Gouvernance responsable de l'IA : Comment la norme ISO/IEC 42001 : 2023 permet-elle de répondre aux exigences imposées aux Système d'intelligence artificielle à haut risque par l'IA ACT ?

Système d'Intelligence Artificielle à haut risque

Eléonore SCARAMOZZINO, Avocate

SOMMAIRE

Gouvernance responsable de l'IA : Comment la norme ISO 42001 permet-elle de répondre aux exigences imposées aux Système d'intelligence artificielle à haut risque par le règlement sur l'intelligence artificielle ?

Système d'Intelligence Artificielle à haut risque

Un système d'IA est défini par le règlement sur l'intelligence artificielle (règlement 2024/1689 du 13 juin 2024) comme 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* » (article 3.1)

Le règlement fondé sur une approche par les risques, distingue quatre catégories de systèmes d'intelligence artificielle, ceux présentant des risques inacceptables pour les droits fondamentaux et qui sont interdits, depuis le 5 février 2025, les Systèmes d'intelligence artificielle à haut risque (SIAHR), qui sont au centre de la régulation sur l'IA, les SIA évalués à faibles risques, guidés par un principe de transparence (art 52)et enfin, les SIA présentant un risque minime, qui ne sont quant à eux pas régulés.

Les SIAHR sont régis par les articles 6-49 (sur un total de 113 articles) du Règlement (IA Act). Ils doivent respecter des exigences, dont notamment la gestion des risques, la gouvernance des données et la transparence. Comme pour le règlement général des données (RGPD, Règlement UE 2016/698), une

documentation décrivant les mesures adoptées pour être compliance aux exigences de l'IA Act est obligatoire.

L'application de ce règlement est progressive. En effet, il sera effectif le 2 aout 2026 pour les SIAHR de l'annexe II, le 2 aout 2027, pour les systèmes relevant de réglementation harmonisée comme les dispositifs médicaux embarquant une IA, et notamment les robots chirurgicaux, et enfin, à compter du 31 décembre 2030 pour les SIAHR visés à l'annexe X (système d'information à grande échelle mis sur le marché avant le 1er aout 2027). Au vu du niveau d'exigences réglementaires, les entreprises doivent se préparer à la mise en conformité avec cette réglementation. Cependant, les normes d'application sont toujours en cours d'élaboration par le Comité européen de normalisation et le Comité européen de normalisation en électronique et en électrotechnique pour obtenir le marquage CE. Bien que ces normes harmonisées ne soient pas obligatoires pour être compliance avec le réglement, leur respect fait naître néanmoins une présomption de conformité.

Dans son rapport[1] intitulé « ***EU AI ACT ISO/IEC 42001 : 2023, A guide to Implementation March 2025*** », AI & Partners[2] analyse comment la norme ISO/IEC12001:2023 permet de répondre à certaines exigences de l'AI Act, et notamment pour gérer les risques liés à l'IA. Le document présente la norme ISO en tant que première norme mondiale pour le management de l'intelligence artificielle, instaurant un **cadre pour la gouvernance, la transparence et la gestion des risques**.

Avant de présenter la norme ISO /IEC 42001 : 2023 au regard des dispositions de l'IA Act pour les SIAHR (II), il convient de rappeler brièvement ces exigences (I).

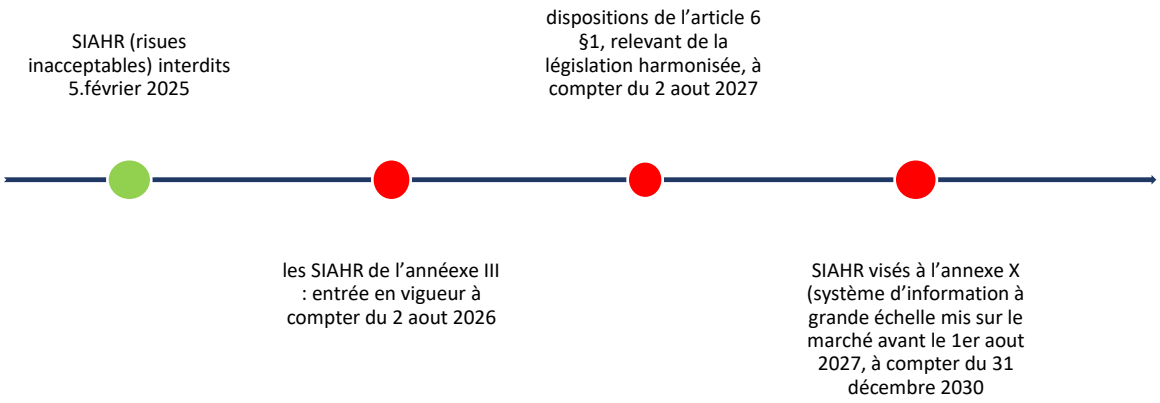
[1] Home | AI & Partners

[2] *AI & Partners defends and extends the digital rights of users at risk around the world. By combining direct technical support, comprehensive policy engagement, global advocacy, grassroots professional services, regulatory interventions, and participating in industry groups such as AI Commons, we fight for fundamental rights in the artificial intelligence age*.

I-Pyramide des risques de l'IA ACT

Entrée en vigueur progressive

Le règlement sur les SIA ou « *Artificial Intelligence Act* » (AI Act) entre en application en aout 2026, puis en 2027 et 2030, sauf pour les SIA présentant des risques inacceptables, dont leur application est entrée en vigueur le 5 février 2025.



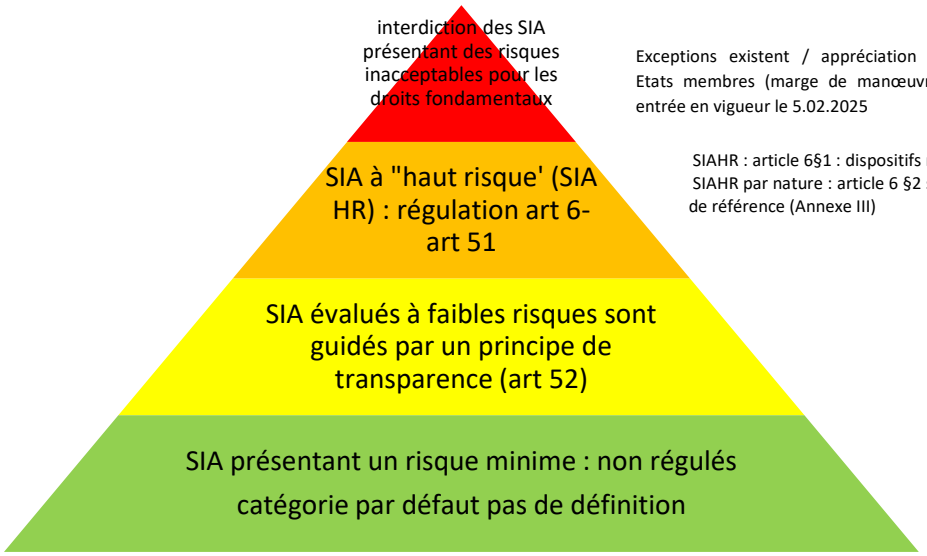
Pour utiliser une IA de confiance et respectueuse des droits fondamentaux de l'UE, ou représentant les risques les plus faibles, le règlement a opté pour une approche fondée sur les risques générés par la technologie en cause, couplée par un cadre souple des codes de conduite volontaires pour les SIA à risques faibles pour les droits fondamentaux. Ce qui est régulé est l'usage et non la technologie, en se fondant sur une pyramide des risques, fondée sur les usages de l'IA.

L'annexe I énumère trois catégories de techniques algorithmiques, en substance

- les systèmes auto-apprenants (machine learning) ;
- les systèmes logiques ;
- les systèmes statistiques.

Sont donc exclus du champ d'application :

- les systèmes d'IA développés ou utilisés exclusivement à des fins militaires sont exclus (art. 2, § 3) ;
- l'utilisateur d'un système d'IA à des fins personnelles et non-professionnelles (art. 3, § 4).



Les SIA à faibles risques

Pour les SIA à faibles risques il est prévu une obligation de transparence vis-à-vis des utilisateurs (art. 52).
Trois hypothèses sont envisagées les SIA :

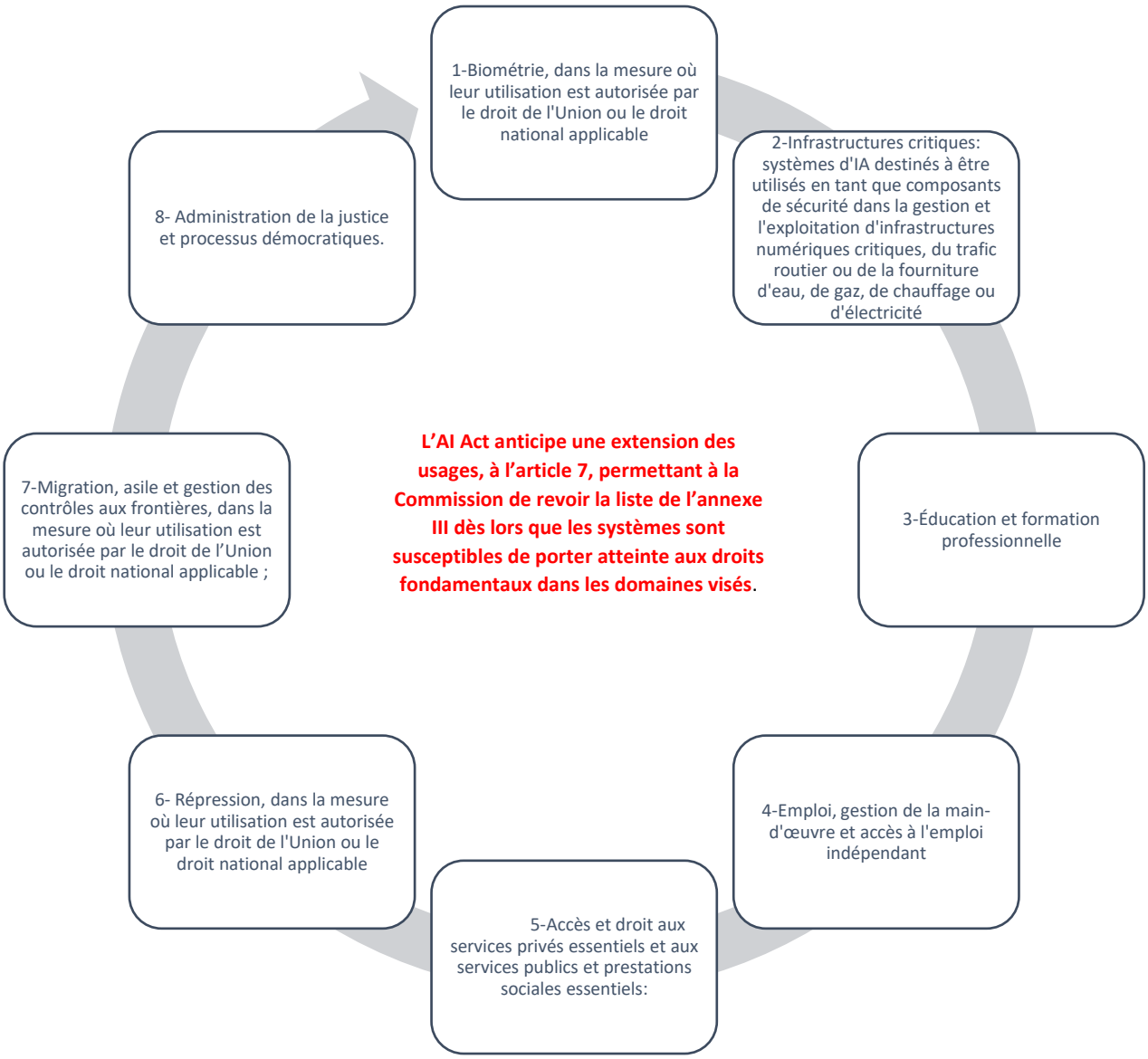
- sont destinés à interagir avec une personne (chatbots),
- ont pour finalité la reconnaissance d'émotion ou la catégorisation biométrique ;
- génèrent ou manipulent des contenus (deep fakes). L'utilisateur doit ici connaître le lien avec le système d'IA, l'objectif étant de limiter les manipulations possibles.

L'obligation n'a cependant pas vocation à s'appliquer aux systèmes autorisés par la loi et destinés à détecter, prévenir, enquêter et poursuivre les infractions pénales.

Les SIA à haut risque : SIAHR

La classification proposée est double (art. 6).

- **Qualification dépendant de conditions** : elle concerne tout système (autonome ou embarqué dans un produit) couvert par une des législations harmonisées de l'UE listées en annexe II et qui est soumis à une évaluation de conformité en vertu de ces mêmes textes.
- **Qualification automatique** : sont considérés comme systèmes à haut risque les systèmes d'IA visés à l'annexe III, qui identifie 8 domaines répertoriant les systèmes d'IA à haut risque au sens de l'article 6, paragraphe 2.



La régulation de ces SIA s'inscrit dans le RGPD, fondée sur une compliance (article 8) avec une documentation attestant cette conformité.

Principales Exigences à respecter par les SIAHR	une procédure continue de gestion des risques (art. 9) visant à prévenir les risques tout au long de son cycle de vie ;
	une politique contraignante de gouvernance des données utilisées pour l'entraînement, la validation ou encore les tests, afin de s'assurer de leur qualité et éviter les biais (art. 10) ;
	l'établissement d'une documentation technique dont le contenu est précisé l'annexe IV (art. 11) ;
	la tenue de fichiers de journalisation-enregistrement automatique des événements, à savoir des informations relatives à leur fonctionnement (art. 12) ;
	une obligation de transparence et d'information au bénéfice des déployeurs (art. 13) ;
	une surveillance humaine du système visant à prévenir ou minimiser les risques (art. 14) ;
	des obligations d'exactitude, de robustesse et de cybersécurité (art. 15) ;
	Des obligations incombant aux fournisseurs d'IA (art 16) ;
	Système de gestion de la qualité (art 17)

Interdiction des SIA présentant des risques inacceptables

L'article 5 du règlement interdit certaines pratiques dans l'utilisation des SIAHR. Ces systèmes qui sont au sommet de la pyramide des risques sont interdits, compte tenu des risques inacceptables qu'ils présentent au regard des droits fondamentaux (police prédictive, contrôle social pratiques de manipulation). Il s'agit d'une protection des individus face à l'innovation technologique.

Sont interdits les SIA visant à :

- manipuler les comportements des individus, conduisant à l'adoption de comportements indésirables ou des décisions non fondées sur leur liberté de choix ;
- conduire à des discriminations soient en exploitant leurs émotions, ou en les catégorisant sur la base de données biométriques pour déduire leurs opinions syndicales/ politiques, convictions religieuses/philosophiques, leur orientation sexuelle, sont prohibés ;
- établir un « social scoring », évaluer ou classifier les personnes en fonction de leur comportement social, de caractéristiques personnelles sont prohibés ;
- évaluer ou prédire le risque qu'une personne commette une infraction pénale, dès lors qu'elle se fonde uniquement sur le profilage de la personne ou de l'évaluation de ses traits de personnalité ou caractéristiques (nationalité, lieu de naissance ou de résidence, nombre d'enfants, niveau d'endettement, etc.). Il en résulte que la police prédictive se heurte au respect de la présomption d'innocence. Toutefois, cette interdiction est très restrictive et

permet l'utilisation de l'IA dans la police, notamment pour évaluer l'implication d'une personne dans une activité criminelle, sous certaines conditions ;

- utiliser des systèmes d'identification biométrique à distance en temps réel dans des espaces publics à des fins répressives, des bases de données de reconnaissance faciale, par le moissonnage non ciblé d'images faciales provenant d'internet ou de la vidéosurveillance. Cependant ces SIA peuvent être utilisés dans le respect de certaines conditions i) pour la recherche de personne, ii) la prévention d'une menace substantielle, et imminente pour la vie ou la sécurité physique de personnes ou d'une menace réelle et actuelle ou réelle et prévisible d'attaque terroriste et iii) la localisation ou l'identification d'une personne soupçonnée d'avoir commis certaines infractions pénales.

Cependant, il convient de souligner que les Etats membres bénéficient d'une marge de manœuvre considérable pour l'implémentation de cette liste d'interdiction de SIA présentant des risques inacceptables.

II-La gestion des risques de l'IA par la norme ISO/IEC 42001 :202

ISO/IEC 42001:2023 introduit la première norme mondiale de **système de management de l'IA (**AIMS), qui propose une approche systématique de la gestion des risques, de la conformité et de l'amélioration continue de l'IA. Cette norme est adaptée notamment pour les entreprises du secteur de la finance et de la santé, qui fournissent ou/et utilisent des SIAHR.

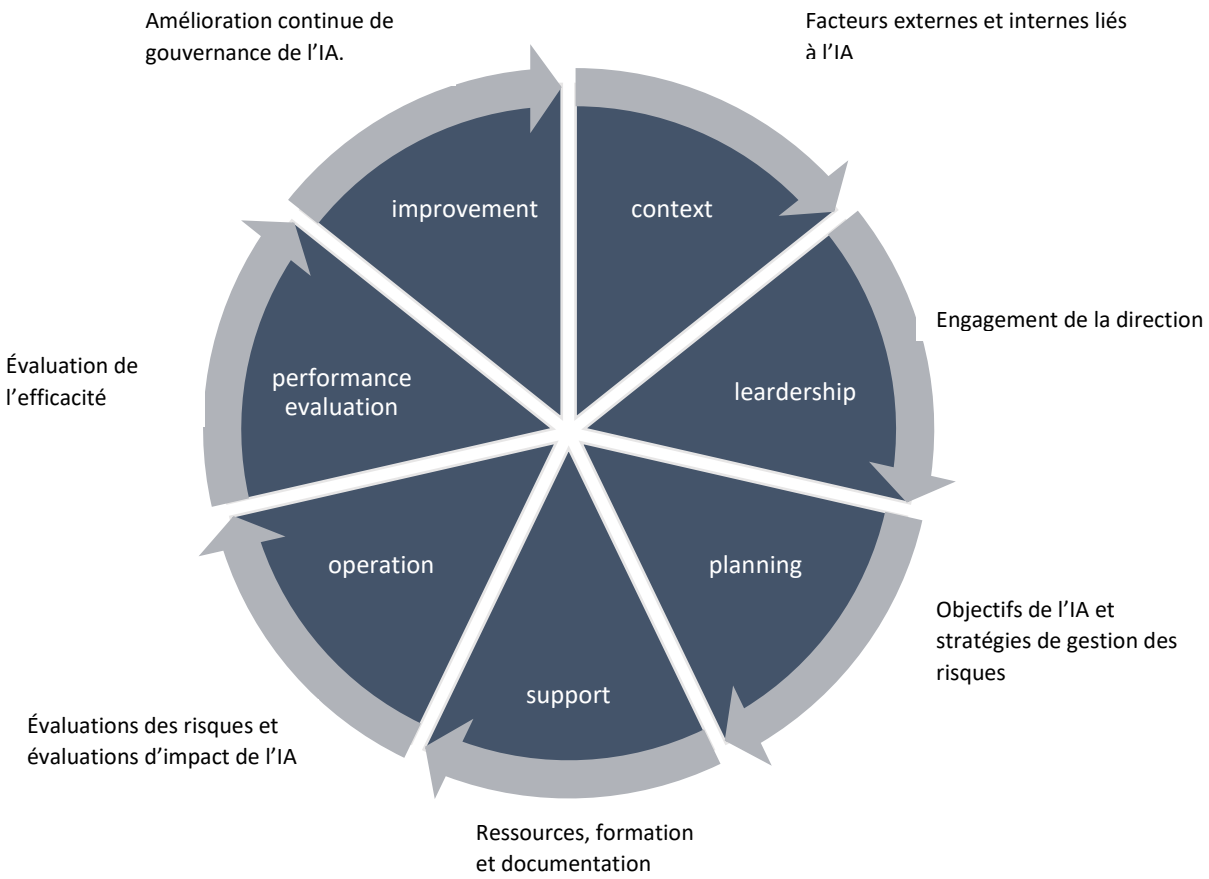
ISO/IEC 42001 établit un cadre structuré pour gérer les risques liés à l'IA, protéger les droits des utilisateurs et démontrer leur engagement vers des pratiques éthiques en matière d'IA.

Exigences de la norme ISO/IEC 42001:	une identification des risques potentiels liés à l'IA, tels que	les biais, les problèmes de sécurité des données, les conséquences imprévues,
les entreprises doivent réaliser	une mise en œuvre des contrôles pour y remédier ;	
	une surveillance continue des systèmes d'IA pour détecter les anomalies et assurer le respect des normes éthiques et juridiques ;	
	une évaluation d'impact de l'IA sur les individus et la société.	

La norme ISO 42001 s'appuie sur des pratiques de management des risques similaires à celles de l'ISO/IEC 23894:2023.

Les principales clauses obligatoires pour établir un système de management de l'IA

Structured approach for responsible AI development and deployment



ISO/IEC 42001 : une gestion responsable et éthique des risques liés à l'IA

Alors que des normes telles que ISO/IEC 27001 traitent de la sécurité de l'information et ISO/IEC 38507 se concentrent sur la gouvernance de l'IA pour les conseils d'administration, ISO 42001 fournit un cadre complet pour la gestion des risques liés à l'IA, les considérations éthiques et l'évaluation des performances.
La norme ISO 42001 en se concentrant spécifiquement sur les systèmes de management de l'IA, facilite la preuve :

- d'une gouvernance responsable de l'IA auprès des organismes de réglementation, des auditeurs et des parties prenantes.

- de la conformité au RGPD, en garantissant des évaluations et une documentation appropriées des risques liés à l'IA.

ISO/IEC 42001 intègre des principes éthiques dans la gouvernance de l'IA en promouvant

- l'équité,
- la transparence,
- la responsabilité.

La norme met l'accent sur la surveillance humaine, en veillant à ce que les systèmes d'IA ne fonctionnent pas sans considérations éthiques.

Comment les dispositions de la norme ISO /IEC 420001 : 2023 permettent-elles de répondre à certaines exigences de l'IA ACT ?

Exigence : Système de gestion de la qualité

L'article 17 du Règlement sur l'IA prévoit que **les fournisseurs de systèmes d'IA à haut risque** mettent en place un système de gestion de la qualité garantissant le respect du présent règlement. Ce système est documenté de manière méthodique et ordonnée sous la forme de politiques, de procédures et d'instructions écrites, et comprend au moins les aspects suivants:

- une stratégie de respect de la réglementation ;
- des techniques, procédures et actions systématiques destinées à la conception et au développement des systèmes d'IA à haut risque ainsi qu'au contrôle, à la vérification de cette conception et à l'assurance de la qualité ;
- des procédures d'examen, de test et de validation, des spécifications techniques ;
- les systèmes et procédures de gestion des données ;
- le système de gestion des risques (article 9) ;
- un système de surveillance après commercialisation (article 72) ;
- procédures relatives au signalement d'un incident grave (article 73) ;
- la gestion des communications avec les autorités compétentes ;
- les systèmes et procédures de conservation de tous les documents et informations pertinents;
- la gestion des ressources ;
- un cadre de responsabilisation définissant les responsabilités de l'encadrement et des autres membres du personnel

Système de gestion de qualité (art 17)	NORME ISO/IEEC 42001
documentation du système de qualité sous la forme de politiques, de procédures et d'instructions écrites	A.2.2.IA Policy : documenter une politique pour le développement ou l'utilisation de SIA
l'intégration de la gestion des SIA à d'autres politiques organisationnelles	A.2.3. Alignment with other organizational policies
examens réguliers du système de gestion de la qualité afin d'en garantir l'efficacité.	A.2.4-Review of the AI policy : La politique en matière d'IA est réexaminée à intervalles réguliers ou en complément si nécessaire afin de garantir sa pertinence, son adéquation et son efficacité
définition des rôles et des responsabilités dans le cadre du système de gestion de la qualité	A.3.2 AI roles and responsibilities : Les rôles et les responsabilités en matière d'IA doivent être définis et attribués en fonction des besoins de l'organisation.
Spécifications techniques et normes dans le cadre du système de gestion de la qualité	A.4.4 Tooling resources : documenter des informations sur les ressources d'outillage utilisées pour le système d'IA
management des ressources	A.4.5 System and computing resources : documenter les informations sur le système et les ressources informatiques utilisées pour le système d'IA.
affectation des ressources humaines dans le système de gestion de la qualité	A.4.6 Human resources
objectif de compliance et de développement responsable	A.6.1.2 Objectives for responsible development of AI system
nécessite de définir des processus pour la conception et le développement de systèmes d'IA.	A6.1.3 Processus de conception et de développement de systèmes d'IA fiables
procédures d'examen, de test et de validation.	A.6.2.5 AI system deployment
plan de déploiement dans le cadre du système de management de la qualité	L'organisation doit documenter un plan de déploiement et s'assurer que les exigences appropriées sont respectées avant le déploiement
procédure d'usage responsable	A.9.2 Processes for responsible use of AI systems
objectif d'usage responsable	A.9.3 Objectives for responsible use of AI system
documentation sur l'usage prévu.	A.9.4 Intended use of the AI system.
cadre des responsabilités de l'encadrement et des autres membres du personnel en ce qui concerne tous les aspects énumérés dans le présent paragraphe.	A.10.2 Allocating responsibilities
nécessite d'assurer l'alignement des fournisseurs sur les objectifs du système d'IA.	A.10.3 Suppliers
communication avec les clients dans le cadre du système de gestion de la qualité.	A.10.4 Customers : approche responsable du développement et de l'utilisation des SIA tient compte des attentes et des besoins de ses clients.

Exigence : Partage d'information sur les incidents graves

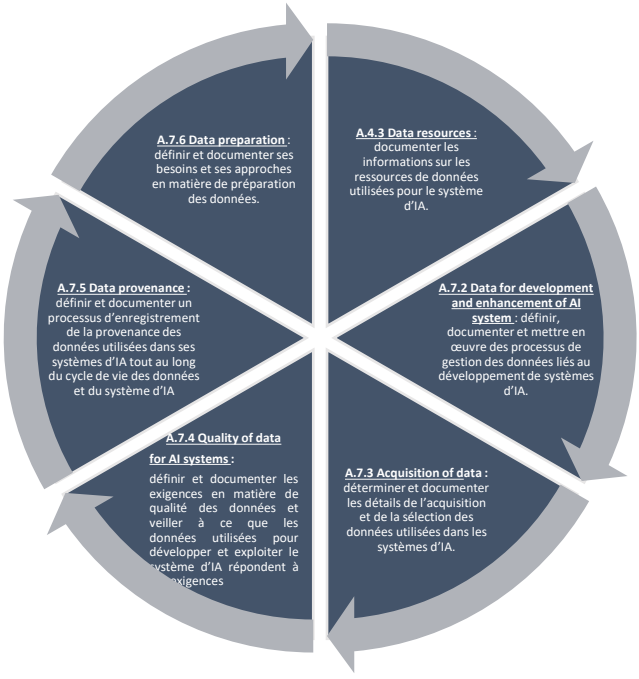
L'article 73 : Signalement d'incidents graves

- définit les procédures de signalement des incidents graves liés aux systèmes d'IA
- décrit les obligations de signalement des incidents.
- Exige des plans de communication en cas d'incident.

Exigence : Données et gouvernance des données

L'article 10 se concentre sur la gouvernance des données, y compris les pratiques de gestion des données, qui englobent la préparation des données.

- Couvre les processus de gestion des données.
- comprend les processus de collecte de données.
- met l'accent sur les exigences de qualité des données.
- exige la documentation de la provenance des données.



Exigence : Système de gestion des risques

Règlement IA	Norme ISO 42001
L'article 9 exige un système de gestion des risques pour évaluer les incidences potentielles. « *2 (…) Il comprend les étapes suivantes:* • *a) l'identification et l'analyse des risques connus et raisonnablement prévisibles que le système d'IA à haut risque peut poser pour la santé, la sécurité ou les droits fondamentaux lorsque le système d'IA à haut risque est utilisé conformément à sa destination;* • *b) l'estimation et l'évaluation des risques susceptibles d'apparaître lorsque le système d'IA à haut risque est utilisé conformément à sa destination et dans des conditions de mauvaise utilisation raisonnablement prévisible;* • *c) l'évaluation d'autres risques susceptibles d'apparaître, sur la base de l'analyse des données recueillies au moyen du système de surveillance après commercialisation visé à l'article 72;* • *d) l'adoption de mesures appropriées et ciblées de gestion des risques, conçues pour répondre aux risques identifiés en vertu du point a).* »	A.5.2 AI system impact assessment process L'organisation doit établir un processus d'évaluation des conséquences potentielles pour les individus et les sociétés qui peuvent découler du système d'IA tout au long de son cycle de vie.

Exigence : Documentation technique

La documentation technique est établie de manière à démontrer que le système d'IA à haut risque satisfait aux exigences énoncées dans la présente section et à fournir aux autorités nationales compétentes et aux organismes notifiés les informations nécessaires sous une forme claire et intelligible pour évaluer la conformité du système d'IA avec ces exigences.

Norme ISO 42001

Exigences : Surveillance-Etude d'impact -compliance et journalisation

Règlement IA	Norme ISO /IEEC 42001
L'article 72 Surveillance après commercialisation par les fournisseurs et plan de surveillance après commercialisation pour les systèmes d'IA à haut risque exigences en matière de surveillance après la mise en marché	**A.6.2.6 AI system operation and monitoring** -définit et documente les éléments nécessaires au fonctionnement continu du système d'IA. Au minimum, cela devrait inclure la surveillance du système et des performances, les réparations, les mises à jour et l'assistance
Article 27 : étude d'impact sur les droits fondamentaux (*Fundamental Rights Impact Assessment* FRIA) Nouvel instrument d'évaluation des risques : procédure technico-juridique qui permet d'évaluer les incidences négatives sur les droits fondamentaux résultant de l'utilisation de l'IA et d'adopter des mesures pour réduire les effets négatifs afin qu'ils deviennent acceptables. Elle permet de montrer que le déployeur a analysé toutes les diligences requises pour réduire les risques de violations et donc sa responsabilité ne saurait être engagée L'article 27 §1 limite le champ d'application de l'évaluation de risques aux déployeurs qui sont des organismes publics ou des déployeurs de droit privé mais fournissant des services publics.	**A.5.4 Assessing AI system impact on individuals and groups of individuals** évaluer et documenter les impacts potentiels des systèmes d'IA sur des individus ou des groupes d'individus tout au long du cycle de vie du système **A.5.5 Assessing societal impacts of AI systems** : évaluer et documenter les impacts sociétaux potentiels de ses systèmes d'IA tout au long de leur cycle de vie
Article 8 : Respect des exigences les systèmes d'IA à haut risque doivent être conformes aux exigences de l'article 8.2, compte tenu de leur destination et de l'état de l'art en matière de technologies d'IA. Cela comprend la spécification et la documentation des exigences relatives aux nouveaux systèmes d'IA ou aux améliorations.	**A.6.2.2 AI system requirements and specification** L'organisme doit spécifier et documenter les exigences relatives aux nouveaux systèmes d'IA ou aux améliorations matérielles apportées aux systèmes existants.
Article 19 : Journaux générés automatiquement Les fournisseurs de SIAHR assurent la tenue des journaux générés automatiquement par leurs systèmes, dans la mesure où ces journaux se trouvent sous leur contrôle.	**A.6.2.8 Système d'IA d'enregistrement des journaux d'événements** déterminer à quelles phases du cycle de vie du SIA, la tenue de registres d'événements doit être activée, mais au minimum lorsque le SIA est utilisé

Il en résulte que la certification à la norme ISO/IEEC 42001 permet d'être compliance avec les exigences suivantes de l'IA ACT :

Article 17 : Système de gestion de la qualité	
article 73 : Partage d'information sur les incidents graves	
article 10 : Données et gouvernance des données	
article 9 : système de gestion des risques	
article 11 : documentation technique	
article 72 : Surveillance après commercialisation par les fournisseurs et plan de surveillance après commercialisation pour les systèmes d'IA à haut risque	
article 27 Analyse d'impact des systèmes d'IA à haut risque sur les droits fondamentaux	
article 73 Signalement d'incidents graves	
article 8 Respect des exigences	
article 19 Journaux générés automatiquement	

Conclusion

Compte tenu que la norme ISO/IEC 42001:2023 fournit un cadre pour la mise en œuvre de système de gestion de l'IA, elle marque une étape dans le développement d'une gouvernance de l'IA structurée, éthique et responsable. Les organisations de soins de santé s'appuient sur la norme ISO 42001 pour renforcer la conformité, atténuer les risques liés à l'IA et renforcer la confiance du public. En intégrant des évaluations des risques, des garanties éthiques et une surveillance continue dans l'écosystème d'IA, cette norme permet une gouvernance de l'IA, responsable, efficace et durable.

A suivre ….

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

Eléonore Scaramozzino
Avocat
Constellation Avocats
escaramozzino@constellation.law

constellation • law