

LES BRIEFS DE L'IA RESPONSABLE

5

Conformité des systèmes et modèles d'IA (générative)

JANVIER 2025

Introduction

Ce *Brief de l'IA responsable* rend compte de ses travaux et propose des pratiques et des initiatives pour faire en sorte d'asseoir une conformité des systèmes et modèles d'IA (générationnelle) développés, déployés et/ou mis à disposition par les organisations vis-à-vis de toutes les législations, les règlements et autres cadres s'appliquant.

Il s'appuie sur les retours d'expérience des membres d'Impact AI et diverses études et initiatives récentes afin d'offrir une vue d'ensemble sur la conformité de l'IA (générationnelle) avec une multitude de lois, proposées ou promulguées, qui ont vu ou sont sur le point de voir le jour dans le monde entier, ainsi que la gouvernance (globale) à mettre en place dans un tel contexte.



1/ Le contexte

Le Règlement européen sur l'IA (EU AI Act) est entré en vigueur le 1^{er} août 2024 à la suite de [sa publication au Journal officiel de l'Union Européenne \(UE\)](#). Cette loi phare suit une approche « fondée sur le risque » comme cela a été couvert dans un précédent Brief de l'IA responsable et exprime à ce titre un ensemble d'exigences essentielles (de haut niveau) sur les organisations, les systèmes d'IA et les modèles d'IA à usage général (GPAI), en fonction du niveau de risque. Les exigences établies sont d'autant plus strictes que le risque vis-à-vis du cas d'utilisation (système) ou de la technologie (modèle) de nuire à la santé, à la sécurité, et/ou aux droits fondamentaux est élevé. L'AI Act est la première réglementation du genre dans le monde, elle est susceptible d'établir demain une référence mondiale pour la réglementation de l'IA comme peut l'être, à certains égards, le Règlement général pour la protection des données à caractère personnel (RGPD). Des « ponts » existent par ailleurs entre ces deux règlements.

Les nouveaux fournisseurs de modèles d'IA à usage général (à risque systémique ou non) pourront s'inscrire à partir du 2 août 2025 dans le cadre d'un code de pratique qui établit des obligations en matière de transparence des modèles et de leurs données d'entraînement, de gestion des droits d'auteur et, le cas échéant, de gestion des risques systémiques. Les modèles existants devront s'y conformer pour août 2027. [Ce code en cours d'élaboration](#) par le [Bureau européen de l'IA](#) est attendu pour avril 2025.

Par ailleurs, les systèmes à haut risque doivent respecter des contraintes particulières de gestion de risques, de qualité et de conformité auxquels ces derniers devront répondre à partir du 2 août 2026. À cette fin, le [comité technique mixte 21 du CEN-CENELEC](#) qui élabore les normes d'IA en Europe a reçu un mandat de normalisation, dès le mois de mai 2023, pour le développement d'un ensemble de « [Normes Harmonisées](#) » attendues pour mai 2025. Ces normes visent à définir les exigences techniques concrètes à satisfaire pour ces systèmes et, à la condition qu'elles soient publiées au Journal officiel de l'UE, conféreront une présomption légale de conformité aux systèmes d'IA développés

conformément à ces textes. Dans cette perspective, le comité technique mixte 21 adapte les normes internationales de l'ISO/IEC comme la norme [42001 Système de management pour l'IA \(AIMS\)](#) et en élabore de nouvelles pour soutenir la mise en œuvre et l'opérationnalisation du Règlement européen sur l'IA (cf. [programme de travail](#)). Les normes devront répondre à des caractéristiques précises et, en 2025, quatre normes structurantes sont notamment en cours d'élaboration vis-à-vis d'un cadre de gestion de l'IA digne de confiance, des risques, de la qualité et de l'évaluation de la conformité.

La norme ISO/IEC 42001 AIMS n'est adaptée aux conditions européennes que dans une [mesure relative](#). En effet, au-delà d'un non-alignement dans la définition du risque, la loi sur l'IA décrit les exigences relatives à un produit, alors que la norme 42001 décrit les exigences relatives à une organisation. En outre, les normes internationales de l'ISO/IEC sont davantage des bonnes pratiques et des lignes directrices que les organisations peuvent utiliser comme guide ; des cadres de certification peuvent néanmoins en être dérivées, comme en témoigne le projet au stade approbation de norme [ISO/IEC FDIS 42006 Exigences applicables aux organismes chargés de l'audit et de la certification des systèmes de gestion de l'intelligence artificielle](#). L'approche européenne, quant à elle, exige des critères vérifiables et applicables spécifiques en vue du marquage CE de ces « produits » d'IA.

Les membres d'Impact AI observent dans le monde une vague imminente de mesures réglementaires ou d'initiatives de même nature. En réponse aux diverses préoccupations de la société, des impacts de l'IA sur la sécurité nationale, les gouvernements réglementent l'IA. L'OCDE propose à ce titre [un référentiel en temps réel](#) de plus de 1 000 initiatives politiques en matière d'IA et de gouvernance provenant de 69 pays.

Ce paysage en évolution rapide impose aux organisations de toute taille un examen minutieux et à la clé des investissements conséquents, ne serait-ce que pour identifier les réglementations, les législations, les codes de conduite qui leur sont applicables en tant qu'organisation et/ou



vis-à-vis de leurs systèmes (et modèles) d'IA, pour en comprendre ensuite les exigences ainsi exprimées et les intégrer dans leurs cadres de gouvernance et systèmes de gestion internes (des risques, de la qualité, etc.).

Ce même paysage suggère également la nécessité d'aider à tracer une voie éclairée en vue d'une gouvernance internationale de l'IA. Le besoin d'une gouvernance efficace au niveau mondial ne cesse de croître afin de garantir l'interopérabilité de ces efforts nationaux en faveur

d'une IA sûre, sécurisée et digne de confiance, pour que les avantages de l'IA soient largement partagés et que les risques importants à l'échelle mondiale soient gérés de manière efficace.

Aujourd'hui, de nombreux gouvernements, institutions internationales et membres des secteurs privés et à but non lucratif sont engagés dans des initiatives qui vont dans le sens de ces objectifs. Il n'en reste pas moins que nous n'en sommes qu'au début de la gouvernance mondiale de l'IA en vue d'assurer une conformité globale.

2/ Le partage d'expérience des membres Impact AI

Naaia : la première solution certifiée ISO 42001

Fondée en 2021, **Naaia** apporte une réponse opérationnelle aux besoins de mise en conformité accélérée des systèmes d'IA. Son objectif est d'aider les entreprises à concilier confiance, performance et compliance pour accélérer l'adoption de l'IA. Naaia a mis en place une solution SaaS multi réglementations (AI Act, Chine, Corée du Sud, US...) et multi référentiels (ISO 42001, NIST, LNE...) qui accompagne les organisations pour atteindre leur cible de conformité des Systèmes d'IA. Sa solution est la première en France certifiée ISO 42001 et parmi les premières dans le monde. « Suite à une qualification du produit d'IA dans la plateforme, nous générons un plan d'action unique pour chaque système d'IA, sur la base d'un certain nombre de principes : transformer les obligations réglementaires en actions opérationnelles, fournir des templates et matcher entre eux les différents référentiels afin d'assurer le plus haut niveau de conformité et rationaliser les efforts », explique Isabelle Perl, Chief Marketing Officer de Naaia.

Mais il reste à sensibiliser les entreprises aux exigences de la régulation, les accompagner dans ce chemin de conformité avec des conseils, des outils, des méthodologies. Il est indispensable, en premier lieu, d'identifier les IA interdites, de cartographier les produits d'IA puis d'identifier le véritable niveau de risque, afin d'apporter une réponse à la fois pertinente mais adaptée. Il faut donc engager dès maintenant des démarches de gouvernance opérationnelle et d'outillage. Pour Isabelle Perl, les entreprises mettront en place « une gouvernance de transition » pour les deux prochaines années afin d'apprendre en marchant.

Microsoft France : prendre exemple sur les autres grandes activités régulées comme l'aérien ou le nucléaire

Les conversations sur les politiques en matière d'IA se sont multipliées ces derniers mois. Elles ont commencé à être ponctuées de



références répétées à des acronymes désignant les différentes institutions internationales qui régissent aujourd'hui de grands secteurs de l'économie, comme l'aviation civile, l'énergie nucléaire ou encore les flux de capitaux à l'échelle mondiale.

Microsoft qui a souhaité vouloir aller plus loin et en savoir plus sur la manière dont les approches de la gouvernance de secteurs comme l'aviation civile pouvaient s'appliquer à un ensemble de technologies, qui ne seraient jamais assemblées dans un hangar ou guidées par des officiers de contrôle du trafic aérien. Il s'agissait également d'en savoir plus sur les engagements nucléaires qui ont vu le jour à une tout autre époque pour réglementer une technologie prometteuse en tant qu'outil, mais

qui n'avait été utilisée que comme arme.

En effet, l'histoire nous enseigne depuis longtemps que la façon dont la technologie transforme notre monde dépend en partie de l'efficacité avec laquelle elle est gouvernée, et que la gouvernance internationale est vitale pour les technologies qui ne connaissent pas de frontières.

À la lumière de ces analyses, Microsoft se réjouit de pouvoir présenter et partager [Gouvernance globale : Objectifs et leçons pour l'IA](#), c'est-à-dire une collection de perspectives externes sur les institutions internationales de différents domaines, rassemblée avec ses propres réflexions sur les objectifs et les cadres de la gouvernance mondiale de l'IA.

3/ Les bonnes pratiques identifiées par les membres Impact IA

À travers les études de cas et les analyses du recueil [Gouvernance globale : Objectifs et leçons pour l'IA](#), l'histoire et l'évolution d'institutions telles que l'Organisation de l'aviation civile internationale et le Conseil de stabilité financière ainsi retracées, ainsi que le partage de réflexions sur leurs succès et leurs défis afin d'éclairer la gouvernance mondiale de l'IA, trois enseignements de haut niveau pour l'IA peuvent être tirés :

→ **1. Comme pour l'aviation civile et les flux mondiaux de capitaux, la gouvernance de l'IA comporte trois niveaux interdépendants :**

- i) les normes industrielles ;
- ii) la réglementation nationale ;
- iii) la gouvernance internationale.

→ **2. Au niveau de la gouvernance internationale, trois résultats sont importants pour l'IA :**

- i) une gouvernance des risques significative au niveau mondial ;
- ii) une interopérabilité réglementaire ;
- iii) un progrès inclusif.

→ **3. Quatre fonctions de gouvernance internationale permettront d'atteindre ces résultats :**

- i) la surveillance et la gestion des risques mondiaux ;
- ii) la définition de normes ;
- iii) l'établissement d'un consensus scientifique ;
- iv) le renforcement d'un accès approprié aux ressources nécessaires.



4/ Les défis et questions en suspens

Nos membres soulignent l'importance de maintenir à jour une veille réglementaire. Cette veille doit couvrir les différentes géographies [pertinentes] en termes de régulation de l'IA, intégrer les interactions entre les régulations de l'IA et d'autres régulations connexes [protection des données personnelles, cybersécurité, responsabilité sociale des entreprises, etc.] et anticiper même en Europe les régulations encore à venir autour de l'IA [par exemple les directives des régulateurs, les actes

délégués de la réglementation européenne sur l'IA, la directive responsabilité, ainsi que les futures jurisprudences].

Une telle veille peut inclure la participation à des groupes de travail interentreprise comme notre collectif Impact AI afin de :

- **Se préparer dans l'attente des standards** harmonisés et évaluer les alternatives à la standardisation ;
- **S'outiller pour préparer la conformité.**

5 / Pour aller au-delà

• [Règlement - UE - 2024/1689 du Parlement européen et du Conseil du 13 juin 2024 établissant des règles harmonisées concernant l'intelligence artificielle](#)

• [IAPP - EU AI Act](#)

• [Second Draft of the General-Purpose AI Code of Practice published, written by independent experts](#)

• [Analysis of the preliminary AI standardisation work plan in support of the AI Act](#)

• [Harmonised Standards for the European AI Act](#)

• [Regulating Under Uncertainty: Governance Options for Generative AI](#)

• [Global Governance: Goals and Lessons for AI](#)

• [Hub France IA : L'évaluation de conformité selon l'AI Act](#)

• [Guide de mise en œuvre de l'AI Act par Numeum](#)

Remerciements

à Philippe Beraud (Microsoft), Emilie Sirvent Hien (Orange)
et Matthieu Capron (Crédit Agricole) pour leur engagement
et leur contribution essentielle à ces Briefs de l'IA responsable.
à tous les participants aux ateliers qui ont bien voulu
partager leurs expériences et cas d'usage.

Retrouvez-nous sur
www.impact-ai.fr

pour rejoindre nos initiatives et retrouver nos divers travaux !

