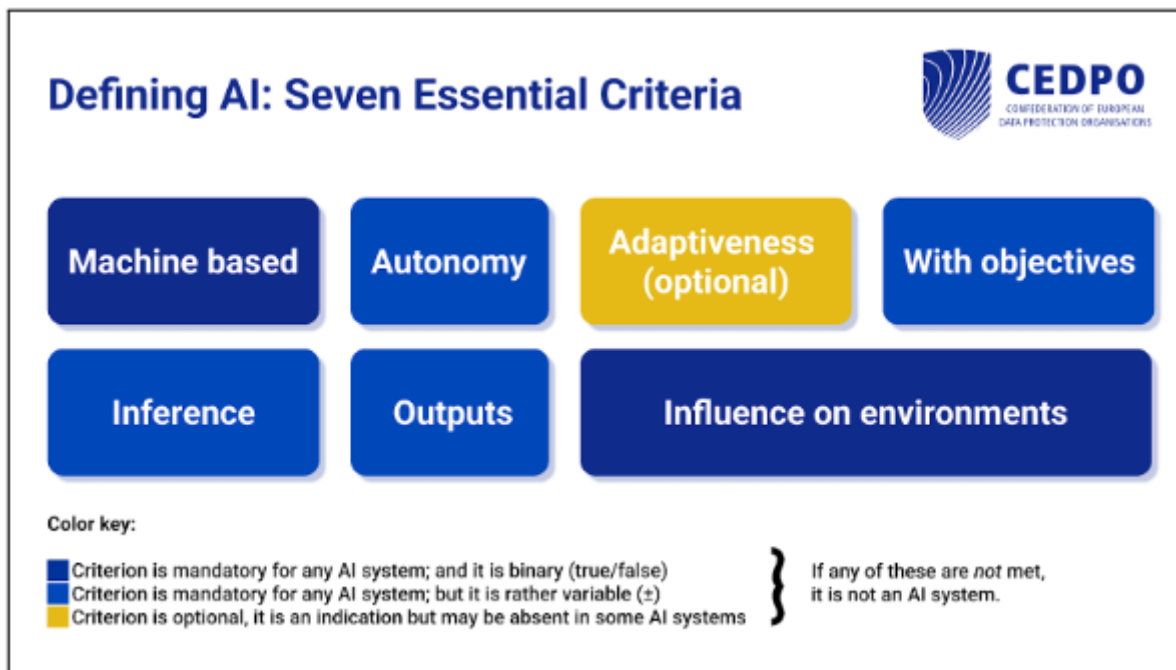


Figure 1: Les sept critères de définition d'un système d'IA

2. Critères applicables aux systèmes d'IA

a) Un système automatisé

Tous les systèmes d'IA sont développés et exécutés sur des machines : matériel (éléments physiques) et logiciel (code, instructions, programmes, systèmes d'exploitation et applications). Les systèmes d'IA peuvent s'appuyer sur de nombreux composants.

L'expression « système automatisé » est entendue largement et inclut des systèmes émergents s'écartant de l'informatique traditionnelle, notamment

- les systèmes d'informatique quantique ;
- les systèmes informatiques biologiques ou organiques (informatique biologique).

b) Niveaux d'autonomie variables

Pour déterminer si un système relève d'un système d'IA, le niveau d'autonomie est un facteur crucial. Les systèmes conçus pour fonctionner avec un degré raisonnable d'indépendance dans leurs actions satisfont au critère d'autonomie de la définition.

Un aspect clé de l'autonomie tient au niveau d'implication et d'intervention humaines. Par exemple, un système comme ChatGPT qui requiert des entrées fournies manuellement pour générer de manière indépendante une sortie, manifeste « un certain degré d'indépendance

d'action », car il est capable de produire une sortie sans qu'un humain n'en spécifie précisément le contenu.

c) Capacité d'adaptation après déploiement

Un système d'IA « peut faire preuve d'une capacité d'adaptation après son déploiement ». Cela renvoie à des capacités d'auto-apprentissage, c'est-à-dire que le système peut automatiquement apprendre, découvrir de nouveaux motifs ou identifier des relations dans les données au-delà de ce sur quoi il a été initialement entraîné. Le système peut ainsi modifier son comportement en cours d'usage et produire des résultats différents pour des entrées identiques.

À la différence d'autres critères, tous essentiels pour définir un système d'IA, la capacité d'adaptation après déploiement n'est pas décisive. C'est plutôt un indicateur. Par conséquent, un système initialement entraîné qui ne manifeste aucune adaptation une fois déployé n'est pas exclu de la définition de système d'IA.

d) Objectifs

Les « objectifs » d'un système d'IA renvoient à ses buts internes, qui déterminent la façon dont il opère. Ils peuvent être explicitement définis par des fonctions d'optimisation ou implicitement dérivés de motifs appris lors d'interactions avec l'environnement. Ces objectifs peuvent différer de « l'usage prévu » assigné par le fournisseur dans un contexte donné, car ils se concentrent sur les résultats que le système doit atteindre, sans impliquer nécessairement la façon dont il est intégré ou utilisé. Ainsi, une IA peut développer des objectifs non formulés, comme c'est le cas avec l'apprentissage par renforcement ou avec des modèles de grande taille. Cette caractéristique est essentielle pour décrire un système d'IA, mais elle ne suffit pas à elle seule pour le distinguer d'autres algorithmes.

e) Inférence

L'« inférence » en IA désigne la capacité d'un système à générer, à partir des entrées reçues, des sorties telles que des prédictions, du contenu, des recommandations ou des décisions. Ce processus implique de tirer des conclusions à partir de données, souvent via des techniques d'IA comme l'apprentissage automatique ou des approches fondées sur la logique et la connaissance. Dans le contexte de l'IA, l'inférence couvre à la fois la production de résultats à partir d'entrées lors de la phase d'utilisation et la dérivation de modèles ou d'algorithmes à partir de données lors de la phase de construction. Cette capacité d'inférer distingue les systèmes d'IA des logiciels traditionnels, qui s'appuient sur des règles prédéfinies et ne génèrent pas de sorties fondées sur des motifs appris ou dérivés.

f) Sorties

La capacité d'un système à générer des sorties, par exemple des prédictions, du contenu et des recommandations, qui peuvent influencer des environnements physiques ou virtuels à partir des entrées reçues et en recourant à des approches d'apprentissage automatique ainsi qu'à des approches logiques et fondées sur la connaissance, est fondamentale. C'est ce qui distingue ces systèmes d'autres formes de logiciels. Les sorties se répartissent en quatre catégories principales : **prédictions, contenu, recommandations et décisions**. Les **prédictions** permettent d'estimer des valeurs inconnues avec une intervention humaine minimale. Le **contenu** renvoie à la création de nouveaux contenus, tel que du texte ou des images. Les **recommandations** suggèrent des actions ou des produits à partir des données utilisateur. Les **décisions** sont des conclusions rendues par le système, qui, en pratique, automatisent un jugement humain. Les systèmes d'IA, notamment ceux utilisant l'apprentissage automatique, excellent dans le traitement de données complexes et la génération de sorties plus nuancées que les systèmes non-IA.

g) Influence sur les environnements

Un élément important de la définition d'un système d'IA est sa capacité à produire des sorties qui peuvent influencer des environnements physiques et virtuels. Cela souligne que les systèmes d'IA agissent et ont un impact sur les environnements où ils opèrent. L'expression « environnements physiques ou virtuels » couvre les objets tangibles physiques, comme un bras robotisé, et les cadres virtuels, y compris les espaces numériques, les flux de données et les écosystèmes logiciels. Un agent d'IA qui réserve des billets d'avion est un exemple de système d'IA influençant un environnement physique.

3. Conclusion

Le règlement IA propose une définition complète des systèmes d'IA. Il décrit ces systèmes au travers de sept critères clés exposés dans ce document. La leçon principale de l'exercice de définition est que ces systèmes sont, en substance, des logiciels développés à l'aide de techniques comme l'apprentissage automatique, les méthodes statistiques et d'autres approches logiques ou fondées sur la connaissance, qui aboutissent à des décisions, prédictions ou recommandations susceptibles d'influencer nos environnements physiques et virtuels.

Cette définition complexe cherche à saisir la nature dynamique et en constante évolution de l'IA tout en établissant des bases juridiques claires pour la superviser. En ancrant les systèmes d'IA dans ce cadre réglementaire, l'UE cherche un équilibre entre stimulation de l'innovation et protection de la santé, de la sécurité et des droits fondamentaux. Cela promeut la transparence, la responsabilité et la confiance dans le développement et le déploiement des technologies d'IA dans l'Union et pour l'avenir.