QUESTIONS SUR LA PROPOSITION DE LA COMMISSION EUROPEENNE POUR LA REGLEMENTATION DE L'IA - JUILLET 2021

GROUPE DE TRAVAIL - BANQUES ET AUDITABILITE - HUB FRANCE IA

TABLE DES MATIERES

INTRODUCTION

Le Groupe de travail Banque et Auditabilité du Hub France IA, qui regroupe des experts IA de trois grandes banques françaises, BNP Paribas, la Banque Postale et Société Générale, souhaite apporter une réponse à la Commission Européenne dans le cadre de la consultation ouverte sur le projet de régulation des systèmes IA (AI Act).

Les membres du groupe de travail soulèvent un ensemble d'interrogations et apportent des cas d'exemples de systèmes IA qualifiés à « haut risque » relatifs aux applications déployées dans le secteur bancaire, en analysant les problèmes identifiés.

En tant qu'experts du secteur bancaire, nous amenons notre connaissance de la mise en pratique de réglementations et actions d'audit auxquelles les banques sont soumises sur une grande partie de leurs activités et des systèmes qui sont déployés par les différentes structures. Les banques sont déjà organisées pour mettre en œuvre une réglementation comme celle proposée par la CE. En effet, une partie des mesures présentées dans le cadre de la proposition est couverte par les processus d'audit déjà en place.

Nous sommes ainsi convaincus de l'importance de réglementer les systèmes d'IA au même titre que sont aujourd'hui réglementés les systèmes bancaires, dans une perspective d'uniformisation des procédures et d'extension des régulations existantes à tous les systèmes qualifiés à « haut risque ». En tant que citoyens également, nous comprenons l'importance de réglementer de tels systèmes, et aussi d'identifier de façon précise les systèmes d'IA dont le déploiement serait totalement interdit en Europe.

Néanmoins, à la lecture approfondie de la proposition, elle nous apparaît, pour les systèmes à « haut risque », particulièrement complexe à mettre en place, voire impossible dans certains cas d'usages innovants de l'IA, en raison de divers aspects critiques :

- La définition proposée de l'Intelligence Artificielle est extrêmement large ;
- L'emploi de certains termes est parfois sujet à interprétation d'une part pour la définition d'un système à haut risque, mais surtout dans les attentes au niveau de la mise en œuvre des exigences de conformité ;
- Le périmètre d'application est parfois mal identifié et nous avons longuement débattu sur les spécificités des cas d'usage décrits comme à haut-risque ;
- Un nombre important d'exigences de conformité nous apparaît comme irréalisables en pratique. Nous pensons qu'une approche proposant des principes plutôt que des exigences spécifiques serait à même de s'intégrer dans les organisations actuelles.

Les questions que nous identifions font ainsi apparaître un véritable risque d'insécurité juridique absolument préjudiciable pour l'investissement et l'innovation.

Enfin, compte tenu de la faible maturité de certaines technologies d'IA, mais également de certaines entreprises ou certains métiers vis-à-vis de l'IA, cette règlementation apparait particulièrement contraignante et laisse craindre un vrai ralentissement de l'innovation pour les entreprises européennes.

Globalement, nous sommes en accord avec le besoin de mettre en place progressivement une règlementation de l'IA, mais dont le contenu s'adapte au niveau de risque encouru par les citoyens européens à l'usage de la technologie et à la portée en termes d'utilisateurs de cette technologie. Nous nous interrogeons sur la nécessité d'établir le risque ex ante. En effet si certains cas d'usage peuvent être proscrits sur des considérations éthiques, c'est bien le contexte de l'utilisation d'une technologie qui en définit le risque réel. Faire porter de lourdes contraintes a priori peut mener simplement à un déplacement de l'innovation en dehors de l'union européenne.

En synthèse

- il est nécessaire de réglementer,
- la vision par le risque est pertinente,
- la réglementation doit s'appliquer à tous les fournisseurs d'IA y compris les « petits fournisseurs »,
- une réglementation ex ante ne semble pas applicable de manière générale en particulier sur un périmètre aussi large que celui présenté dans la proposition,
- la réglementation englobe de façon trop vaste tout type de système d'IA et de technologie, ce qui rend complexe et trop coûteuse sa mise en pratique.

La réglementation de l'IA ne doit pas être un frein à l'innovation ou limiter la compétitivité des entreprises.

PRECISION DES TERMES & PERIMETRE D'APPLICATION

DEFINITION DE L'IA

Dans le secteur bancaire, de nombreuses solutions en production sont portées par des solutions classiques de règles ou de statistiques et pourraient donc dorénavant, au vu de la définition très extensive de l'IA (annexe 1), être soumises à la réglementation. Si au sein des banques, nous avons une démarche de recensement des solutions et algorithmes IA, nous n'avons pas identifié la nécessité de couvrir aussi largement le domaine de l'IA du point de vue des attentes des régulateurs nationaux qui, eux, s'intéressent plus particulièrement aux IA de type « apprentissage automatique » dont la mise en œuvre peut en effet entraîner des modifications des processus d'audit actuellement en place. En revanche, élargir le domaine de l'IA à des domaines plus anciens (b et c de l'annexe) reviendrait à faire porter de nouvelles contraintes à des applications déjà largement utilisées.

La consultation lancée en février 2020 avait d'ailleurs fait apparaître une demande dans le même sens (exposé des motifs § 3.1) :

> *« Les parties intéressées ont pour la plupart demandé que l'IA soit définie de manière stricte, claire et précise »*

Recommandation 1 : restreindre la définition de l'IA en annexe 1 au § a)

SYSTEME IA A HAUT RISQUE

On doit noter que « système IA » n'est jamais précisément défini. Dans la section 1.1 (Exposé des motifs), les objectifs de la proposition de réglementation indiquent vouloir :

> *« … veiller à ce que les systèmes d'IA mis sur le marché de l'Union et utilisés soient sûrs et respectent la législation en vigueur en matière de droits fondamentaux et les valeurs de l'Union »*

Les systèmes IA à risque sont ensuite définis « en extension » dans le Titre III article 6 et l'annexe III :

> *« les systèmes d'IA visés à l'annexe III sont … considérés comme à haut risque »*

Un **système IA** serait donc un système contenant un composant IA (IA défini en Annexe I), mis sur le marché et utilisé sur le marché européen.

Cependant, dans l'introduction (11), un usage « *secondaire* » est indiqué : si un système IA n°1 non mis sur le marché (par exemple développé dans un pays non européen ou également pour un usage interne) est ensuite utilisé pour produire des résultats par un système (IA ou non d'ailleurs) n°2 qui, lui, est mis sur le marché, alors le risque du système IA n°1 doit être évalué :

> *« Compte tenu de leur nature numérique, certains systèmes d'IA devraient relever du présent règlement même lorsqu'ils ne sont ni mis sur le marché, ni mis en service, ni utilisés dans l'Union »*

Cet usage secondaire s'applique-t-il exclusivement aux systèmes IA n°1 développés en dehors de la Communauté, ou également aux solutions IA développées en interne dans les entreprises ? Dans ce second cas, la vérification de la conformité ex ante pourrait s'avérer très difficile et coûteuse.

Une définition précise de « système IA » devrait donc être donnée en limitant explicitement l'usage secondaire (11) aux systèmes développés hors CE et excluant les usages internes aux sociétés, la clé étant bien toujours la finalité de l'usage (« destination » définie à l'article 3) et le périmètre d'utilisation (en termes d'usages et d'utilisateurs finaux). On doit demander que la finalité du système IA soit spécifiée de façon beaucoup plus précise, par exemple sous la forme d'une liste d'usages prévus.

> *«destination», l'utilisation à laquelle un système d'IA est destiné par le fournisseur, y compris le contexte et les conditions spécifiques d'utilisation, telles que précisées dans les informations communiquées par le fournisseur dans la notice d'utilisation, les indications publicitaires ou de vente, ainsi que dans la documentation technique;*

Recommandation 2 : préciser la définition de système IA

Un **système IA à haut risque** est un système IA utilisé pour l'un des usages cités en annexe III, donc nécessairement un système mis sur le marché européen. Il serait intéressant de définir des « critères de risque » (santé, sécurité des personnes, droits fondamentaux, etc.) sur lesquels les usages prévus pourraient être évalués, ce qui préfigurerait le système de gestion des risques décrit à l'article 9. Un même système IA, selon l'usage particulier visé (et faisant partie de sa finalité déclarée), pourrait être défini « à haut risque » avec plus de finesse selon les critères de risque impactés.

La proposition de mise en conformité existante étant largement prescriptive, avec un périmètre très large, implique un impact très fort sur la compétitivité. L'impact économique évalué au § 3.3. Analyse d'impact de l'exposé des motifs dépasse très largement le simple coût d'obtention du label CE, puisqu'il faut lui adjoindre la mise en place des éléments prévus au titre III :

> *Les coûts de mise en conformité correspondants sont estimés entre 6 000 EUR et 7 000 EUR pour la fourniture d'un système d'IA à haut risque moyen d'une valeur d'environ 170 000 EUR d'ici 2025*

On peut estimer les coûts de mise en place à 60 à 100 k€ pour 6 k€ de coût d'obtention de label. Il est donc important de limiter ce risque majeur pour la compétitivité européenne. Nous proposons d'ajuster cette approche par les risques, de sorte de définir deux niveaux, définis selon l'analyse des critères de risque associés aux cas d'usage de la finalité déclarée du système IA :

- Une liste très limitée et établie selon un processus défini au préalable de cas d'usage nécessitant l'application de la régulation telle que définie à la lettre,
- Une autre liste pour laquelle la réglementation érige des principes de bonne pratique à suivre tout en laissant aux acteurs économiques la flexibilité de les mettre en œuvre en accord avec leurs processus internes.

Il serait nécessaire d'élaborer une analyse d'impact sur les coûts (et la perte de bénéfice des secteurs « à haut risque ») plus complète que celle indiquée au §3.3.

Recommandation 3 : dans les domaines à haut risque, préciser deux sous-catégories de système IA à haut risque permettant de différencier deux niveaux de mise en œuvre de la conformité. Préciser les impacts économiques complets des secteurs à haut risque concernés.

Par ailleurs, la définition en extension du périmètre de tels systèmes pourrait faire apparaître un risque juridique si l'usage visé n'apparait pas dans la liste au moment de la mise sur le marché, mais que celle-ci est ensuite modifiée. L'Article 7 (Modifications de l'annexe III) semble indiquer que la liste des domaines de l'annexe III (points 1 à 8) ne peut pas être modifiée :

> *« La Commission est habilitée à adopter des actes délégués conformément à l'article 73 afin de mettre à jour la liste figurant à l'annexe III en y ajoutant des systèmes d'IA à haut risque lorsque les deux conditions suivantes sont remplies:*
> *(a) les systèmes d'IA sont destinés à être utilisés dans l'un des domaines énumérés à l'annexe III, points 1 à 8;*
> *(b) les systèmes d'IA présentent un risque de préjudice pour la santé et la sécurité, ou un risque d'incidence négative sur les droits fondamentaux, qui, eu égard à sa gravité et à sa probabilité d'occurrence, est équivalent ou supérieur au risque de préjudice ou d'incidence négative que présentent les systèmes d'IA à haut risque déjà visés à l'annexe III »*

Un usage ne figurant pas dans la liste des « domaines énumérés à l'annexe III » ne peut donc pas y être intégré ensuite. Cependant s'il y figure et que la liste de l'Annexe III est donc modifiée, comment les acteurs seront-ils notifiés ? la mise en conformité sera-t-elle être rétroactive ? dans quels délais ?

De même, si nous restons dans la perspective d'un établissement des risques ex ante fondé sur des cas d'usage, quels sont les processus qui seront mis en place afin de réévaluer au fil du temps le niveau de risque des cas d'usage ? Le recours aux « actes délégués » est très imprécis et leur périmètre devrait être très clairement encadré.

Recommandation 4 : préciser les règles de modification de l'Annexe III pour les systèmes IA à haut risque par le biais d'actes délégués.

PROBLEMES CONCRETS POUR L'IMPLEMENTATION DE LA MISE EN CONFORMITE

Nous avons identifié de nombreux cas où l'implémentation pratique de la mise en conformité semble extrêmement coûteuse, voire impossible. Notre crainte repose principalement sur la mise en place d'une réglementation qui sera vue comme un frein à l'innovation permise par l'IA, compte tenu de ces contraintes de mise en conformité. Par ailleurs, nous ne voyons pas comment il serait possible d'appliquer la mise en conformité sur certains systèmes déjà en production mais pour lesquels les exigences requises ne peuvent être rétroactivement mises en œuvre. Pour chaque contrainte de mise en conformité devrait être conduite une analyse d'impact sur l'innovation.

Nous listons ici quelques cas qui nous apparaissent critiques.

MISES A JOUR FREQUENTES

Une fois qu'une solution comprenant un composant d'apprentissage a été mise sur le marché, elle doit être monitorée pour contrôler ses performances. Habituellement, les données dérivent peu à peu et leur distribution

s'écartant de la distribution des données d'apprentissage, les performances vont finir par se dégrader. On prévoit donc toujours des procédures de réapprentissage des modèles, la plupart du temps non automatiques, mais en général sur levée d'alertes (il existe cependant des systèmes auto-apprenants, qui vont déclencher automatiquement ces réapprentissages). Quand ce réapprentissage est effectué, une modification est apportée au système IA d'origine, mais alors faut-il le revalider ? Il est indiqué au § 5.2.3 :

> *« En cas de modification substantielle des systèmes d'IA, ceux-ci devront faire l'objet de réévaluations ex ante de la conformité »*

La question de la définition précise de « **modification substantielle** » est alors cruciale, puisque si le réapprentissage peut être considéré comme apportant une modification substantielle, le **réapprentissage au fil de l'eau devient impossible**, ce qui de fait va limiter considérablement l'intégration d'approche IA innovantes dans les systèmes.

Cependant la définition du terme dans le document n'est pas assez claire :

> *« (23) « modification substantielle», une modification apportée au système d'IA à la suite de sa mise sur le marché ou de sa mise en service, qui a une incidence sur la conformité de ce système avec les exigences énoncées au titre III, chapitre 2, du présent règlement ou entraîne une modification de la destination pour laquelle le système d'IA a été évalué »*
> *« (66) Conformément à la notion communément établie de modification substantielle pour les produits réglementés par la législation d'harmonisation de l'Union, il convient que les systèmes d'IA fassent l'objet d'une nouvelle évaluation de la conformité chaque fois qu'ils subissent une modification susceptible d'avoir une incidence sur leur conformité avec le présent règlement ou que la destination du système change »*
> *Nous recommandons d'ajuster la définition d'une modification substantielle en l'orientant par l'usage du système, ses performances ou la modification de son niveau de risque. Cela faciliterait la mise en œuvre de la réglementation pour le réapprentissage qui ne vient pas nécessairement modifier la finalité du système, mais au contraire, améliorer les résultats obtenus.*

Recommandation 5 : préciser la définition de « modification substantielle »

La procédure indiquée dans le document pour intégrer le réapprentissage dans la mise en conformité est la suivante :

- Décrire les modifications anticipées dans la documentation technique :

> *« Article 13 Les systèmes d'IA à haut risque sont accompagnés d'une notice d'utilisation … comprenant … (b) les caractéristiques, les capacités et les limites de performance du système d'IA à haut risque, …*
> *(c) les modifications du système d'IA à haut risque et de ses performances qui ont été prédéterminées par le fournisseur au moment de l'évaluation initiale de la conformité, le cas échéant »*

- Si les modifications apportées ont été décrites, alors elles ne constituent pas une modification substantielle :

> *« Article 43, 4. … Pour les systèmes d'IA à haut risque qui continuent leur apprentissage après avoir été mis sur le marché ou mis en service, les modifications apportées au système d'IA à haut risque et à ses performances qui ont été déterminées au préalable par le fournisseur au moment de l'évaluation initiale de la conformité et font partie des informations contenues dans la documentation technique visée à l'annexe IV, point 2 f), ne constituent pas une modification substantielle »*

Mais si les modifications vont au-delà de ce qui a été décrit :

> *« 5.2.3. En cas de modification substantielle des systèmes d'IA, ceux-ci devront faire l'objet de réévaluations ex ante de la conformité (notamment lorsque les modifications vont au-delà de ce qui est prédéterminé par le fournisseur dans sa documentation technique et vérifié lors de l'évaluation ex ante de la conformité) »*

Pour que les réapprentissages réguliers soient possibles, la documentation technique initiale doit donc adopter une description « large » des modifications possibles : comment ces limites seront-elles appréciées par l'organisme certificateur ? Si la modification est considérée comme substantielle, quels seront les délais pour la mise en conformité ? Si le réapprentissage nécessite une mise en conformité systématique il est à craindre une perte de performance des systèmes d'IA intégrant du réapprentissage qui ne seront alors plus mis à jour aussi régulièrement. Ces pertes de performance se traduiront par des coûts en augmentation, qui, in fine, seront supportés par les utilisateurs finaux.

Une autre approche classique en IA, pour les systèmes d'apprentissage, consiste à développer le système dont on évalue les performances techniques, puis, quand celles-ci sont satisfaisantes, l'évaluation « sur le terrain » pour valider que ces performances techniques se traduisent bien en performances métier. Après itérations, on décide alors de mettre sur le marché. Comment peut-on évaluer sur le marché la solution pour arriver à la solution finale ? Est-ce que ce type d'approche entre dans le champ de la réglementation si elle est utilisée pour la mise en œuvre d'un système à haut risque ? ou bien ces évaluations sont-elles considérées comme des tests au sens de l'article 9 : Les tests des systèmes d'IA à haut risque sont effectués, selon les besoins, à tout moment pendant le processus de développement et, en tout état de cause, avant la mise sur le marché ou la mise en service. Les tests sont effectués sur la base de métriques et de seuils probabilistes préalablement définis, qui sont adaptés à la destination du système d'IA à haut risque.

De même, comment certifier des approches de type « A/B testing » ? Comment les certifier ex-ante alors que la solution finale n'est pas encore totalement définie ? Est-ce que ce type d'approche entre dans le champ de la réglementation si elle est utilisée pour la mise en œuvre d'un système à haut risque ?

Recommandation 6 : préciser le processus à suivre en cas de réapprentissage

ALGORITHMES PRE-ENTRAINES PAR TIERS

De plus en plus, les composants IA sont développés en utilisant des « solutions de développement » comme une plateforme de développement de solutions IA (e.g. Watson d'IBM, Palantir, DataRobot, Dataiku …), des bibliothèques d'algorithmes open source (e.g. PyTorch, scikitlearn …) ou des modèles pré-entraînés disponibles en open source (e.g. yolov3, FaceNet, BERT et dérivés …). Si une entreprise développe un système IA à haut risque en utilisant de telles solutions de développement, on doit considérer (Article 28) qu'il est le fournisseur du système IA à haut risque et que, comme il a apporté une « modification substantielle » à la solution de développement (c) ci-dessous) le fournisseur n'est plus assujetti aux contraintes de conformité :

> *Tout distributeur, importateur, utilisateur ou autre tiers est considéré comme un fournisseur aux fins du présent règlement et est soumis aux obligations incombant au fournisseur au titre de l'article 16 dans toutes les circonstances suivantes:*
> *(a) il met sur le marché ou met en service un système d'IA à haut risque sous son propre nom ou sa propre marque;*
> *(b) il modifie la destination d'un système d'IA à haut risque déjà mis sur le marché ou mis en service;*
> *(c) il apporte une modification substantielle au système d'IA à haut risque.*
> *2. Lorsque les circonstances visées au paragraphe 1, point b) ou c), se produisent, le fournisseur qui a initialement mis sur le marché ou mis en service le système d'IA à haut risque n'est plus considéré comme un fournisseur aux fins du présent règlement.*

Cependant, les conditions de l'Article 16 (Obligations incombant aux fournisseurs de systèmes d'IA à haut risque) peuvent devenir impossibles à remplir : par exemple, pour l'utilisation de systèmes pré-entraînés sur des jeux de données qui ne sont pas disponibles. Doit-on considérer qu'un réentrainement partiel est une modification « substantielle » ?

Dans ce cas, les fournisseurs de ces solutions de développement devraient-ils eux-mêmes être évalués ex ante ? Vérifier la conformité de ces solutions de développement ne peut être à la charge des producteurs « secondaires », mais les grandes entreprises du numérique, sources des solutions de développement, accepteront-elles de faire certifier leurs solutions ? Il conviendrait donc de préciser ces cas.

Comment traiter alors les systèmes d'IA qui s'appuient sur des solutions IA génériques dont l'usage final ni le niveau de risque ne peuvent être établis a priori a ? Comment se met en place le processus de certification ? Quel poids supplémentaire va porter sur les fournisseurs de ces solutions génériques pour permettre aux intégrateurs de les exploiter dans le respect des exigences de certification ?

Par ailleurs, comment l'exigence de conformité portant sur la mise à disposition du code source et des données (Article 64) peut-elle être mise en œuvre lorsqu'un composant IA tiers est intégré dans le système IA final ? En effet, il parait impossible d'exiger l'accès au code source d'un fournisseur tiers portant en particulier une solution propriétaire.

Recommandation 7 : préciser les conditions de certification de système IA à haut risque comprenant un composant IA exploitant une solution de développement provenant d'un tiers

BIAIS

Les systèmes IA utilisant des techniques d'apprentissage doivent utiliser des jeux de données satisfaisant des critères de qualité, notamment en matière de biais (Articles 10, 14 et 15) :

> *« Les jeux de données d'entraînement, de validation et de test sont assujettis à des pratiques appropriées en matière de gouvernance et de gestion des données. Ces pratiques concernent en particulier … :*
> *(f) un examen permettant de repérer d'éventuels biais*
> *(g) la détection d'éventuelles lacunes ou déficiences dans les données, et la manière dont ces lacunes ou déficiences peuvent être comblées.*
> *Les jeux de données d'entraînement, de validation et de test sont pertinents, représentatifs, exempts d'erreurs et complets. »*
> *« Les mesures prévues … donnent aux personnes chargées d'effectuer un contrôle humain, en fonction des circonstances, la possibilité …*
> *(b) d'avoir conscience d'une éventuelle tendance à se fier automatiquement ou excessivement aux résultats produits par un système d'IA à haut risque («biais d'automatisation »)… »*
> *« Les systèmes d'IA à haut risque qui continuent leur apprentissage après leur mise sur le marché ou leur mise en service sont développés de telle sorte que les éventuels biais dus à l'utilisation de résultats comme données d'entrée pour les opérations futures («boucles de rétroaction») fassent l'objet d'un traitement adéquat au moyen de mesures d'atténuation appropriées.*

Cependant, on sait que tous les jeux de données comprennent des erreurs et sont biaisés. Si on a pu détecter des lacunes (g) on les a évidemment comblées ! Quant aux biais, ils sont difficiles à identifier et potentiellement difficiles à supprimer. On peut donc tout au plus mettre en œuvre des « pratiques appropriées » et prévoir des « traitements adéquats ». Par ailleurs, il conviendrait de restreindre l'étude des biais aux variables sensibles (telles que définies dans le RGPD).

Recommandation 8 : préciser la notion de « biais » et la nature des exigences de prise en compte des biais

TRAÇABILITE

Pour les systèmes IA à haut risque, il faut mettre en place ex ante un processus de mise en conformité qui aboutira à la délivrance d'un label CE (§2.3 Exposé des motifs) :

> *Pour les systèmes d'IA à haut risque, les exigences en matière de données de haute qualité, de documentation, de traçabilité, de transparence, de contrôle humain, d'exactitude et de robustesse*

Ces exigences comprennent des contraintes d'enregistrement des événements (article 12) tout au long du cycle de vie du système IA à haut risque :

1. *La conception et le développement des systèmes d'IA à haut risque prévoient des fonctionnalités permettant l'enregistrement automatique des événements («journaux») pendant le fonctionnement de ces systèmes. Ces fonctionnalités d'enregistrement sont conformes à des normes ou à des spécifications communes reconnues.*
2. *Les fonctionnalités d'enregistrement garantissent un degré de traçabilité du fonctionnement du système d'IA tout au long de son cycle de vie qui soit adapté à la destination du système.*
3. *En particulier, ces fonctionnalités permettent de surveiller le fonctionnement du système d'IA à haut risque dans l'éventualité de situations ayant pour effet que l'IA présente un risque au sens de l'article 65, paragraphe 1, ou entraînant une modification substantielle, et facilitent la surveillance après commercialisation visée à l'article 61.*

Les fonctionnalités d'enregistrement automatique doivent être limitées : en effet, « tout » enregistrer aboutit rapidement à de très grandes quantités de données. La traçabilité devient rapidement extrêmement complexe à mettre en œuvre car elle va nécessiter une capacité de stockage importante sur toute la durée de vie du système d'IA. Il convient donc de limiter le périmètre d'enregistrement : selon quelles « normes ou à des spécifications communes reconnues » ?

Jusqu'où va la traçabilité : quelle est la profondeur de l'historique à conserver ? doit-on conserver les journaux de tous les tests et toute la base d'apprentissage initiale ? Comment la CE imagine-t-elle les cas où il est nécessaire de stocker la vie des données et des modèles de 20 millions de clients ? Cela ne semble pas faisable et probablement trop coûteux. Par exemple, sur le credit scoring, il y a un cadre de validation qui est déjà en place et lourd, mais moins lourd cependant que celui proposé par la CE. Par ailleurs, l'exigence de conserver les données sur de grandes profondeurs d'historique peut être en désaccord avec RGPD.

L'exigence de conservation exhaustive des « événements » aboutissant à une volumétrie trop importante est donc irréaliste.

Recommandation 9 : préciser les exigences d'enregistrement en termes de profondeur d'historique, périmètre des données, etc. ainsi que les « normes ou à des spécifications communes reconnues »

Le processus de mise en conformité comporte également des exigences d'accès aux données (Annexe VII et article 64) :

> *La documentation technique est examinée par l'organisme notifié. À cette fin, l'organisme notifié se voit accorder un accès complet aux jeux de données d'entraînement et de test utilisés par le fournisseur, y compris par l'intermédiaire d'interfaces de programmation ou d'autres moyens et outils appropriés permettant un accès à distance.*

Ces exigences sont à aligner avec les contraintes RGPD, et, de façon réaliste, être limitées exclusivement aux données du système en production. Par ailleurs, l'accès à distance aux données pose un risque de sécurité qui ne semble pas acceptable. En particulier, les systèmes « sensibles » (sécurité financière : fraude, blanchiment…) doivent être protégés en accès.

Recommandation 10 : restreindre les exigences d'accès aux données du système sur le marché, et éviter absolument les accès à distance.

TRANSPARENCE

Les exigences de transparence (Article 13), et notamment les informations contenues dans la notice d'information (§3) devraient être restreintes pour éviter les menaces liées aux attaques cyber ou concernant les systèmes « sensibles » (sécurité financière : fraude, blanchiment…).

Recommandation 11 : restreindre les exigences de transparence pour protéger la cybersécurité des systèmes IA, et plus particulièrement les systèmes sensibles (fraude, blanchiment).

EXCEPTIONS (INTERET LEGITIME)

La procédure de régulation proposée se considère comme complémentaire aux régulations déjà en place (§1.2 de l'exposé des motifs) :

> *Le caractère horizontal de la proposition requiert une cohérence parfaite avec la législation de l'Union existante applicable aux secteurs dans lesquels des systèmes d'IA à haut risque sont déjà utilisés ou sont susceptibles de l'être dans un avenir proche.*

Cependant, il peut se poser des questions de hiérarchie des normes. Ainsi, dans le cas de certaines applications ou certains services proposés par les banques (lutte contre le blanchiment d'argent, lutte contre la fraude, devoir de diligence), des règlementations sont déjà en place, sont-elles alors considérées comme prioritaires par rapport à la réglementation IA de la CE ?

En particulier dans le document il est indiqué une « cohérence parfaite … » sans que pour autant une priorité spécifique soit exprimée. Dès lors, est-ce qu'il peut y avoir des exceptions « légitimes » liées aux contraintes légales en vigueur. Par exemple, sur certains cas d'usage il n'est pas légalement possible de stocker les données et les modèles, la démarche de transparence est complexe voire impensable, la mise à jour des modèles est faite de façon tellement régulière qu'il n'apparait pas possible de repasser dans le processus de régulation à chaque itération etc…

Recommandation 12 : préciser la hiérarchie des normes

EXEMPLES DE CAS D'USAGE IDENTIFIES A « HAUT RISQUE » (EXISTANTS OU POSSIBLES)SCORE DE CREDIT

Ce domaine concerne évidemment très fortement les banques (attendus, 37) :

> *« les systèmes d'IA utilisés pour évaluer la note de crédit ou la solvabilité des personnes physiques devraient être classés en tant que systèmes d'IA à haut risque »*

Plusieurs questions de périmètre se posent :

1. Est-ce que cela englobe tout type de crédit ou uniquement les crédits « essentiels » ? Le périmètre d'application devrait être défini plus précisément.
2. Ensuite, est-ce que cela inclut également le monitoring du crédit ou est-ce qu'on reste uniquement au niveau de l'octroi dès lors que le monitoring ne supprime pas de service essentiel ? Qu'en est-il des modèles d'évaluation de la solvabilité d'une personne qui peuvent ensuite être utilisés pour établir un score d'octroi ou pour suivre le niveau de risque associé à un prêt existant ?
3. Question sur l'emploi du terme « personne physique ». Il est fait mention ici de la personne physique, pas du citoyen. Dans le cas des sociétés unipersonnelles (personne morale) représentées par une seule personne physique, les auto-entrepreneurs, doit-on considérer que la demande de crédit de la personne morale est classifiée comme à haut risque ?
4. L'IA étant dans le document définie de manière très large/vaste, est-ce que des solutions basées sur des systèmes de règles métiers ou même parfois des calculs fait sur des fichiers Excel et qui font de l'évaluation de « solvabilité » vont rentrer sous le coup de la régulation ?

USAGES PAR LES AUTORITES REPRESSIVES

Les dispositifs basés sur l'Intelligence Artificielle mis en œuvre par les banques à ce titre peuvent concerner :

a) la lutte contre le blanchiment d'argent et le terrorisme,
b) lutte contre la fraude interne et externe,
c) devoir de diligence pour la connaissance client.

Ces systèmes pourraient être considérés comme des composantes de la sécurité (au sens de Art.6 1.) des opérations financières et semblent relever de l'Annexe III.6.g et seraient donc considérés comme à haut risque ou pourraient le devenir à la faveur d'une modification telle que prévue à l'article 7. Est-ce bien le cas ? Dans l'affirmative, les délais induits par les obligations prévues au chapitre 3 ou la transparence requise au Titre IV n'iront-elles pas à l'encontre des obligations réglementaires visées ? La clause 6.g peut-elle être précisée ?

Recommandation 13 : exclure explicitement les système IA dédiés à la sécurisation des opérations financières, à la lutte anti-fraude, anti blanchiment ou anti-terroriste de la liste des systèmes à haut risque.

BIOMETRIE

Les cas d'usage connus à date de l'usage de la biométrie sont potentiellement les suivants :

- Authentification client par reconnaissance faciale ou par reconnaissance d'empreinte digitale dans nos applications mobiles pour accéder à l'application et pour effectuer les opérations sensibles (enregistrement nouveau RIB, souscription produit)
- Authentification client dans les centres d'appels par reconnaissance d'empreinte vocale
- Authentification par biométrie dans les ATM (https://www.americanbanker.com/payments/news/how-biometric-atms-are-entering-mainstream-use)
- Authentification par empreinte digitale pour sécuriser des paiements directement sur les cartes de paiement (https://www.thalesgroup.com/fr/europe/france/dis/banque/cartes/biometrique-emv
- Authentification par empreinte digitale ou reconnaissance d'IRIS des collaborateurs (ou de prestataires) accrédités à pénétrer dans des locaux sensibles (type datacenter)
- Usage de la reconnaissance faciale pour comparer le visage du client à la photo de sa pièce d'identité lors du processus d'onboarding à distance (exemple fintech https://www.ubble.ai/fr/accueil/#contact-ubble)

Ces usages sont aujourd'hui majoritairement fondés sur des solutions externes à l'entreprise. Ainsi dans ces cas d'usage, les banques ne font qu'intégrer une solution existante au sein d'un programme de développement d'outil IT. Notre compréhension est que dans ces cas-là il n'appartient pas à la Banque de certifier le système IA, mais c'est le fournisseur de la solution qui doit obtenir la certification (voir Recommandation 7).

CONTACT

Hub France IA

8-10 rue Charles V, 75004 Paris.

E-mail : contact@hub-franceia.fr