

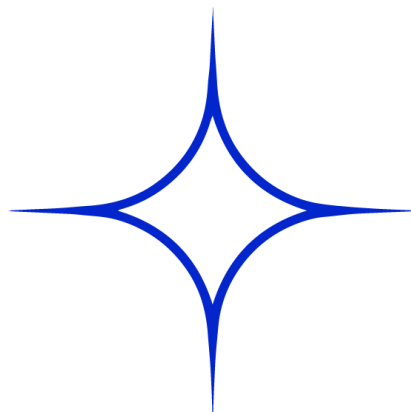
IA & Santé : Accès aux données de qualité et Développement de la Puissance de calcul publique

Les grandes lignes de l'IA Act

Modernisation de l'approche de la protection des données, Nécessité de développer la puissance de calcul publique, Adaptation du DMA pour préserver la concurrence sur la chaîne de valeur l'IA Générative

Rapport Notre ambition pour la France : un optimisme raisonné de la Commission de l'Intelligence Artificielle

Eléonore SCARAMOZZINO, Avocate, Constellation Avocats





IA & Santé :
Accès aux données de qualité
et Développement de la Puissance de calcul publique

Les grandes lignes de l'IA Act, Modernisation de l'approche de la protection des données, Nécessité de développer la puissance de calcul publique, Adaptation du DMA pour préserver la concurrence sur la chaîne de valeur l'IA Générative

Rapport Notre ambition pour la France : un optimisme raisonné de la Commission de l'Intelligence Artificielle

Par Eléonore Scaramozzino, Avocat, Constellation Avocats

SOMMAIRE

L'AI ACT : adoption par le Parlement européen	5
Les grandes lignes.....	5
Implémentation de cette réglementation	7
<i>La mise en œuvre au niveau européen,</i>	7
<i>La mise en œuvre au niveau national</i>	8
<i>Les normes</i>	8
Les recommandations de la Commission de l'IA : application au domaine de la santé.....	9
Les prérequis de l'IA applicables au domaine de la Santé.....	10
Le Rôle du DMA sur le maintien de la concurrence sur la chaîne de valeur de l'IA Générative	11
Les enjeux concurrentiels sur la chaîne de valeur de l'IA Générative :.....	11
Rappel sur le Règlement sur les marchés numériques : DMA.....	13
Le recours au DMA pour maintenir la concurrence sur la chaîne de valeur de l'IA générative ?	15
Modernisation de l'approche de la donnée personnelle par le RGPD pour faciliter l'innovation tout en protégeant l'individu:	16
Nécessité de développer la puissance de calcul en France et en Europe	18



IA & Santé : Accès aux données de qualité et Développement de la Puissance de calcul publique

Les grandes lignes de l'IA Act, Modernisation de l'approche de la protection des données, Nécessité de développer la puissance de calcul publique, Adaptation du DMA pour préserver la concurrence sur la chaîne de valeur l'IA Générative

Rapport Notre ambition pour la France : un optimisme raisonné de la Commission de l'Intelligence Artificielle

Par Eléonore Scaramozzino, Avocat, Constellation Avocats

Version du 10.04.2024



Le 13 mars 2024, le Parlement européen adoptait en plénière, le règlement sur l'Intelligence Artificielle (IA Act) par 523 votes pour, 46 contre et 49 abstentions. L'IA Act vise à mettre en place des garanties relatives à l'intelligence artificielle à usage général et protéger les droits fondamentaux, tout en encourageant l'innovation. Le même jour, la Commission sur l'Intelligence Artificielle¹ remettait au Président de la République, son rapport intitulé : « *IA : une ambition pour la France* ». L'IA générative constitue un tournant dans l'innovation. L'écosystème a besoin de sécurité juridique, des données de qualité, des supercalculateurs et des financements dédiés à l'IA. Ce rapport n'est pas axé principalement sur l'IA dans le domaine de la santé, mais il propose des mesures intéressantes pour

¹ La Commission de l'Intelligence artificielle est présidée par Philippe Aghion, "Professeur au Collège de France et Anne Bouverot, Présidente du Conseil d'Administration de l'Ecole Normale Supérieure. Elle est composée 13 experts de l'IA : Gilles Babinet, Joëlle Barral, Alexandra Bensamoun, Nozha Boujemaa, Benard Charlès, Luc Juia, Yann Le Cun, Aerhur Mensch, Cédric O, Isabel Ryl, France Salis-Madinier, Martin Tisné, Gaël Varoquaux.

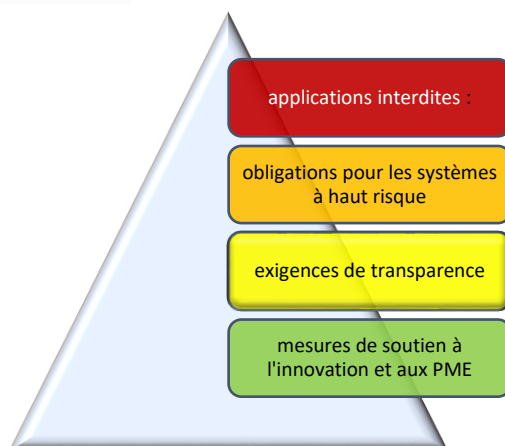
faciliter la circulation des données. Les auteurs s'interrogent sur l'équilibre à trouver entre la protection des données et leur accessibilité. Ils soulignent l'importance des données pour les technologies d'apprentissage automatique. La limitation de l'accès aux données pour protéger la confidentialité, la propriété intellectuelle, conduit à restreindre l'innovation ou à la laisser entre les mains de quelques acteurs capables de collecter plus de données et de supporter le coût de la régulation. D'autres leviers, comme la disponibilité de la puissance de calcul, l'investissement, le personnel expert dans le domaine sont examinés. *« Le rapport restitue les termes du débat et apporte un ensemble de réponses afin de créer un cercle vertueux. Celles-ci reflètent les convictions de notre Commission. Ces convictions, forgées entre septembre 2023 et février 2024, se fondent sur l'expertise des membres et des rapporteurs. Elles sont étayées par la consultation de 7 000 personnes et l'audition de 600 experts et parties prenantes ».*

Avant d'examiner les propositions de la Commission de l'Intelligence Artificielle, rappel des grandes lignes du Règlement sur les systèmes d'Intelligence Artificielle adopté par le parlement européen.

L'AI ACT : adoption par le Parlement européen

Les grandes lignes

Le règlement établissant des règles harmonisées concernant l'intelligence artificielle (législation sur l'intelligence artificielle) et modifiant certains actes législatifs de l'Union prévoit des obligations pour les systèmes d'IA en fonction de leurs risques potentiels et de leur niveau d'impact.



Applications interdites

- qui menacent les droits des citoyens, y compris les systèmes de catégorisation biométrique utilisant des caractéristiques sensibles et l'extraction non ciblée d'images faciales sur Internet ou par vidéosurveillance pour créer des bases de données de reconnaissance faciale
- Reconnaissance des émotions sur le lieu de travail et dans les établissements d'enseignement,
- la notation sociale,
- la police prédictive (basée sur le profilage d'une personne ou sur l'évaluation de ses caractéristiques)
- IA manipulant le comportement humain ou explicitant les vulnérabilités des personnes seront également interdites

exemptions pour les services répressifs

L'utilisation des systèmes d'identification biométrique par les services répressifs est en principe interdite, sauf dans des situations clairement énumérées et étroitement définies.

	<p>Les systèmes d'identification biométrique "en temps réel" ne peuvent être déployés que si des garanties strictes sont respectées. Leur utilisation est par exemple limitée dans le temps et géographiquement et soumise à une autorisation judiciaire ou administrative préalable spécifique. Ces systèmes pourront notamment servir à la recherche d'une personne disparue ou à la prévention d'une attaque terroriste. L'utilisation de tels systèmes d'identification biométrique à distance "a posteriori" est considérée comme un cas d'utilisation à haut risque, nécessitant une autorisation judiciaire liée à une infraction pénale.</p>
<p>Des obligations pour les systèmes à haut risque</p>	<p>Des exemptions pour les services répressifs</p> <p>L'utilisation des systèmes d'identification biométrique par les services répressifs est en principe interdite, sauf dans des situations clairement énumérées et étroitement définies. Les systèmes d'identification biométrique "en temps réel" ne peuvent être déployés que si des garanties strictes sont respectées.</p> <p>Leur utilisation est par exemple limitée dans le temps et géographiquement et soumise à une autorisation judiciaire ou administrative préalable spécifique.</p> <p>Ces systèmes pourront notamment servir à la recherche d'une personne disparue ou à la prévention d'une attaque terroriste.</p> <p>L'utilisation de tels systèmes d'identification biométrique à distance "a posteriori" est considérée comme un cas d'utilisation à haut risque, nécessitant une autorisation judiciaire liée à une infraction pénale.</p>
<p>Exigences de transparence</p>	<p>Ces systèmes doivent faire l'objet d'une évaluation et d'une réduction des risques, être assortis de registres d'utilisation, être transparents et précis et être soumis à une supervision humaine. Les citoyens auront le droit de déposer des plaintes concernant les systèmes d'IA et de recevoir des explications sur les décisions basées sur des systèmes d'IA à haut risque qui ont une incidence sur leurs droits</p> <p>Les systèmes d'IA à usage général et les modèles sur lesquels ils sont basés devront respecter des exigences de transparence et la législation européenne sur les droits d'auteurs, ainsi que la publication de résumés détaillés des contenus utilisés pour leur entraînement.</p> <p>Les systèmes d'IA à usage général plus puissants présentant un risque systémique seront soumis à des exigences supplémentaires.</p> <ul style="list-style-type: none"> • Par exemple, des évaluations de modèles devront être effectuées, les risques systémiques devront être évalués et atténués et les incidents devront être signalés. • De plus, les images et les contenus audio et vidéo artificiels ou manipulés (aussi appelés "deep fakes") doivent être clairement signalés comme tels.
<p>mesures de soutien à l'innovation et aux PME</p>	<p>Des "bacs à sable réglementaires" et des essais en conditions réelles devront être mis en place par les autorités nationales et mis à disposition des PME et des start-ups pour développer et tester des IA innovantes avant leur mise sur le marché.</p>

Implémentation de cette réglementation

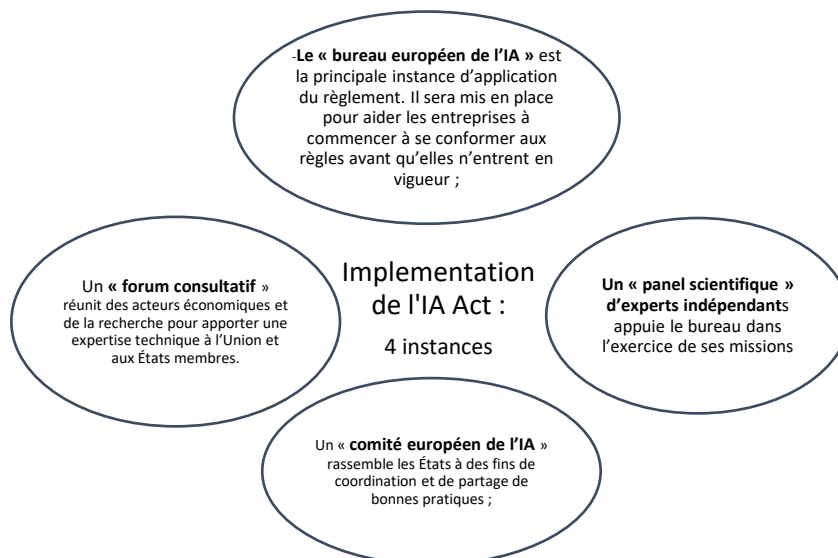
L'adoption de ce règlement est un point de départ. Cependant, l'implémentation est complexe.

La législation entrera en vigueur 20 jours après sa publication au Journal officiel et sera pleinement applicable 24 mois après son entrée en vigueur, à l'exception

A l'exception	après la date d'entrée en vigueur
de l'interdiction des pratiques interdites	6 mois,
des codes de pratique	9 mois
des règles concernant l'IA à usage général, notamment la gouvernance	12 mois
des obligations pour les systèmes à haut risque	36 mois

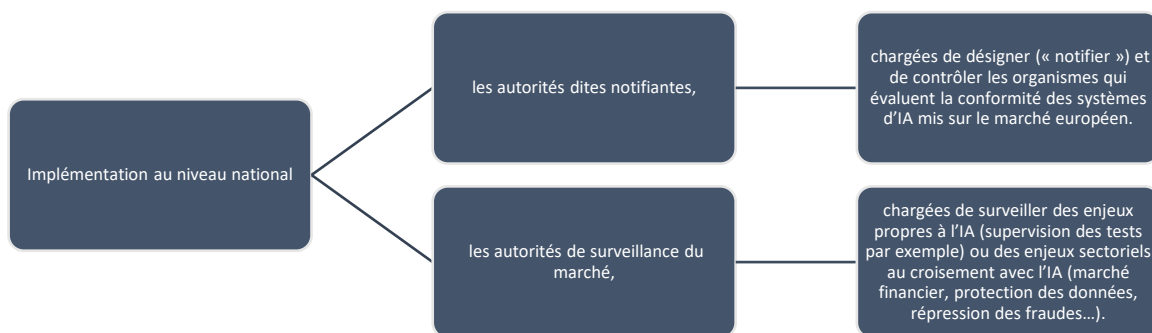
La mise en œuvre au niveau européen,

4 instances sont créées.



La mise en œuvre au niveau national

Au niveau national, la mise en œuvre du règlement sur l'IA s'appuie sur deux ensembles d'acteurs.



Les normes

- En Europe, les normes harmonisées sur lesquelles repose l'application du règlement sur l'IA sont appelées à jouer un rôle central. La Commission européenne a donné mandat au Comité européen de normalisation appelé CEN/CENELEC pour les établir, avec un objectif ambitieux : être prêt dès l'entrée en vigueur du règlement, en 2026.
- Au plan mondial, de nombreuses enceintes mènent des travaux afin de définir les standards et méthodes nécessaires à l'évaluation des systèmes d'IA, qu'il s'agisse d'organismes de normalisation (l'ISO, au niveau mondial, le NIST, aux États-Unis, le BSI, au Royaume-Uni...) ou de forums « privés ».

Les recommandations de la Commission de l'IA : application au domaine de la santé



Les usages de l'IA dépendent de

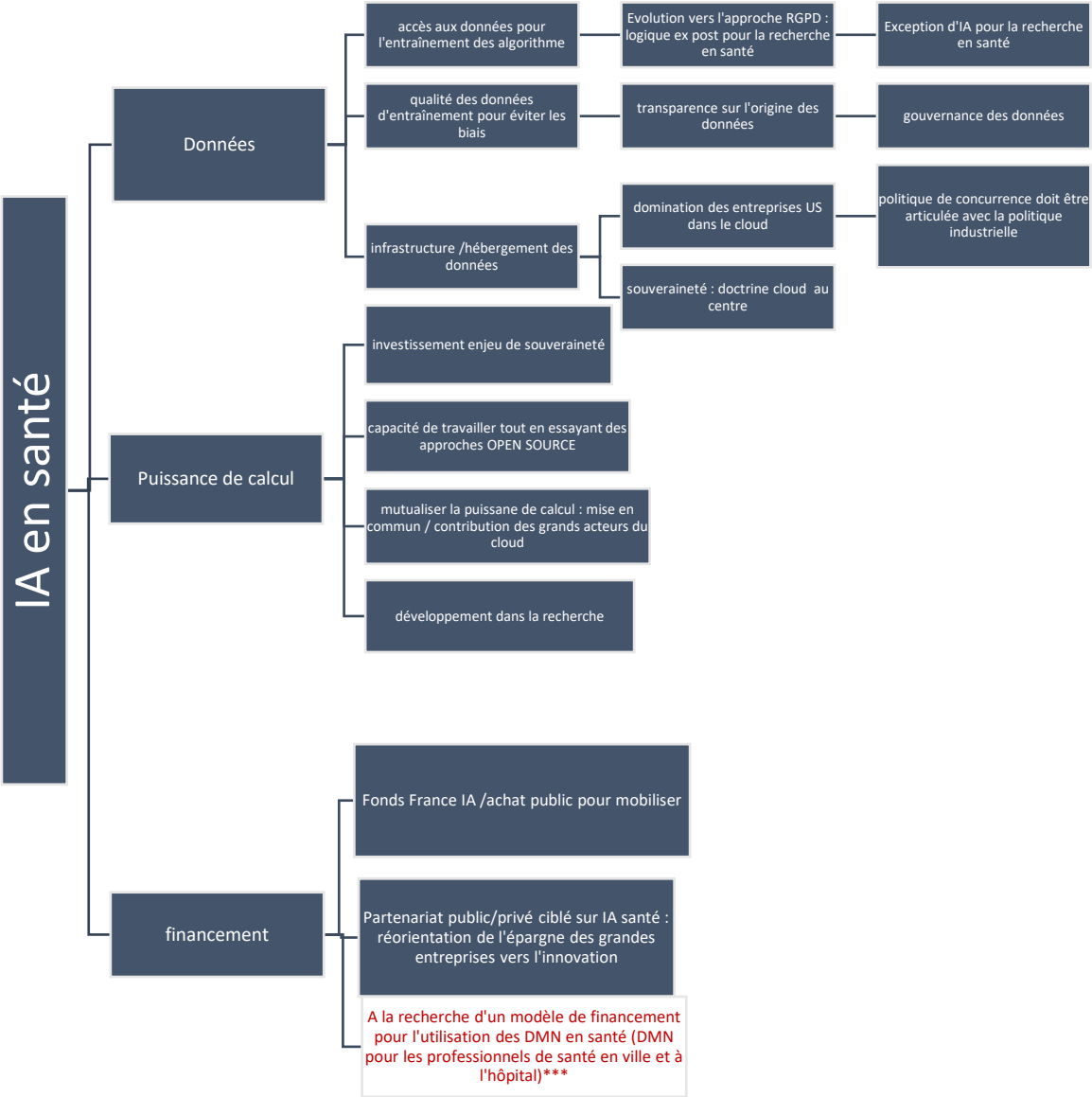
- la technologie : nouvelles capacités, performances, contrôle, sûreté et open source ;
- des ressources : talents, données, puissance de calcul, financement ;
- la confiance : renforcée par l'évaluation des SI, audit, sécurité juridique, sensibilisation ;
- des bénéfiques : activité économique, emploi, productivité amélioration des services publics et de la qualité de vie au travail.

Pour l'IA dans le domaine de la santé, le rapport recommande notamment de :

« **faciliter l'accès aux données** : en matière de données à caractère personnel, modernisation du mandat de la Commission nationale de l'informatique et des libertés (Cnil) et de son collègue, suppression de certaines **procédures d'autorisation préalable d'accès aux données de santé et réduction des délais de réponse** ; *en matière culturelle, mise en place de l'infrastructure technique favorisant l'entraînement des modèles d'IA dans le respect des droits de propriété intellectuelle* ; »

Les prérequis de l'IA applicables au domaine de la Santé

Le développement de l'IA nécessite des données, de la puissance de calcul et un financement dédié.





***Sur la question du financement, la Commission de l'IA n'a pas abordé la question du modèle économique des dispositifs médicaux numériques avec IA, nous l'avons rajouté car c'est une question spécifique à la santé qui est en cours de réflexion.

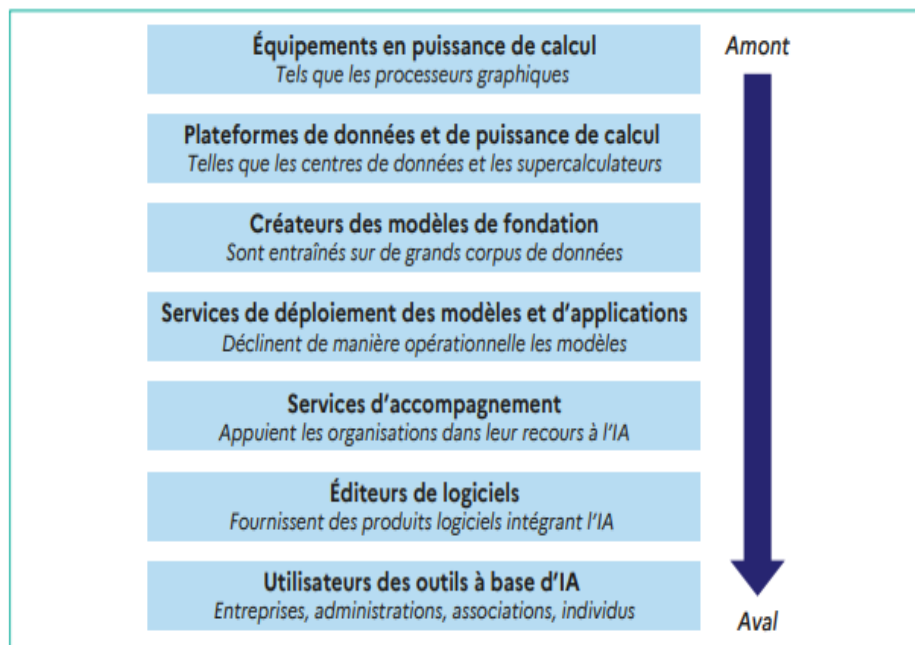
Le plan France 2030 a notamment consacré des fonds pour le développement de l'IA en Santé. Cependant, comme le précisent les auteurs, le soutien public ne doit pas viser un rattrapage permanent en courant derrière les pionniers, mais une supériorité sur certaines briques de la chaîne de valeur, en sélectionnant quelques domaines de compétition pour les remporter. Cependant, une intervention excessive ou mal dimensionnée pourrait nuire de manière disproportionnée à l'innovation et nous priver de ses bénéfices.

Le Rôle du DMA sur le maintien de la concurrence sur la chaîne de valeur de l'IA Générative

Digital Market Act, ci-après DMA : le règlement (UE) 2022/1925 relatif aux marchés contestables et équitables dans le secteur numérique et modifiant les directives (UE) 2019/1937 et (UE) 2020/1828 (Règlement sur les marchés numériques).

Les enjeux concurrentiels sur la chaîne de valeur de l'IA Générative :

Dans son rapport, la Commission de l'IA montre que la chaîne de valeur de l'IA générative, les ingrédients indispensables aux créateurs de modèles de fondation, que sont les données et les infrastructures de calcul. La chaîne de valeur de l'IA générative se compose d'entreprises dont l'IA est au cœur du modèle d'affaires, et d'entreprises qui utilisent l'IA et l'intègrent dans un modèle d'affaires préexistant ou adapté. Il y a donc **une économie de l'IA** et **une économie par l'IA**.



Source Rapport IA Notre ambition pour la France – Commission de l'Intelligence Artificielle p.22

Le développement de l'IA confirme et renforce le risque de concentration de la chaîne de valeur du numérique. Anticiper l'évolution concurrentielle de la chaîne de valeur de l'IA pour garantir un marché contestable est nécessaire pour permettre l'entrée de nouveaux acteurs. Les acteurs américains dominent très largement l'amont de la chaîne de valeur de l'IA générative. Les processeurs graphiques (GPU) sont aujourd'hui les équipements de puissance de calcul les plus indispensables au fonctionnement des systèmes d'IA générative. Or, une seule entreprise — américaine — détient actuellement 80 % des parts de marché mondial de conception des processeurs graphiques. Comme l'indique l'autorité de la concurrence dans son avis de juin 2023², le secteur du cloud, en amont de la chaîne de valeur de l'IA, est dominé par trois grands acteurs, les « hyperscalers » (Amazon Web Services, Google Cloud Platform et Microsoft Azure). En 2021, ces trois acteurs représentaient 80 % de la croissance des dépenses en infrastructures et applications de services cloud public en France. Ces entreprises sont en mesure d'entraver la bonne concurrence sur la chaîne de valeur de l'IA générative.

² Avis 23-A-08 du 29 juin 2023 portant sur le fonctionnement concurrentiel de l'informatique en nuage ("cloud")



Le **Digital Markets Act (DMA)**, a pour objectif de permettre aux marchés dans le secteur numérique de rester contestables et équitables. Cette réglementation vise à empêcher qu'une grande plateforme jouissant d'une position de « **contrôleur d'accès** » (**gatekeeper**) vis-à-vis d'un grand nombre d'utilisateurs n'abuse pas de cette position en empêchant des entrants potentiels d'accéder à ces utilisateurs, et donc à ce marché.

Rappel sur le Règlement sur les marchés numériques : DMA

Notification pour les Entreprises dépassant le seuil de l'art 3.3-2 DMA

Dans le cadre de la mise en œuvre du DMA, la Commission a adopté le 14 avril 2023, le règlement d'exécution ainsi que ses deux annexes, lesquelles comprennent le formulaire que les entreprises qui atteignent les seuils fixés à l'article 3, paragraphe 2, du règlement sur les marchés numériques doivent utiliser pour la notification à effectuer à la Commission aux fins de leur désignation comme contrôleur d'accès (ci-après le «formulaire GD»), ainsi que les spécifications relatives au format et à la longueur des documents présentés au titre du règlement sur les marchés numériques. Aux termes de l'article 3, paragraphe 3, du DMA, lorsque des entreprises fournissant des services de plateforme essentiels atteignent les seuils fixés à l'article 3, paragraphe 2, dudit règlement, elles doivent en informer la Commission.

Le 3 juillet 2023, la Commission a reçu des notifications d'Alphabet, d'Amazon, d'Apple, de ByteDance, de Meta, de Microsoft et de Samsung au titre de l'article 3, paragraphe 3, premier alinéa, du DMA. Le 5 septembre 2023, la Commission a désigné Alphabet, Amazon, Apple, ByteDance, Meta et Microsoft comme contrôleurs d'accès pour un total de 22 services de plateforme essentiels, et a adopté une décision en vertu de laquelle elle ne désignait pas Samsung comme contrôleur d'accès en ce qui concerne son navigateur internet.



Source : Rapport de la Commission européenne du 6.3.2024 COM(2024) 106 final Rapport annuel sur le règlement UE) 2022/1925 relatif aux marchés contestables et équitables dans le secteur numérique et modifiant les directives (UE) 2019/1937 et (UE) 2020/1828 (Règlement sur les marchés numériques)

Les contrôleurs d'accès doivent se conformer aux obligations énoncées aux articles 5, 6, 7 et 15 du DMA six mois après leur désignation, de sorte que les entreprises désignées au 5 septembre 2023 devaient se conformer à ces obligations au plus tard le 7 mars 2024.

Informations sur les concentrations : Art 14 DMA

Conformément à l'article 14 du DMA, les contrôleurs d'accès sont tenus d'informer la Commission de tout projet de concentration, lorsque les entités qui fusionnent ou la cible de la concentration fournissent des services de plateforme essentiels ou d'autres services dans le secteur numérique, ou permettent la collecte de données.



Fonction de vérification de la Conformité : art 28 DMA

Les contrôleurs d'accès sont tenus de mettre en place une fonction de vérification de la conformité, qui est indépendante des fonctions opérationnelles du contrôleur d'accès et fait appel à un ou plusieurs responsables de la conformité disposant d'une autorité, d'une stature et de ressources suffisantes pour être en mesure de contrôler le respect dudit règlement par le contrôleur d'accès et de le conseiller à ce sujet. Dans son rapport du 6 mars 2024, sur la mise en œuvre du règlement sur les marchés numériques, la Commission a indiqué avoir contrôlé la mise en place d'une telle fonction de vérification de la conformité par chaque contrôleur d'accès désigné afin de s'assurer qu'il satisfait aux exigences énoncées à l'article 28 du règlement sur les marchés numériques. À la suite de discussions avec la Commission et de la publication par cette dernière de lignes directrices concernant ces obligations, tous les contrôleurs d'accès désignés ont nommé des responsables de la conformité dans le respect des principes énoncés à l'article 28 du DMA et en ont communiqué les détails à la Commission.

Le recours au DMA pour maintenir la concurrence sur la chaîne de valeur de l'IA générative ?

Pour la Commission de l'IA

- Le DMA devrait permettre de s'attaquer **aux comportements anticoncurrentiels en aval de la chaîne de valeur** : on peut penser à celui de la vente liée, ou à la réutilisation de données personnelles collectées grâce à l'utilisation d'un premier produit (par exemple la collecte de données personnelles via un système de messagerie pour faire de la publicité ciblée sur un réseau social appartenant à la même plateforme).
- Le DMA devrait également permettre de s'intéresser aux **problèmes de concentration sur certains marchés numériques**, situation dans laquelle un faible nombre d'acteurs voire un seul acteur domine le marché, grâce à l'obligation faite aux « contrôleurs d'accès » de déclarer l'ensemble de leurs acquisitions.
- **Les limites du DMA : conçu à travers le prisme de la plateforme.** Le DMA est en mesure d'apporter des solutions aux enjeux concurrentiels présents sur la chaîne de valeur de l'IA tant que ces enjeux sont reliés à une problématique de plateforme. *« Il ne sera probablement pas suffisant pour couvrir tous les enjeux concurrentiels. Pour le Commission IA Ce règlement mériterait donc d'être complété pour tenir compte des spécificités de l'IA. »*



Enfin, la Commission IA recommande d'envisager un changement de doctrine de la politique de concurrence, permettant d'anticiper les concentrations plutôt que d'attendre de pouvoir les constater.

Modernisation de l'approche de la donnée personnelle par le RGPD pour faciliter l'innovation tout en protégeant l'individu:

Le rapport souligne l'importance de la qualité des données accessibles pour l'IA générative et recommande de repenser la gouvernance de la donnée.

La Commission IA préconise :

- **Suppression de certaines procédures d'autorisation préalable** et réduction des délais de réponse de la Commission nationale de l'informatique et des libertés (Cnil). Les auteurs souhaitent basculer vers une logique RGPD (surveillance ex post) et supprimer les autorisations préalables existant dans le domaine de la santé avec la loi Informatique et Liberté modifiée. Cette autorisation préalable constitue un frein pour les auteurs qu'il convient de lever pour faciliter l'exploitation des données de santé pour une finalité de recherche. La Cnil a instauré une procédure simplifiée de déclaration de conformité à des méthodologies de référence (MR). En pratique, la procédure simplifiée reste l'exception par rapport à la procédure d'autorisation préalable car le moindre écart de conformité avec les exigences de la MR implique de basculer vers la procédure de décision d'autorisation individuelle de la CNIL.
- **Adaptation du RGPD**
 - **Sur le plan technique, plusieurs notions clés du RGPD sont ainsi d'un maniement malaisé face au fonctionnement de l'IA.**
 - La notion de « *responsable du traitement* », pour laquelle la répartition des responsabilités entre le développeur qui a procédé à l'entraînement d'une IA générative et qui la met à disposition de tiers et l'utilisateur final du système pour ses propres besoins n'apparaît pas forcément aller de soi.
 - La notion de *finalité du traitement*, qui conditionne la nature des données pouvant légalement être utilisées et sur laquelle porte le consentement des personnes concernées est également plus complexe à appréhender, eu égard

aux nombreuses utilisations possibles d'une IA générative une fois celle-ci entraînée.

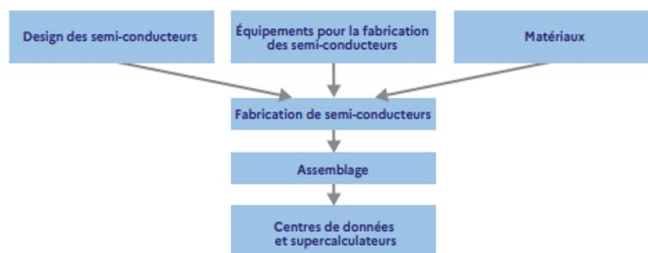
- **Sur le plan des principes,**
 - la notion même de donnée personnelle, qui constitue la clé d'application du RGPD, suscite des interrogations dans un contexte croissant d'utilisation de données collectives.
 - En l'état, juridiquement, seul un processus d'anonymisation des données personnelles permet de « sortir » du régime de protection des données personnelles du RGPD. Or, la technologie ouvre de plus en plus loin des possibilités de réidentification de données anonymisées.
- **Une gestion plus collective des données** (aujourd'hui embryonnaire dans le RGPD) pourrait améliorer la protection des intérêts et l'exercice des droits, notamment face aux acteurs mondiaux de la donnée. En effet, que ce soit au travers d'associations, de syndicats ou de tout autre collectif organisé, des accords relatifs aux traitements des données et à l'utilisation de systèmes d'IA pourraient permettre d'accroître l'effectivité de la garantie des droits de chacun.
- **Respect de l'application du principe de transparence des données d'entraînement des grands modèles d'IA.** Prévu par le règlement européen relatif à l'IA, ce principe doit permettre le respect du droit de la propriété littéraire et artistique. Il doit pouvoir être mis en œuvre de la façon la plus simple possible, pour les développeurs de modèles d'IA comme pour les ayants droit, en particulier par l'élaboration de standards pour la publication des informations sur les modèles d'IA et la mise en œuvre du droit de retrait.
- **Réforme du mandat de la Cnil,** pour y ajouter un objectif d'innovation, de réviser la composition de son collège et de renforcer ses moyens. Cette évolution impliquera un ajustement de la composition de son collège, pour qu'une palette plus large de compétences soit représentée (innovation, recherche...), et un **renforcement de ses moyens de fonctionnement.**

Nécessité de développer la puissance de calcul en France et en Europe

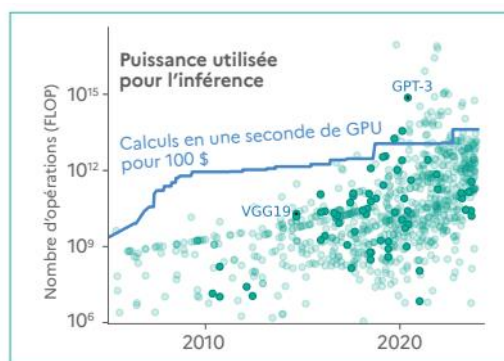
La puissance de calcul :

- est un *ingrédient incontournable de l'IA générative* ;
- constitue le socle de la chaîne de valeur de l'IA générative ;
- consiste en un mélange de matériel électronique distribué et de services associés.

En effet, les calculs requis pour l'entraînement ou l'utilisation (« l'inférence ») d'un modèle d'IA sont réalisés par des réseaux de semi-conducteurs assemblés, formant des serveurs eux-mêmes rassemblés au sein de supercalculateurs ou de centres de données. Lorsqu'une entreprise, une administration ou un particulier utilise un modèle d'IA, il recourt donc aux services de calcul de plusieurs supercalculateurs ou de centres de données.



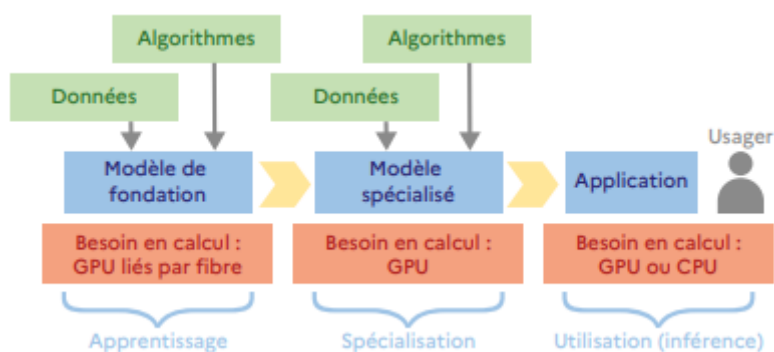
Source Commission de l'IA : Rapport Une Ambition pour la France p.92



Graphique 11 : Évolution de la puissance et du coût de l'inférence des modèles d'IA.
 Source : calculs de la Commission IA (2024).
 Lecture : Chaque point représente un modèle d'IA et la puissance nécessaire pour pour l'inférence.
 La courbe indique l'évolution du nombre de calculs réalisables en une seconde de GPU pour 100 \$.

Source Commission de l'IA : Rapport Une Ambition pour la France p.92

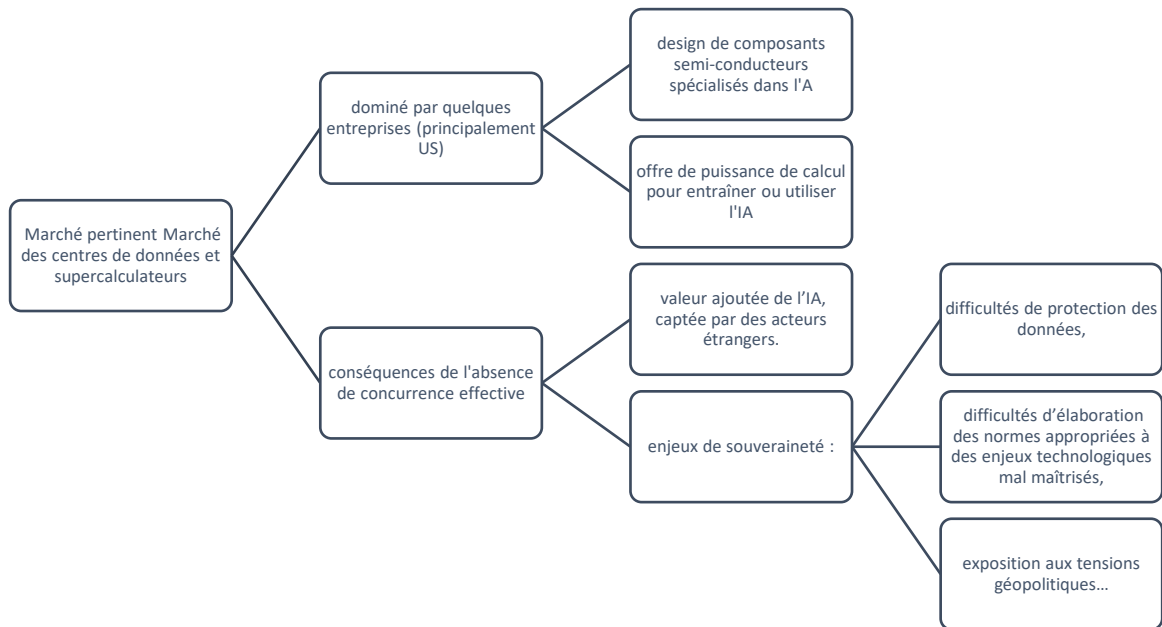
Le rapport souligne les différentes étapes nécessaires à la création d'une application d'IA générative et les différents réseaux de composants semi-conducteurs qui vont réaliser les calculs nécessaires. Les étapes d'apprentissage et de spécialisation nécessitent un calcul massif à réaliser de manière intensive, ceux-ci sont réalisés par des GPU (graphics processing unit), qui ont été créés initialement pour le jeu vidéo et non l'IA. L'étape d'inférence, quant à elle, requiert beaucoup moins de puissance de calcul. Elle peut faire appel aux GPU ou à des composants semi-conducteurs plus répandus : les CPU (central processing unit).



« un modèle de fondation permet la création de multiples modèles spécialisés dans plusieurs domaines, chacun de ces modèles spécialisés étant ensuite utilisé par de multiples usagers. Au total, l'entraînement et l'inférence créent un besoin très important de puissance de calcul tout au long de la chaîne de valeur »

Extrait Rapport Une ambition pour l'IA p.94

Des efforts importants sont réalisés aujourd'hui pour diminuer la puissance de calcul nécessaire à l'entraînement et à l'inférence, car c'est l'un des principaux postes de coûts dans la construction et l'utilisation d'un modèle d'IA, et l'un des obstacles à sa diffusion. La compétition se joue autant sur la performance des modèles que sur la capacité à développer des IA dont le coût d'utilisation est acceptable. Comme le précise le rapport, le marché est dominé par quelques entreprises, principalement américaines, détenant individuellement ou collectivement jusqu'à 80 % des parts de marché mondial de certains segments d'activité, que ce soit dans le design de composants semi-conducteurs spécialisés dans l'IA, ou l'offre de puissance de calcul pour entraîner ou utiliser des IA.



La Commission sur l'IA souligne que « **le retard de l'Europe en matière de puissance de calcul publique et installée par des acteurs privés pose des difficultés, notamment pour les usages sensibles de l'IA, sur lesquels il n'est pas envisageable d'entraîner les modèles ou les déployer sur de la puissance de calcul extra européenne** »

Cette absence de puissance de calcul publique pose problème pour le développement de l'IA dans la santé, compte tenu de la protection des données.

Proposition de la Commission de l'IA

La Commission de l'IA considère que la France et l'Europe doivent impérativement devenir un pôle majeur de la puissance de calcul installée avec trois objectifs :

- fournir une puissance de calcul publique pour les cas d'usage sensibles ;
- fournir une puissance de calcul accessible et abordable pour stimuler la recherche et le développement des start-up en IA ;
- être en mesure d'entraîner et d'utiliser sur le sol européen les modèles d'IA les plus avancés.

Du côté de l'offre, la Commission sur l'IA recommande à la fois :

- d'accélérer les travaux d'extension des supercalculateurs français et européens de taille « exascale »,

- de lancer à court terme une opération d'achat groupé pour l'écosystème et de fixer un objectif pour l'implantation de centres de calcul sur le territoire européen, assorti à la fois d'une garantie publique d'utilisation de la puissance de calcul et d'un accompagnement à l'implantation et au raccordement électrique.

La Commission de l'intelligence artificielle a réalisé un travail remarquable d'analyse et de proposition pour déployer et maîtriser l'IA. L'Etape de la mise en œuvre de ses recommandations est déterminante pour que le développement de l'IA en France et en Europe, et notamment en santé. On attend avec impatience la suite

A suivre



Pour toute information complémentaire ou question vous pouvez contacter directement l'auteur

Eléonore Scaramozzino
Avocat
Constellation Avocats
escaramozzino@constellation.law

