CENTRE DE DROIT
BANCAIRE ET FINANCIER

Systèmes d'intelligence artificielle

Les catégories de la règlementation européenne

Par Yannick Caballero Cuevas le 16 décembre 2024

Le règlement européen sur l'intelligence (RIA) adopte une approche axée sur les risques qu'un système d'intelligence artificielle (SIA, cf. Caballero Cuevas, cdbf.ch/1382) peut poser pour la santé, la sécurité et les droits fondamentaux des individus. Les SIA sont divisés en quatre catégories, respectivement les SIA présentant un risque inacceptable, les SIA à haut risque, les SIA présentant un risque limité et les SIA présentant un risque minime. Le présent commentaire se concentre sur les trois premières catégories de SIA.

A. Les SIA présentant un risque inacceptable

L'art. 5 RIA interdit les SIA dont la mise sur le marché, la mise en service ou l'utilisation présentent un risque inacceptable. Le paragraphe 1 prévoit une liste exhaustive des pratiques interdites. À titre d'exemples, nous pouvons mentionner :

- les SIA de crédit social (let. c) ;
- les SIA évaluant de manière préventive le risque qu'une personne commette une infraction pénale (let. d) ;
- les SIA inférant les émotions d'une personne physique sur le lieu de travail et dans les établissements d'enseignement (let. f) ;
- les SIA de catégorisation individuelle biométrique qui permettent de déduire notamment la race, l'opinion politique, l'affiliation syndicale, l'orientation sexuelle d'une personne physique (cf. let. g) ;
- les SIA de surveillance biométrique en temps réel dans des espaces publics à des fins répressifs (let. h) ;
- les SIA visant à manipuler les comportements humains afin de priver les personnes physiques de leur libre arbitre (let. a).

Malgré l'interdiction formelle, le RIA prévoit des portes dérobées (*backdoors*) qui permettent, à certaines conditions, l'utilisation de SIA présentant un risque inacceptable. C'est notamment le cas pour les SIA de surveillance biométrique en temps réel, lorsqu'ils sont, par exemple, utilisés pour rechercher des victimes d'enlèvement, de traite d'êtres humains ou d'exploitation sexuelle ainsi que pour prévenir une menace spécifique telle qu'une attaque terroriste. Les conditions de leur utilisation sont prévues aux paragraphes 2 et suivants de l'art. 5 RIA.

À l'heure actuelle, aucun SIA utilisé dans le domaine bancaire et financier ne présente des risques inacceptables. En pratique, l'art. 5 RIA ne devrait donc pas avoir d'incidence sur les

banques et les établissements financiers.

B. Les SIA à haut risque

Les exigences règlementaires du RIA s'appliquent essentiellement aux SIA à haut risque (art. 8 à 15 RIA). L'art. 6 par. 1 RIA prévoit qu'un SIA doit être considéré comme étant à haut risque lorsque les deux conditions suivantes sont cumulativement remplies :

- Premièrement, le SIA est destiné à être utilisé comme composant de sécurité d'un produit couvert par une directive ou un règlement mentionné dans l'Annexe 1 RIA, ou que le SIA constitue lui-même un tel produit. Parmi les législations visées, nous pouvons citer la directive sur les machines, la directive sur les ascenseurs et les composants de sécurité des ascenseurs, le règlement sur les dispositifs médicaux ou encore le règlement sur l'instauration de règles communes dans le domaine de la sûreté de l'aviation civile.
- Deuxièmement, le produit, dont le SIA est un composant de sécurité ou un produit en lui-même, est soumis à une évaluation de conformité par un tiers en vue de la mise sur le marché ou de la mise en service de ce produit conformément aux législations visées à l'Annexe 1 RIA.

Le paragraphe 2 renvoie à l'Annexe 3 RIA qui prévoit une liste exhaustive d'applications à haut risque. Dans le domaine bancaire et financier, une attention particulière devrait être portée aux SIA suivants :

- les SIA utilisés dans la sélection ou le recrutement de candidat (cf. Hirsch, cdbf.ch/1384) ;
- les SIA utilisés dans la prise de décision relative aux conditions des relations professionnelles, la promotion ou le licenciement ;
- les SIA utilisés pour évaluer la solvabilité des personnes physiques ou pour établir leur note de crédit (*credit scoring*) ; et
- les SIA utilisés pour évaluer les risques et la tarification concernant les personnes physiques en matière d'assurance-vie et d'assurance maladie.

Le RIA prévoit, une nouvelle fois, des portes dérobées. En effet, un SIA mentionné dans l'Annexe 3 RIA peut être considéré comme n'étant pas à haut risque lorsqu'il ne présente pas « de risque important de préjudice pour la santé, la sécurité ou les droits fondamentaux des personnes physiques, y compris en n'ayant pas d'incidence significative sur le résultat de la prise de décision » (cf. art. 6 par. 3 RIA). L'appréciation du risque est laissée au fournisseur du SIA. Ce dernier doit ainsi documenter précisément son évaluation avant la mise sur le marché ou en service s'il considère qu'un tel SIA n'est pas à haut risque conformément à l'art. 6 par. 3 RIA. La Commission européenne devra fournir, d'ici le 2 février 2026, des lignes directrices précisant la mise en œuvre de l'art. 6 RIA.

Les SIA considérés comme à haut risque pourront évoluer à l'avenir en fonction des progrès technologiques et de leurs usages. À cette fin, l'art. 7 RIA prévoit que la Commission est autorisée de modifier l'Annexe 3 RIA.

C. Les systèmes présentant un risque limité

Le RIA ne définit pas les SIA présentant un risque limité. Néanmoins, il instaure des obligations de transparence (cf. art. 50 RIA) pour certains types de SIA qui ne sont ni des SIA présentant un risque inacceptable ni des SIA à haut risque. Ces exigences s'appliquent notamment aux SIA destinés à interagir directement avec des personnes physiques, ceux destinés à générer des contenus de synthèse de type audio, image, vidéo ou texte, ou encore ceux destinés à générer ou manipuler des images ou des contenus audio ou vidéo constituant un hypertrucage (*deepfakes*). Selon nous, cette catégorie vise, avant tout, les agents conversationnels et les applications d'IA générative.

Les exigences de transparence visent principalement à signaler aux utilisateurs qu'il s'agit d'une conversation avec une IA, et de désigner le contenu généré par l'IA comme tel.

La qualification d'un SIA comme présentant un risque limité ne dispense pas d'analyser si le modèle d'IA qu'il utilise est un modèle d'IA à usage général (cf. Caballero Cuevas, cdbf.ch/1382). Ainsi, le fournisseur d'un agent conversationnel utilisant un modèle d'IA à usage général devrait notamment être amené à respecter les exigences de transparence (cf. art. 50 RIA), ainsi que celles relatives aux modèles d'IA à usage général (cf. art. 53 RIA, art. 55 RIA).

À notre avis, les banques et établissements financiers pourraient être tenus de se conformer aux exigences de transparence prévues par l'art. 50 RIA, en fonction de leur qualification en tant que fournisseurs ou déployeurs de SIA.

Reproduction autorisée avec la référence suivante: Yannick Caballero Cuevas, Les catégories de la règlementation européenne, publié le 16 décembre 2024 par le Centre de droit bancaire et financier, https://cdbf.ch/1390/