

Dossier de conformité — données d'entraînement (AI Act, article 10)

Squelette généré automatiquement à partir de l'analyse du jeu de données. Les blocs « à compléter par l'expert » nécessitent une revue humaine (provenance, jugement de conformité, mesures d'atténuation).

1. Identité & finalité

- **Système / modèle** : _____
- **Version du dossier** : _____
- **Date d'analyse** : _____
- **Jeu de données** : src
- **Finalité prévue** : _____

À compléter par l'expert — finalité prévue et ce que les données sont censées représenter

2. Provenance & base légale

Origine de chaque source, licences/contrats, et — pour les données personnelles — finalité initiale et base légale RGPD.

À compléter par l'expert — provenance par source (flux séparé §16), licences, base légale RGPD

3. Composition (rempli automatiquement)

- **Volume** : 303 lignes, 14 colonnes.
- **Variables** :

Colonne	Type
age	DOUBLE
sex	DOUBLE
cp	DOUBLE
trestbps	DOUBLE
chol	DOUBLE
fbs	DOUBLE
restecg	DOUBLE
thalach	DOUBLE
exang	DOUBLE
oldpeak	DOUBLE
slope	DOUBLE
ca	DOUBLE
thal	DOUBLE
target	BIGINT

- **Attributs sensibles déclarés** : age, sex.

À compléter par l'expert — périmètre géographique / contextuel / comportemental

4. Préparation (observations automatiques)

- **Valeurs manquantes** : ca (4) -> stratégie de traitement à documenter.

À compléter par l'expert — journal des transformations : collecte, nettoyage, étiquetage, enrichissement, agrégation

5. Qualité (rempli automatiquement)

- **Complétude** : 4 valeur(s) manquante(s) (ca).
- **Exactitude / aberrations** : chol (1).

À compléter par l'expert — représentativité et pertinence vis-à-vis de la finalité

6. Biais

Analyse automatique

Caractéristique	Mesure	Cible	Verdict
Complétude	99.9 %	"e 95.0 %	conforme
Unicité	100.0 %	"e 99.0 %	conforme
Validité	100.0 %	"e 98.0 %	conforme

sex — « target » : **disparate impact** (ratio) 2.41 · **parité statistique** (écart) 13.1 % (!)

Groupe	Effectif	Taux de « target »
1.0	206	22.3 %
0.0	97	9.3 %

Jugement & mesures

À compléter par l'expert — impact sur santé/sécurité/droits fondamentaux et mesures de détection/prévention/atténuation

7. Lacunes & limites

Points détectés à examiner comme lacunes potentielles :

- [low] missing — ca
- [low] missing — tha1
- [low] outliers — chol
- [high] bias — sex

À compléter par l'expert — manques identifiés et traités ; usages hors périmètre

8. Gouvernance & traçabilité

- **Version de l'outil d'analyse** : 0 . 4 . 1
- **Empreinte du jeu (schéma + volume)** : a5aed4387668f53f

À compléter par l'expert — responsabilités, journal d'audit, versioning, maintenance/mise à jour