AI Act - Opérationnalisation

GOUVERNANCE DES DONNÉES (1)

Usage technique des données dans le respect de la réglementation, afin d'identifier, de mesurer et de gérer les risques

Étape 1 ***Collecte*** **des données** → **S'assurer que la collecte est licite, des éventuels consentements et du droit de réutilisation**

Étape 2 ***Connaissance*** **des données** → **Décrire les données, les insérer dans un catalogue, documenter, gérer la qualité, etc.**

Étape 3 ***Prétraitement et enrichissement*** **des données** → **Effectuer les opérations de recodage et d'enrichissement nécessaires au traitement attendu**

(1) La phase et les étapes font référence au livrable « Opérationnaliser la gestion des risques » du GT Banque du Hub France IA

LEXIQUE

SIA = Système d'IA

HUB FRANCE .IA

AI Act – Opérationnalisation

GOUVERNANCE DES DONNÉES

- Cette exigence est prévue pour les **systèmes d'IA à haut risque** (SIA à HR) (art. 10) ;
- Par principe, elle concerne les **données d'entraînement, de validation et de test**. Par exception à ce principe, sont exclus les SIA qui ne font pas appel à des techniques qui impliquent l'entraînement de modèles d'IA. Dans ce cas, cette exigence s'applique uniquement aux jeux de données de test (art. 10, §6) ;
- Des **exigences spécifiques** sont aussi prévues **pour les fournisseurs de modèles d'IA à usage général** (GPAIm) : résumé synthétique des données, protection des données à caractère personnel et respect du droit d'auteur et des droits voisins (art. 53, c et d) ;
- Le **système de gestion de la qualité** (art. 17) et le **système de gestion des risques** (art. 9) doivent contribuer au respect de ces exigences. La gouvernance des données doit permettre d'identifier, gérer les risques pour la santé, la sécurité et les droits fondamentaux.

Préalables nécessaires	Boîte à outils	Notre *check-list* pour vous aider à opérationnaliser l'*AI Act*
		L'analyse du HFIA préconise d'adopter la même approche pour la mise en place d'une politique de gestion de la qualité pour les fournisseurs de GPAIm.
Définition d'un projet et des objectifs métier	• **Questionnaire de pré analyse** des risques du SIA[2]	❑ **Analyse du niveau de risque du SIA**, et qualification du rôle de votre organisation - **Fiche projet** [2] ;
Mise en place d'un système de gestion des risques des SIA à HR (art. 9)	• **Certification LNE « processus pour l'IA »** • « Cadre de gestion des risques en Intelligence Artificielle » du NIST (en attendant les normes harmonisées)	❑ **Politiques de gestion des risques à élaborer, et à mettre à jour** (art. 9) ;
Mise en place d'un système de gestion de la qualité des SIA à HR pour les fournisseurs (art. 17)	• **Norme ISO/IEC 42001:2023** « Système de management de l'intelligence artificielle » (en attendant les normes harmonisées)	❑ **Politiques, procédures et instructions écrites** concernant les systèmes et procédures de gestion des données (art. 17, §1, f) lors de l'acquisition, la collecte, l'analyse, l'étiquetage, le stockage, la filtration, l'exploration, l'agrégation, la conservation des données, et toute autre opération concernant les données, qui est effectuée avant la mise sur le marché, ou la mise en service, de SIA à HR.

[2] Cf. Le livrable « définition d'un projet » et « objectifs métier ». L'explication plus détaillée de la Fiche Projet est disponible dans les « concepts clés » livrés par le GT BAIA.

HUB FRANCE .IA

AI Act - Opérationnalisation

GOUVERNANCE DES DONNÉES

1. *Collecte* des données

Gérer les risques
- Assurer que la collecte respecte la réglementation et les référentiels internes
- **Si données à caractère personnel (DCP), respecter le principe de minimisation**

2. *Connaissance* des données

Évaluer la disponibilité, la quantité et l'adéquation de jeux de données (qualité) (art. 10, §2)
- **Vérifier la pertinence, la représentativité, l'absence d'erreur et de données manquantes, la présence de biais**
- Indiquer les caractéristiques des données collectées
- Les données doivent prendre en compte le cadre géographique, contextuel, comportemental ou fonctionnel (art. 10 §4, et art. 42)

3. *Prétraitement* et *enrichissement* des données

- **Nettoyer, annoter, extraire des caractéristiques, répartir des données**
- **Détecter, prévenir et atténuer les éventuels biais** (art. 10, §2, g)
- Si pertinent, enrichir les données

Boîte à outils

- ***Check-list* des principes de collecte des données** (art. 10, §2, b)
- ***Template* de formulaire d'accès et d'accréditation**
- **Outils de traçabilité des données (*data lineage*)**
- **Registre de traitement** si DCP
- **Modèle de résumé** prévu pour GPAIm (art. 53, §1, d)
- **Outils d'anonymisation/pseudonymisation et/ou données synthétiques**

- **Dictionnaire et/ou catalogue de données**
- **Méthodologie de contrôle de données (qualité)**

- **Guide de bonnes pratiques** (ou même une procédure dans certains cas réglementaires)
- **Outils d'annotation, étiquetage**
- **Outils de détection et suivi des biais**
- **Outils d'anonymisation/pseudonymisation**
- **Données agrégées** et/ou **données synthétiques**
- ***Feature store***
- Cadre d'utilisation du ***feature store***
- ***Tags* des données synthétiques** pour les **GPAIm**

Notre *check-list* pour vous aider à opérationnaliser l'*AI Act* (3)

❑ **Fiches** avec indication de la manière dont les données ont été **obtenues et sélectionnées** (art. 11, annexe IV, §2) ;

❑ Des informations détaillées sur la surveillance et le fonctionnement du SIA, en particulier en ce qui concerne […] les spécifications concernant les **données d'entrée** (art. 11, annexe IV, §3) ;

❑ **Résumé pour le GPAIm** (art. 53).

❑ **Fiches** sur les jeux de données (d'entraînement, validation, test) utilisés :
- Description générale de ces jeux de données ;
- Information sur leur provenance, leur portée, et leurs principales caractéristiques. (art. 11, annexe IV, §2) ;

❑ **Notice d'utilisation** (art. 13, §3, vi) spécifications relatives aux données d'entrée, information pertinente concernant les jeux de données d'entraînement, de validation et de test utilisés.

❑ **Fiches** sur les jeux de données (d'entraînement, validation, test) utilisés, présentant le processus de prétraitement (art. 11, annexe IV, §2) ;

❑ **Description détaillée** sur :
- Les procédures d'étiquetage ;
- Les méthodes de nettoyage ;
- Les mentions des approches d'enrichissement de la donnée (art. 10, §2, c, et annexe IV).

(3) Ces livrables font partie de la documentation technique, attendue dans le cadre de l'*AI Act*, et doivent être mis à jour.

HUB FRANCE IA

AI Act – Définitions clés

LES DONNÉES dans le RIA

Définitions RIA– Article 3

29) « ***données d'entraînement*** », les données utilisées pour entraîner un système d'IA en ajustant ses paramètres entraînables ;

30) « ***données de validation*** », les données utilisées pour fournir une évaluation du système d'IA entraîné et pour régler ses paramètres non entraînables ainsi que son processus d'apprentissage, afin, notamment, d'éviter tout sous-ajustement ou surajustement ;

31) « ***jeu de données de validation*** », un jeu de données distinct ou une partie du jeu de données d'entraînement, sous la forme d'une division variable ou fixe ;

32) « ***données de test*** »,les données utilisées pour fournir une évaluation indépendante du système d'IA afin de confirmer la performance attendue de ce système avant sa mise sur le marché ou sa mise en service ;

33) « ***données d'entrée*** » (*inputs*), les données fournies à un système d'IA ou directement acquises par celui-ci et à partir desquelles il produit une sortie ;

34) « ***données biométriques*** », v. « définitions clés ».

Définitions techniques

Tout d'abord, le « jeu de données » n'est pas défini dans l'*AI Act*. Il s'agit d'un ensemble de données utilisées, soit pour produire un SIA, soit pour l'utiliser.

Les différentes définitions fournies par l'article 3 concernant les données reprennent le langage technique typique de l'apprentissage automatique/machine (*Machine Learning* ou ML). Ces définitions peuvent avoir une acception plus large dans la pratique que dans le texte, et doivent donc être interprétées avec précaution.

Pour simplifier, on peut imaginer la calibration classique d'un modèle d'apprentissage automatique/machine (*Machine Learning* ou ML) en 3 étapes :

- **Apprentissage** : le système s'entraîne sur ces données pour effectuer la tâche attendue ; à cette étape, il s'agit des données ***d'entraînement*** (29).
- **Validation** : estimation de l'erreur de prédiction pour la sélection du modèle et l'optimisation des hyperparamètres ; à cette étape, il s'agit des données ***de validation*** (30).
- **Test** : évaluation de l'erreur de généralisation du modèle final choisi ; à cette étape, il s'agit des données ***de test*** (32).

Il est important de respecter les spécificités de chaque jeu de données à chaque étape, à défaut d'entraîner une surestimation des performances du modèle.

Les **données d'entrée** sont des données utilisées pour la prise de décision du système d'IA en production. Elles ne sont pas spécifiquement mentionnées par l'*AI Act* à l'article 10. Toutefois, le Hub France IA considère que les mêmes exigences en termes de qualité devraient être également garanties pour ces données.

Les **données synthétiques** sont des données générées artificiellement dont l'objectif est d'imiter et simuler les caractéristiques de l'ensemble de données d'origine. Il existe différentes méthodes pour générer des données synthétiques (extraction, reproduction des propriétés des données d'origine, utilisation de connaissances d'experts, ou de processus connus). Les données synthétiques présentent l'avantage considérable de bénéficier d'un régime juridique allégé, ce qui en facilite l'utilisation par rapport à des données réelles.