



République du Bénin

\*\*\*\*\*

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique  
(MESRS)

\*\*\*\*\*

Université d'Abomey-Calavi (UAC)

\*\*\*\*\*

Ecole Polytechnique d'Abomey-Calavi (EPAC)

\*\*\*\*\*

Département de Génie Informatique et Télécommunications (GIT)

\*\*\*\*\*

Option : Réseaux Informatiques et Internet (RII)

**Mémoire de fin de formation pour l'obtention du diplôme d'ingénieur  
de conception grade master**

**Thème :**

---

**Outil automatisé de pré-traitement éthique et transparent  
des données.**

---

Réalisé et soutenu par :

**BONOU Mael Habib Sedjro**

Maître de mémoire :

**Dr SOGBOHOSSOU Medesu**

Enseignant - Chercheur à l'Ecole Polytechnique d'Abomey-Calavi.

Encadreur :

**Dr (MA) HOUNDEJI Ratheil**

Enseignant - Chercheur à l'IFRI

**Année académique : 2023 - 2024**



**République du Bénin**

\*\*\*\*\*

**Ministère de l'Enseignement Supérieur et de la Recherche Scientifique  
(MESRS)**

\*\*\*\*\*

**Université d'Abomey-Calavi (UAC)**

\*\*\*\*\*

**Ecole Polytechnique d'Abomey-Calavi (EPAC)**

\*\*\*\*\*

**Département de Génie Informatique et Télécommunications (GIT)**

\*\*\*\*\*

**Directeur**

**Docteur ALITONOU Guy Alain  
(Professeur Titulaire des Universités du CAMES)**

**Directeur Adjoint  
(Chargé des affaires académiques)**

**PRODJINONTO Vincent  
(Professeur Titulaire des Universités du CAMES)**

**Chef de département**

**Docteur DJARA Tahirou  
(Maître de Conférences des Universités du CAMES)**

**Année académique : 2023 - 2024**

**17<sup>ème</sup> Promotion**

# LISTE DES ENSEIGNANTS AYANT INTERVENU DANS NOTRE FORMATION DE 2019 A 2024

<b>NOM</b>	<b>PRÉNOMS</b>	<b>MATIÈRES ENSEIGNÉES</b>
ABALLO	Théophile	Réseaux Locaux, Conception des Systèmes d'Information (UML), Protocoles : Couches de Connexion de Réseaux et de Communication
ADANHOUNME	Villévo	Analyse et calcul intégral, Analyse vectorielle et Équation différentielle
AGBAHUNGBA	Georges	Méthodologie de Recherche Documentaire
AGBOMAHENAN	Macaire	Électromagnétisme, Technologie des Mesures Électriques
AKOWANOU	Onésime	Électrochimie
ALLOGNON	Elisabeth	Cinématique et Dynamique
ANJORIN	Malahimi	Thermodynamique, Transferts Thermiques
ASSOGBA	Emery	Architecture des ordinateurs, Microprocesseurs et microcontrôleurs
ASSOGBA	Marc	Probabilité et processus aléatoires, Programmation JAVA, Traitement Numérique de la Parole
BOURAIMA	Marcos	Recherche Opérationnelle
BOCCO	Elvarez	Algorithme numérique
CHETANGNY	Patrice	Compatibilité Électromagnétique
CHITOU	Naimoulai	Dessin Technique

<b>NOM</b>	<b>PRÉNOMS</b>	<b>MATIÈRES ENSEIGNÉES</b>
CODJIA	Wilfried	Gestion des Projets Informatiques
CODO	Paul	Statique Graphique et Analytique
COMLAN	Maurice	Programmation Langage C++
DEGAN	Gérard	Analyse Numérique
DEGBO	Basile	Transmission analogique, Rayonnement et antenne, Systèmes et Technologies Radio mobile et sans fil
DEGUENON	Judicaël	Probabilité et Statistique, Équation aux dérivées partielles
DJARA	Tahirou	Systèmes de Gestion de Base de Données, Environnement et développement logiciel, Traitement Numérique de l'Image
DJOGBE	Léopold	Fonctions Électroniques, Transmission Numérique du signal, Circuits radio fréquences et Dispositifs hyperfréquences
DOGUE	Karel	Droit Industriel
DOSSOU	Michel	Introduction Générale aux Télécommunications, Architecture des Réseaux, Dispositifs et transmission optique, Télévision Numérique Terrestre
EGOUNLETY	Richard	Physique des semi-conducteurs et applications
GBAGUIDI	Gérard	Résistance des Matériaux
GUEDJE	François	Initiation à l'algorithmique, Langage et Programmation Informatique
HOUANOU	Agapi	Analyse tensorielle
HOUEDAKO	Vincent	Circuits Électriques

<b>NOM</b>	<b>PRÉNOMS</b>	<b>MATIÈRES ENSEIGNÉES</b>
HOUNDEJI	Ratheil	Intelligence Artificielle et Applications
HOUNGAN	Théophile	Signaux et Systèmes
JOSSOU	Thierry	Électricité générale
MONTEIRO	Léonard	Convertisseurs numériques et circuits logiques programmables
LALEYE	Claude	Base et Contrôle de Qualité
NOUNAGNON	Charles	Métrologie et analyse des données
N'TCHA	Sabi	Outils Statistiques en Milieu Industriel
PRODJINONTO	Vincent	Mécanique des fluides
SANYA	Max Fréjus	Circuits Logiques Combinatoires et Séquentiels, Théorie des Graphes et Réseaux, Réseaux de Télécommunications, Micro ondes et Radar
SEMASSOU	Clarence	Optique Géométrique
SEGUEDEME	Alexis	Anglais Technique V et VI
SEWANOUDE	Damien	Comptabilité Générale, Création et Gestion des Entreprises, Marketing
SOEDE	Casimir	Anglais Technique I, II, III et IV
SOGBOHOSSOU	Médésu	Système d'Exploitation, Langage C, Théorie des Langages et Interpréteurs, Systèmes Temps Réel et Systèmes Embarqués
SOHOUNHLOUE	Dominique	Chimie pour l'ingénieur
SOTINDJO	Patrick	Traitement numérique du signal, Gestion du spectre de fréquences
TOGNIVO	Marc	Éducation Physique et Sportive
ZOUTANGNI	Laurent	Techniques d'Expression et Méthodes de Communications I et II

# DÉDICACE

# REMERCIEMENTS

Mes sincères remerciements :

# LISTE DES SIGLES ET ABRÉVIATIONS

**A**

**B**

**C**

**D**



# LISTE DES TABLEAUX

# LISTE DES FIGURES

1.1	Problèmes liés aux données collectées dans le monde réel. [?]	7
1.2	Flux de travail de base de la préparation des données et de l'ingénierie des caractéristiques dans l'apprentissage automatique et le développement d'applications big data. [4]	9
1.3	Google AutoML [?]	11
1.4	H2O.ai [?]	12
1.5	DataRobot [?]	13
1.6	Trifacta Wrangler (DataWrangler) [?]	13
1.7	IBM Watson Studio (AutoAI) [?]	14
1.8	Comparaison des outils de prétraitement automatisé (outils en lignes, critères en colonnes)	14
2.1	Types de préoccupations et défis liés à l'éthique de l'IA [5]	18
4.1	Logo de Python, le langage de programmation choisi pour ce projet.	31
4.2	Logo de VSCode, l'éditeur de code utilisé pour ce projet.	31
4.3	Logo de Git, le système de contrôle de version utilisé pour ce projet.	32
4.4	Logo de PydanticAI	33
4.5	Logo de Ai Fairness 360.	34
5.1	Représentation des composants de la couche d'entrée des données	38

5.2	Représentation des composants de la couche de traitement principal . . . . .	40
5.3	Représentation des composants de la couche d'équité et d'analyse	42
5.4	Représentation des composants du système d'agent IA . . . . .	43
5.5	Représentation des composants de la partie sortie et rapports . .	44

# LISTE DES SYMBOLES ET UNITÉS DE MESURES

# RÉSUMÉ

# ABSTRACT

# SOMMAIRE

<b>INTRODUCTION GÉNÉRALE</b>	<b>1</b>
<b>I SYNTHÈSE BIBLIOGRAPHIQUE</b>	<b>5</b>
1 AUTOMATISATION DU PRE-TRAITEMENT DES DONNÉES	6
2 CONSIDERATIONS ETHIQUES DANS LE TRAITEMENT AUTOMATISE DES DONNEES	16
3 TRANSPARENCE DANS LES OUTILS AUTOMATISES	22
<b>II MATERIELS ET METHODES</b>	<b>24</b>
4 MATERIELS	25
5 METHODES	35
<b>RÉFÉRENCES BIBLIOGRAPHIQUES</b>	<b>46</b>

# INTRODUCTION GÉNÉRALE

Dans un monde de plus en plus axé sur les données, la capacité à traiter et à analyser de grandes quantités d'informations est devenue un atout stratégique majeur pour les entreprises, les gouvernements et les chercheurs. Cependant, l'intégrité et la qualité des données sont souvent compromises par des erreurs, des biais, et des informations incomplètes. Le pré-traitement des données, qui englobe des étapes telles que le nettoyage, la normalisation et la détection des biais, est essentiel pour garantir des analyses fiables et des décisions éclairées. Historiquement, ce processus a été effectué manuellement, ce qui présente des risques d'erreurs humaines et nécessite un temps considérable.

Au fil des années, plusieurs études ont tenté d'automatiser ces étapes. Par exemple, l'introduction des bibliothèques telles que Pandas et Scikit-learn a permis de faciliter le nettoyage et la préparation des données. De plus, des algorithmes comme l'Analyse en Composantes Principales (ACP) ont été largement utilisés pour réduire la dimensionnalité des données sans en perdre l'essence. Cependant, ces solutions, bien que fonctionnelles, manquent souvent de mécanismes intégrés pour détecter et mitiger les biais présents dans les ensembles de données, posant ainsi des défis éthiques majeurs[1].

Aujourd'hui, l'éthique dans le traitement des données est devenue une préoccupation centrale. Des recherches récentes, telles que celles sur les biais algorithmiques, ont mis en évidence le danger que représentent les biais inconscients dans les systèmes automatisés. Des solutions comme AIF360 d'IBM, les Indi-



cateurs d'équité de Google et Fairlearn ont été développées pour détecter ces biais[2], mais elles restent limitées en termes de couverture et de transparence. Peu d'outils automatisés intègrent à la fois le pré-traitement des données et une analyse exhaustive des biais, tout en générant des rapports détaillés et explicatifs pour assurer la transparence du processus.

C'est dans ce contexte que se pose la problématique de cette étude : comment développer un outil automatisé de pré-traitement des données qui soit à la fois éthique et transparent, tout en répondant aux besoins actuels des acteurs de la Data Science ? Si des progrès notables ont été réalisés dans l'automatisation du pré-traitement, il reste des lacunes importantes dans l'intégration de principes éthiques et de mécanismes de transparence, essentiels pour garantir des analyses justes et impartiales.

Le but de cette étude est donc de développer un outil automatisé qui intègre non seulement les étapes classiques de pré-traitement des données (nettoyage, réduction de dimensionnalité), mais aussi des fonctionnalités permettant de détecter et mitiger les biais potentiels, tout en générant des rapports explicatifs détaillés des actions entreprises. Cet outil vise à répondre aux besoins croissants de transparence et d'éthique dans le traitement des données, en fournissant aux utilisateurs des recommandations claires et justifiées, contribuant ainsi à une meilleure prise de décision basée sur des données fiables et impartiales.

## CONTEXTE, JUSTIFICATION ET PROBLÉMATIQUE

Dans l'ère du Big Data, les organisations exploitent de vastes quantités de données pour orienter leurs décisions stratégiques. Cependant, des erreurs, des biais et des données manquantes compromettent souvent la qualité des analyses, pouvant mener à de mauvaises décisions. Le pré-traitement des données est donc essentiel pour assurer la fiabilité et l'intégrité des analyses.

Le pré-traitement des données inclut des étapes comme le nettoyage, la ré-

duction de dimensionnalité et la détection des biais, souvent effectuées manuellement, ce qui les rend fastidieuses et sujettes aux erreurs. Le manque de transparence pose également des préoccupations éthiques liées à la discrimination et aux biais implicites. L'automatisation du pré-traitement, via des algorithmes avancés et l'apprentissage automatique, permet d'automatiser ces tâches tout en garantissant transparence et éthique. Ces outils peuvent détecter et corriger les biais, nettoyer les données efficacement et générer des rapports détaillés.

Le développement d'un outil automatisé de pré-traitement des données, comme proposé dans ce mémoire, répond à un besoin croissant dans divers domaines, tels que la recherche scientifique, les entreprises et les institutions gouvernementales[3]. En intégrant des bibliothèques de données éthiques et des analyses de corrélation, cet outil vise à améliorer la qualité des données et à garantir des analyses plus fiables et impartiales.

La justification de ce projet repose sur l'importance de disposer de données de haute qualité pour des analyses précises et des décisions éclairées. En automatisant le pré-traitement des données et en assurant une transparence totale, cet outil contribuera à renforcer la confiance dans les analyses de données et à promouvoir des pratiques éthiques dans la gestion des données.

## **HYPOTHÈSES DE RECHERCHE**

L'intégration d'un outil automatisé de pré-traitement des données, doté de mécanismes de détection et de mitigation des biais, améliore non seulement la transparence et l'éthique dans le traitement des données, mais permet également de générer des rapports détaillés et auditable sur chaque étape du pré-traitement. Ces rapports renforcent la traçabilité des actions entreprises, réduisent les erreurs humaines et optimisent le processus de préparation des données pour des analyses plus fiables et impartiales.

### OBJECTIFS DE L'ÉTUDE

Ce travail a pour objectif de développer un outil automatisé de pré-traitement des données qui garantit une approche éthique et transparente en intégrant des mécanismes de détection des biais, de génération de rapports détaillés, et d'audit complet des étapes de pré-traitement.

Plus spécifiquement, il s'agira :

- ✓ d'identifier et classer automatiquement les types de données (quantitatives, qualitatives, et démographiques) afin de faciliter le pré-traitement initial et la détection des biais potentiels ;

Resultat attendu : Une classification précise des données pour chaque ensemble de données analysé, permettant d'orienter le processus de nettoyage et d'analyse des biais.

- ✓ d'automatiser le nettoyage des données et la réduction de dimensionnalité, tout en intégrant des algorithmes de détection et de mitigation des biais, pour assurer la qualité des données ;

Resultat attendu : Des ensembles de données propres, normalisés, et débarrassés de biais explicites, prêts pour l'analyse.

- ✓ de générer des rapports détaillés et auditable expliquant chaque étape du pré-traitement, les techniques utilisées, et les actions de mitigation des biais entreprises, tout en offrant un audit du code pour vérifier sa conformité éthique ;

Resultat attendu : Des rapports compréhensibles et exploitables, documentant les actions entreprises pour garantir la transparence et la traçabilité de toutes les opérations de pré-traitement.

---

---

# *Première partie*

---

## **SYNTHÈSE BIBLIOGRAPHIQUE**

# AUTOMATISATION DU PRE-TRAITEMENT DES DONNÉES

Le pré-traitement des données, crucial pour l'analyse, est traditionnellement manuel et sujet à des erreurs. Avec l'essor du Big Data, l'automatisation a amélioré l'efficacité et la précision de ces tâches comme le nettoyage et la réduction de dimensionnalité.

Dans ce premier chapitre, nous allons explorer dans un premier temps le concept d'automatisation du pretraitement des données. Ensuite nous aborderons le sujet des considérations éthiques dans le traitement automatisé des données ainsi que de la transparence dans les outils automatisés.

Enfin, nous ferons un état de l'art des réalisations qui ont déjà été faites dans le cadre de notre études.

L'automatisation du pré-traitement des données est une étape critique dans les projets d'analyse de données, notamment dans les contextes de Big Data. Le pré-traitement est nécessaire pour garantir la qualité et l'exploitabilité des données, mais il est également connu pour être une phase chronophage, représentant entre 50 et 80 pourcent du temps consacré à un projet d'analyse. Sans une préparation adéquate des données, même les meilleurs modèles d'apprentissage automatique risquent de produire des résultats imprécis[1]. De plus, les données collectées à partir de n'importe quelle source restent incomplètes, bruyantes et incohérentes, ce qui entraîne des problèmes dans l'analyse des données. Il est donc nécessaire de rectifier les problèmes à l'avance, qui peuvent être classés en trois groupes :

trop de données, trop peu de données et des données fragmentées. La figure suivante explique les problèmes posés par les données sous forme de tableaux [?].

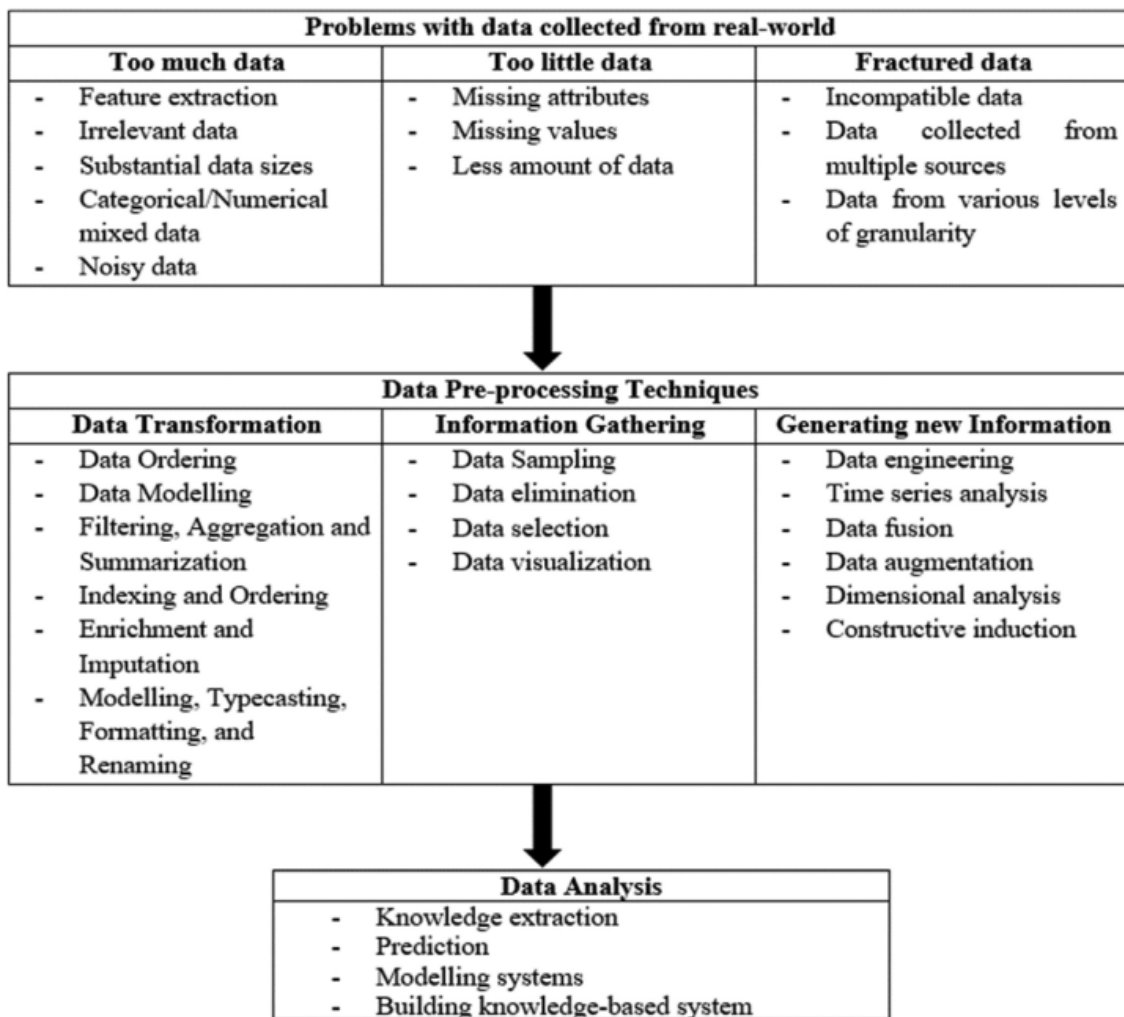


Figure 1.1 – Problèmes liés aux données collectées dans le monde réel. [?]

## 1.1 Clarification conceptuelle

Le pré-traitement des données est une étape cruciale dans les pipelines d'apprentissage automatique et d'intelligence artificielle (IA). Il englobe des opérations telles que le nettoyage des données, la gestion des valeurs manquantes, la détection et la correction des anomalies, ainsi que la normalisation et la réduction de dimensionnalité. Ces processus sont essentiels pour préparer les données à une analyse efficace et pour garantir la qualité des résultats obtenus. Historiquement, ces tâches étaient réalisées manuellement, ce qui était à la fois chronophage et sujet à des erreurs humaines. Cependant, avec l'avènement des outils modernes comme Pandas et Scikit-learn, une grande partie de ces étapes a été automatisée, permettant ainsi de gagner en efficacité et en précision.

L'approche moderne de l'IA vise à concevoir des algorithmes capables d'apprendre directement à partir des données. Cette approche a permis des avancées significatives, notamment dans le domaine de l'apprentissage supervisé et de l'apprentissage profond (deep learning). Cependant, malgré ces progrès, toutes les étapes de traitement des données dans les pipelines d'apprentissage profond n'ont pas été entièrement automatisées. Traditionnellement, les données doivent être collectées manuellement, prétraitées et enrichies par des techniques d'augmentation de données avant de pouvoir être utilisées pour l'entraînement des modèles. Récemment, des techniques spécialisées pour automatiser ces tâches ont émergé, répondant au besoin croissant de traiter de grands volumes de données complexes et hétérogènes pour les applications de big data et d'apprentissage automatique.

Aujourd'hui, des systèmes de traitement de données end-to-end basés sur des techniques d'apprentissage automatique automatisé (AutoML) sont capables de transformer des données brutes en caractéristiques utiles pour des tâches de big data, en automatisant toutes les étapes intermédiaires. Cela inclut le prétraitement automatisé (nettoyage des données, imputation des valeurs manquantes, encodage des données catégorielles), l'augmentation des données (y compris la génération de données synthétiques via des méthodes d'IA générative) et l'ingénierie des caractéristiques (extraction, construction et sélection automatisées

des caractéristiques). Ces systèmes permettent non seulement d'automatiser des tâches spécifiques, mais aussi d'optimiser simultanément toutes les étapes du pipeline d'apprentissage automatique.

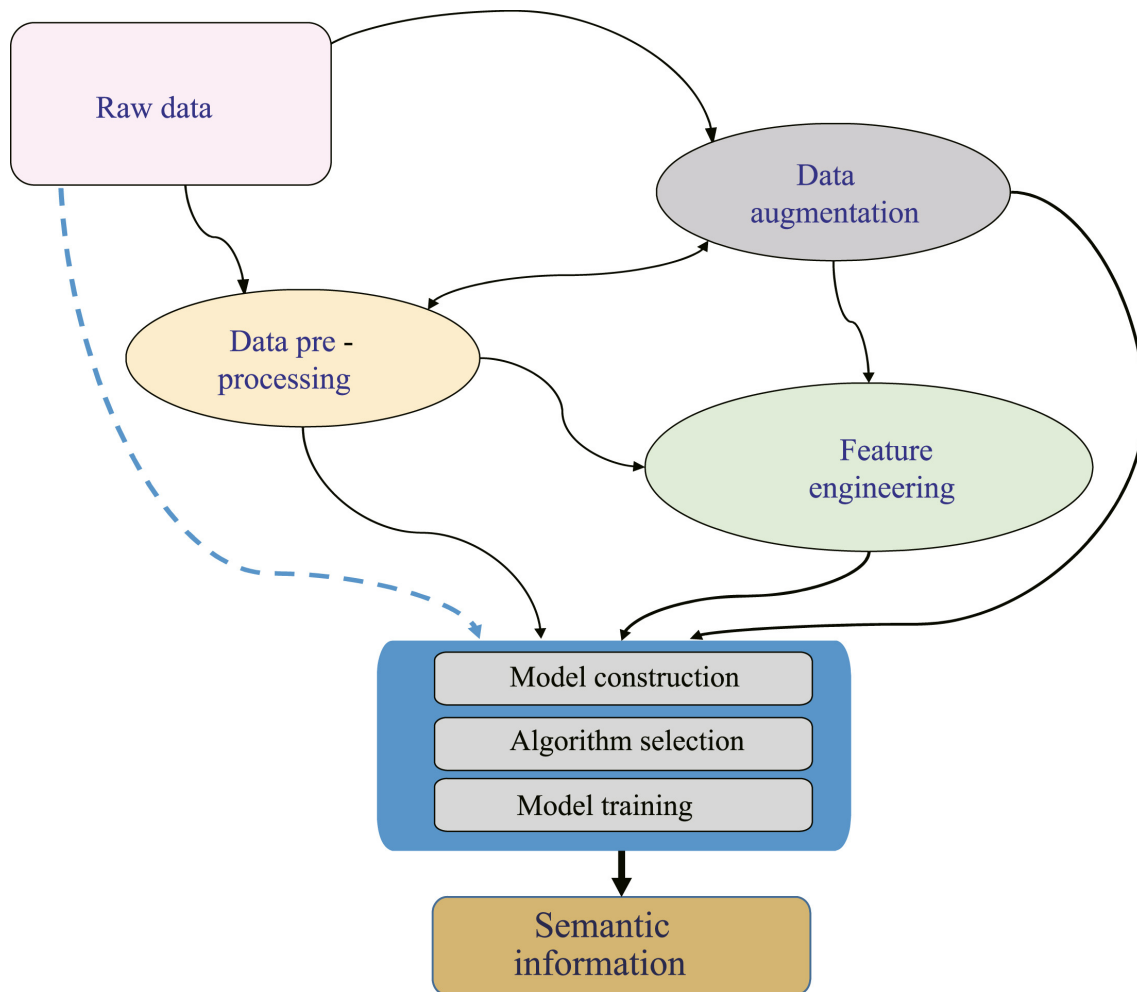


Figure 1.2 – Flux de travail de base de la préparation des données et de l'ingénierie des caractéristiques dans l'apprentissage automatique et le développement d'applications big data. [4]

En résumé, l'automatisation du pré-traitement des données est motivée par la nécessité de gérer des volumes croissants de données complexes tout en garantissant leur qualité et leur utilité pour les modèles d'IA. Les outils modernes et les techniques d'AutoML jouent un rôle clé dans cette évolution, rendant les pipelines de traitement des données plus efficaces, transparents et adaptés aux défis actuels du big data et de l'apprentissage profond.[1].





## 1.2 Solutions existantes

Aujourd'hui, plusieurs solutions logicielles, comme H2O-AutoML, AutoSklearn et DataRobot, intègrent des fonctions de pré-traitement automatisé. Ces outils sont capables de détecter les types de données (numériques, catégorielles, etc.) et de réaliser certaines transformations de base. Cependant, la majorité de ces outils nécessitent une intervention humaine pour les configurations plus avancées, telles que la sélection des méthodes de transformation adaptées aux besoins spécifiques des utilisateurs. L'une des principales lacunes de ces solutions réside dans leur incapacité à proposer des recommandations pertinentes pour choisir les meilleurs paramètres de transformation, ce qui peut affecter la qualité des résultats obtenus[4].

### 1.2.1 Description et comparaison de quelques outils

#### 1.2.2 Google AutoML

- **Description** : Google AutoML est une plateforme cloud qui automatise la création de modèles de machine learning, y compris le prétraitement des données. Il est conçu pour les utilisateurs ayant peu d'expérience en machine learning.
- **Points forts** :
  - Automatisation complète du prétraitement (nettoyage, encodage, normalisation).
  - Intégration avec les services Google Cloud.
  - Support pour les données structurées et non structurées (texte, images).



Google's AutoML

Figure 1.3 – Google AutoML [ ? ]

### 1.2.3 H2O.ai

- **Description** : H2O.ai est une plateforme open-source et commerciale qui automatise l'ensemble du pipeline de machine learning, y compris le pré-traitement des données. Elle est connue pour son framework **H2O** et son produit **Driverless AI**.
- **Points forts** :
  - Automatisation avancée du prétraitement (gestion des valeurs manquantes, sélection de caractéristiques).
  - Ingénierie des caractéristiques automatisée.
  - Détection des biais et explication des modèles.



Figure 1.4 – H2O.ai [?]

### 1.2.4 DataRobot

- **Description** : DataRobot est une plateforme d'apprentissage automatique automatisé (AutoML) qui permet de créer, déployer et gérer des modèles de machine learning. Elle automatise également le prétraitement des données.
- **Points forts** :
  - Nettoyage et transformation automatiques des données.
  - Sélection et ingénierie des caractéristiques.
  - Détection des biais et explication des modèles.

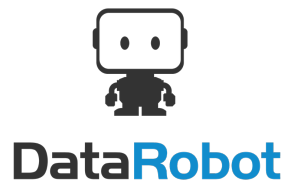


Figure 1.5 – DataRobot [ ?]

### 1.2.5 Trifacta Wrangler (DataWrangler)

- **Description** : Trifacta Wrangler est un outil de préparation de données qui utilise l'IA pour automatiser le nettoyage et la transformation des données. Il est conçu pour les utilisateurs techniques et non techniques.
- **Points forts** :
  - Interface visuelle intuitive pour le prétraitement des données.
  - Détection automatique des schémas de données.
  - Transformation automatisée des données.



Figure 1.6 – Trifacta Wrangler (DataWrangler) [ ?]

### 1.2.6 IBM Watson Studio (AutoAI)

- **Description** : IBM Watson Studio est une plateforme cloud qui inclut une fonctionnalité appelée AutoAI pour automatiser le prétraitement des données et la construction de modèles.

— **Points forts :**

- Nettoyage et transformation automatiques des données.
- Sélection et ingénierie des caractéristiques.
- Détection des biais et explication des modèles.



Figure 1.7 – IBM Watson Studio (AutoAI) [ ? ]

### 1.2.7 Tableau Comparatif

<i>Outils</i>	<i>Google AutoML</i>	<i>H2O.ai</i>	<i>DataRobot</i>	<i>Trifacta Wrangler</i>	<i>IBM Watson Studio (AutoAI)</i>
<i>Niveau d'intervention humaine</i>	Très faible	Très faible	Très faible	Faible à modéré	Très faible
<i>Gestion de l'aspect éthique</i>	Détection des biais, mais limitée	Détection et correction des biais	Détection des biais et transparence	Qualité des données, gestion manuelle des biais	Détection des biais et transparence
<i>Automatisation du prétraitement</i>	Complète	Complète	Complète	Automatisation partielle	Complète
<i>Support des données</i>	Structurées et non structurées	Structurées	Structurées	Structurées	Structurées et non structurées
<i>Intégration cloud</i>	Google Cloud	Multi-cloud	Multi-cloud	Multi-cloud	IBM Cloud
<i>Coût</i>	Payant	Open-source (H2O) et payant (Driverless AI)	Payant	Payant	Payant
<i>Facilité d'utilisation</i>	Très facile	Modérée à avancée	Facile	Très facile	Facile

Figure 1.8 – Comparaison des outils de prétraitement automatisé (outils en lignes, critères en colonnes)

### 1.3 Analyse des Axes de Comparaison

#### Niveau d'intervention humaine nécessaire

- **Google AutoML, H2O.ai, DataRobot, et IBM Watson Studio (AutoAI)** nécessitent un **niveau d'intervention humaine très faible**. Ces outils automatisent presque entièrement le prétraitement des données.
- **Trifacta Wrangler** nécessite un **niveau d'intervention faible à modéré** en raison de son interface visuelle, qui permet à l'utilisateur de valider les transformations.

#### Gestion de l'aspect éthique

- **H2O.ai, DataRobot, et IBM Watson Studio** offrent des fonctionnalités avancées pour la **détection des biais** et la **transparence des modèles**.
- **Google AutoML** propose des outils pour détecter les biais, mais la gestion éthique dépend davantage de l'utilisateur.
- **Trifacta Wrangler** se concentre sur la qualité des données, mais la gestion des biais et des aspects éthiques est moins automatisée.

# CONSIDERATIONS ETHIQUES DANS LE TRAITEMENT AUTOMATISE DES DONNEES

## 2.1 Clarification conceptuelle

L'essor de l'intelligence artificielle (IA) et des technologies de traitement automatisé des données a transformé la manière dont les organisations prennent des décisions, interagissent avec leurs clients et gèrent leurs opérations. Cependant, cette révolution technologique s'accompagne de défis éthiques majeurs, notamment en ce qui concerne la transparence, la responsabilité, et l'équité des systèmes automatisés. Alors que les entreprises adoptent de plus en plus ces technologies pour améliorer leur efficacité et leur compétitivité, les questions éthiques liées à l'utilisation des données et à l'automatisation des processus deviennent centrales.

Les systèmes de traitement automatisé des données, bien que puissants, peuvent introduire des biais, compromettre la vie privée des individus, et générer des décisions injustes ou discriminatoires si elles ne sont pas correctement conçues et surveillées. Par exemple, des algorithmes d'IA utilisés pour le recrutement ont montré des préjugés sexistes en favorisant les candidats masculins, reflétant les biais présents dans les données historiques utilisées pour leur entraînement. Ces exemples soulignent la nécessité de mettre en place des cadres éthiques robustes pour guider le développement et l'utilisation des technologies de traitement au-

tomatisé des données.

Dans ce contexte, il est essentiel de clarifier les concepts clés liés à l'éthique dans le traitement automatisé des données et d'explorer les principes qui peuvent guider une utilisation responsable de ces technologies. Cette section vise à fournir une compréhension approfondie des enjeux éthiques associés à l'automatisation du traitement des données, en mettant l'accent sur la responsabilité des données, l'alignement des valeurs et la redevabilité algorithmique[?].



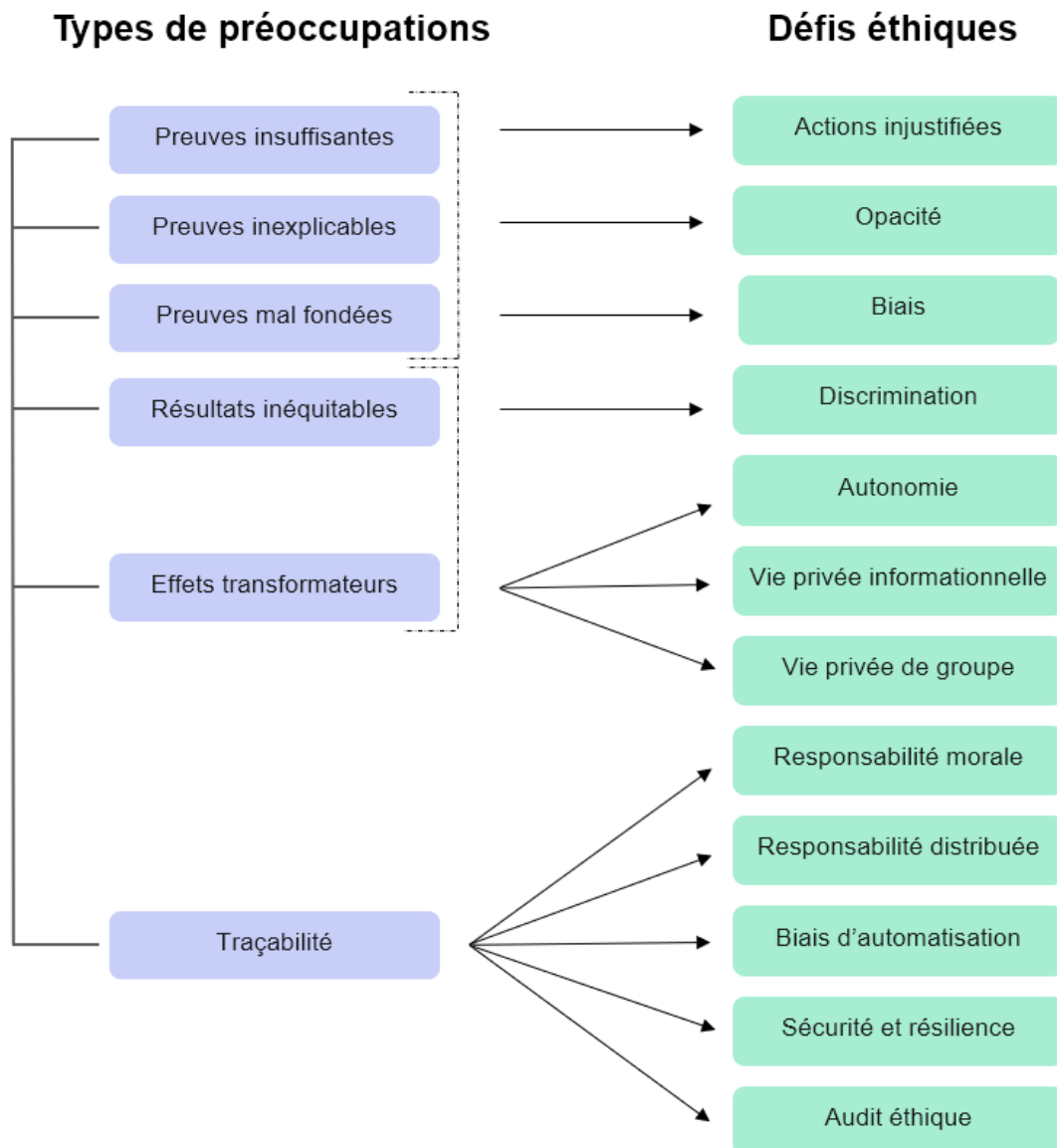


Figure 2.1 – Types de préoccupations et défis liés à l'éthique de l'IA [5]

### 2.1.1 Responsabilité des données

La responsabilité des données est un pilier central de l'éthique dans le traitement automatisé des données. Elle englobe la gestion éthique des données tout au long de leur cycle de vie, depuis leur collecte jusqu'à leur stockage, leur utilisation et leur partage. Les organisations doivent garantir que les données sont utilisées de

manière transparente, sécurisée et respectueuse de la vie privée des individus. Cela inclut la mise en place de mesures pour prévenir les violations de données, assurer la confidentialité, et obtenir le consentement éclairé des utilisateurs.

Selon une étude de l'IBM Institute for Business Value, 81 pour cent des consommateurs ont exprimé une préoccupation accrue quant à l'utilisation de leurs données par les entreprises, et 75 pour cent ont déclaré être moins enclins à faire confiance aux organisations pour la gestion de leurs informations personnelles. Ces chiffres soulignent l'importance de la responsabilité des données dans la construction de la confiance entre les entreprises et leurs clients.

### **2.1.2 Alignement des valeurs**

L'alignement des valeurs fait référence à la capacité des systèmes automatisés à prendre des décisions qui reflètent les valeurs et les normes éthiques de la société. Cela implique que les algorithmes d'IA doivent être conçus pour éviter les biais et les discriminations, tout en respectant les principes d'équité et d'inclusion. Par exemple, un système de traitement automatisé des données utilisé dans le domaine de la santé doit être capable de prendre des décisions qui respectent les droits des patients et les normes médicales éthiques.

Cependant, des cas réels montrent que les systèmes d'IA peuvent mal interpréter les données et produire des résultats biaisés. Par exemple, un algorithme utilisé pour détecter des lésions cancéreuses a appris à associer la présence d'une règle de mesure sur les photographies à un risque accru de cancer, sans comprendre la raison sous-jacente. Ce type de problème met en évidence la nécessité de concevoir des systèmes qui alignent leurs décisions sur des valeurs éthiques claires.

### **2.1.3 Redevabilité algorithmique**

La redevabilité algorithmique concerne la capacité à expliquer et à justifier les décisions prises par les systèmes automatisés. Cela inclut la transparence des processus décisionnels et la capacité à identifier les responsables en cas de résultats erronés ou préjudiciables. Par exemple, si un algorithme de recrutement

rejette injustement des candidats qualifiés, il doit être possible de retracer les raisons de cette décision et de corriger le système.

La redevabilité algorithmique est particulièrement importante dans des domaines sensibles comme la justice, la santé, ou les services financiers, où les décisions automatisées peuvent avoir un impact significatif sur la vie des individus. Selon l'étude de l'IBM Institute for Business Value, deux tiers des dirigeants techniques considèrent la redevabilité algorithmique comme une priorité majeure, reflétant la demande croissante de transparence et d'explicabilité dans les systèmes d'IA.

#### **2.1.4 Inclusion et équité**

L'inclusion et l'équité sont des principes fondamentaux dans le traitement automatisé des données. Les systèmes d'IA doivent être conçus pour éviter les discriminations et garantir que tous les individus, indépendamment de leur genre, origine ethnique ou statut socio-économique, bénéficient de décisions justes et équitables. Cela nécessite une attention particulière à la diversité des données utilisées pour entraîner les algorithmes, ainsi qu'à la détection et à la correction des biais.

Par exemple, Amazon a dû abandonner un système de recrutement automatisé qui favorisait les candidats masculins en raison de biais présents dans les données historiques. Cet exemple illustre l'importance de l'inclusion et de l'équité dans la conception des systèmes automatisés.

## **2.2 Normes éthiques dans l'IA**

Les normes éthiques dans l'intelligence artificielle (IA) sont essentielles pour garantir que les systèmes d'IA soient développés et utilisés de manière responsable, équitable et respectueuse des droits humains. Les principes éthiques de l'IA incluent la responsabilité des données, l'alignement des valeurs, la rede-

vabilité algorithmique et l'inclusion [6]. Ces principes visent à minimiser les biais, à protéger la vie privée, et à assurer que les décisions prises par les systèmes d'IA soient transparentes et justifiables. Par exemple, l'étude de l'IBM Institute for Business Value souligne que 81 pour cent des consommateurs sont préoccupés par l'utilisation de leurs données, ce qui renforce la nécessité de normes éthiques strictes pour maintenir la confiance du public. De plus, il est mis en avant cinq principes clés pour une IA éthique : la bienveillance, la non-malfaisance, l'indépendance, la justice, et l'interprétabilité. Ces principes sont alignés avec les recommandations de l'UNESCO et de l'IEEE, qui insistent sur la nécessité de créer des systèmes d'IA qui profitent à l'humanité tout en respectant les droits fondamentaux. Enfin, l'étude mentionne que des pays comme la Jordanie intègrent désormais des programmes éducatifs sur l'éthique de l'IA, ce qui montre une prise de conscience croissante de l'importance de ces normes dans le développement technologique.[7].

## TRANSPARENCE DANS LES OUTILS AUTOMATISES

La transparence dans les outils automatisés est un pilier central pour garantir la confiance et la compréhension des systèmes d'IA. La transparence implique que les organisations doivent être claires sur la source des données, le fonctionnement des algorithmes et les objectifs des systèmes d'IA [6]. Cela permet aux utilisateurs de comprendre comment les décisions sont prises et de détecter d'éventuelles erreurs ou biais. Par exemple, les études révèlent que deux tiers des dirigeants techniques considèrent la redevabilité algorithmique comme une priorité, ce qui inclut la capacité à expliquer les décisions prises par les systèmes d'IA. De plus, la transparence est cruciale pour éviter les discriminations et les injustices. Les développeurs d'IA doivent fournir des explications claires sur les processus décisionnels, en particulier dans des domaines sensibles comme la santé, la justice et l'emploi. Par exemple, les études mentionnent un cas où un algorithme de recrutement d'Amazon a été abandonné en raison de biais sexistes, ce qui illustre l'importance de la transparence pour identifier et corriger de tels problèmes. Enfin, la transparence est également liée à la protection des données, car les utilisateurs doivent savoir comment leurs informations sont collectées, utilisées et protégées.

Malgré son importance, la transparence dans les systèmes d'IA pose plusieurs défis. L'un des principaux défis est la complexité des algorithmes d'ap-

prentissage profond, qui fonctionnent souvent comme des "boîtes noires"[6]. Cela signifie que même les développeurs peuvent avoir du mal à expliquer comment les décisions sont prises, ce qui rend la transparence difficile à mettre en œuvre. Les attentes des consommateurs en matière de transparence sont élevées, mais les entreprises peuvent hésiter à divulguer des informations sensibles sur leurs algorithmes pour des raisons de propriété intellectuelle ou de concurrence. De plus, il existe de nombreux défis liés à la transparence, en particulier dans les pays en développement comme la Jordanie. Le manque de ressources techniques et de formation peut limiter la capacité des organisations à rendre leurs systèmes d'IA transparents. Par exemple, les algorithmes de détection de la couleur de peau et du genre peuvent avoir des taux d'erreur élevés (jusqu'à 16 pour cent pour les visages noirs), ce qui souligne la difficulté de garantir une transparence totale dans des systèmes complexes. Enfin, un autre défi est la nécessité de trouver un équilibre entre la transparence et la protection des données, car une trop grande transparence pourrait exposer des informations sensibles sur les individus.[8].

Ce chapitre nous a permis de démontrer que l'automatisation du pré-traitement des données permet de gagner en efficacité, mais elle soulève des défis majeurs en matière d'éthique et de transparence. Les solutions actuelles, bien qu'avancées, nécessitent encore des ajustements pour répondre aux exigences d'équité et de transparence imposées par les législations en vigueur. Des normes éthiques, comme celles de l'UNESCO, encadrent l'utilisation de l'IA et insistent sur l'importance de préserver les droits de l'homme dans un monde dans lequel les décisions sont de plus en plus influencées par les données et les algorithmes.

---

---

## *Deuxième partie*

---

### **MATERIELS ET METHODES**

## MATERIELS

Cette section présente en détail l’environnement technique, les outils et les méthodes utilisés pour développer FairAutoClean, l’outil de prétraitement automatique et éthique des données.

### 4.1 Analyse des besoins fonctionnels et techniques

#### 4.1.1 Analyse des besoins fonctionnels

Le développement de FairAutoClean répond à plusieurs besoins fonctionnels essentiels :

#### **Prétraitement automatique et efficient des données**

L’objectif de cette étape est de préparer les données pour l’analyse en les nettoyant et en les transformant de manière à les rendre exploitables. Cela inclut la suppression des doublons, la gestion des valeurs manquantes, la détection et le traitement des valeurs aberrantes, la conversion des colonnes de date, l’encodage des colonnes catégoriques, et la normalisation des colonnes numériques.

##### — **Suppression des doublons :**

- Identifier et éliminer les enregistrements redondants pour garantir l’unicité des données.
- Assurer que chaque ligne du jeu de données représente une entrée unique, ce qui améliore la qualité des analyses ultérieures.



— **Gestion des valeurs manquantes :**

- **Imputation par régression** : Utiliser des modèles de régression pour prédire et remplacer les valeurs manquantes dans les colonnes numériques.
- **Imputation KNN** : Utiliser l'algorithme des k-plus proches voisins pour imputer les valeurs manquantes en tenant compte des relations entre les variables.
- **Suppression intelligente** : Supprimer les lignes contenant un nombre excessif de valeurs manquantes tout en préservant la structure globale du jeu de données.

— **Gestion des valeurs aberrantes :**

- **Détection par IQR** : Utiliser la méthode de l'intervalle interquartile (IQR) pour identifier les valeurs extrêmes dans les données.
- **Traitement par winsorisation** : Limiter l'impact des valeurs aberrantes en les remplaçant par des valeurs situées aux bornes de l'intervalle acceptable.

— **Détection et conversion des colonnes de date :**

- **Identification automatique** : Détecter les formats de date dans les colonnes et les convertir en un format standard.
- **Standardisation** : Uniformiser les formats de date pour faciliter les analyses temporelles.

— **Encodage des colonnes catégoriques :**

- **Label Encoding** : Convertir les variables catégoriques ordinales en valeurs numériques pour les modèles d'apprentissage automatique.
- **One-Hot Encoding** : Créer des colonnes binaires pour chaque catégorie dans les variables nominales, permettant une représentation sans ordre implicite.

— **Normalisation des colonnes numériques :**

- **Standardisation** : Transformer les données pour avoir une moyenne de 0 et un écart-type de 1, ce qui est utile pour les algorithmes sensibles à l'échelle.
- **Min-Max Scaling** : Redimensionner les données dans un intervalle spécifique (par exemple,  $[0, 1]$ ) pour les algorithmes basés sur des distances.
- **Robust Scaling** : Utiliser des statistiques robustes (médiane et intervalle interquartile) pour normaliser les données en présence de valeurs aberrantes.

### Réduction de dimension des données

La réduction de dimension vise à simplifier les données en réduisant le nombre de variables tout en conservant l'information essentielle. Cela améliore l'efficacité des modèles d'apprentissage automatique et réduit le temps de calcul.

- **PCA (Principal Component Analysis)** :
  - **Préservation de la variance** : Réduire la dimensionnalité en projetant les données sur des axes orthogonaux qui maximisent la variance.
  - **Configuration flexible** : Permettre à l'utilisateur de choisir le nombre de composantes principales ou le pourcentage de variance à conserver.
- **Autoencoders** :
  - **Réduction non-linéaire** : Utiliser des réseaux de neurones pour apprendre une représentation compressée des données, adaptée aux structures complexes.
  - **Architecture adaptative** : Adapter la profondeur et la complexité du modèle en fonction de la nature des données.

### Détection automatique des colonnes sensibles

La détection des colonnes sensibles est cruciale pour garantir la conformité aux réglementations éthiques et légales, en particulier en ce qui concerne la protection des données personnelles.

— **Détection algorithmique :**

- **Listes de mots-clés :** Utiliser des listes prédéfinies pour identifier les colonnes susceptibles de contenir des informations sensibles (par exemple, "nom", "adresse", "salaire").
- **Analyse des patterns :** Examiner les noms de colonnes pour détecter des motifs courants associés à des données sensibles.

— **Détection par IA :**

- **Modèles de langage (LLM) :** Utiliser des modèles de langage avancés pour analyser le contexte des colonnes et évaluer leur sensibilité.
- **Évaluation de la sensibilité :** Attribuer un score de sensibilité sur une échelle de 0 à 10 pour chaque colonne, en fonction du risque pour la vie privée.

## **Détection des biais algorithmiques**

Dans certains cas de figures, il peut arriver que les données aient déjà été traitées au préalable, ou qu'un traitement supplémentaire a été prévu d'être fait. Dans ces cas, il est utile de pouvoir analyser le code qui fera ce prétraitement, afin de détecter les potentiels biais qu'il pourrait introduire.

— **Analyse syntaxique du code :**

- **Identification des patterns discriminatoires :** Examiner le code source pour détecter des motifs ou des opérations qui pourraient introduire des biais, tels que des filtres ou des sélections basées sur des caractéristiques sensibles (genre, race, âge, etc.).
- **Vérification des opérations de filtrage :** Analyser les algorithmes de filtrage et de sélection pour s'assurer qu'ils ne favorisent pas ou ne défavorisent pas injustement certains groupes d'individus.

— **Analyse par IA :**

- **Utilisation de modèles de langage (LLM) :** Employer des modèles de langage avancés pour évaluer les implications éthiques du code, en identifiant les zones à risque et en suggérant des améliorations.

- **Génération de recommandations** : Produire des recommandations concrètes pour modifier le code afin de réduire les biais et d'améliorer l'équité des décisions algorithmiques.

## Visualisation et audit

La visualisation et l'audit des données et des processus sont cruciaux pour assurer la transparence, la redevabilité et la qualité des systèmes d'IA. Ces étapes permettent de comprendre les données, de détecter les anomalies et de documenter les opérations effectuées.

- **Exploratory Data Analysis (EDA)** :
  - **Génération de rapports de profiling** : Créer des rapports détaillés qui résument les caractéristiques des données, telles que les distributions, les valeurs manquantes et les corrélations entre variables.
  - **Visualisations statistiques** : Produire des graphiques et des diagrammes (histogrammes, boîtes à moustaches, etc.) pour visualiser les distributions des données et identifier les tendances ou les anomalies.
- **Traçabilité** :
  - **Journalisation des opérations** : Enregistrer de manière détaillée toutes les opérations effectuées sur les données, y compris les transformations, les nettoyages, et les analyses, pour assurer une transparence totale.
  - **Rapport d'audit complet** : Générer un rapport d'audit qui inclut des métriques de performance, des indicateurs de qualité des données, et des évaluations des biais, permettant une vérification externe et une conformité aux normes éthiques et réglementaires.

### 4.1.2 Analyse des besoins techniques

Pour développer un outil automatisé de prétraitement éthique et transparent des données, plusieurs besoins techniques doivent être identifiés et satisfaits. Ces besoins incluent le choix d'un langage de programmation adapté, la sélection d'un environnement de développement efficace, et l'utilisation de bibliothèques et d'outils spécialisés pour la gestion des données, l'apprentissage automatique, et la visualisation. Voici une liste des outils et technologies qui seront utilisés pour répondre à ces besoins :

- **Langage de programmation** : Python
- **Environnement de développement** : VSCode, Git, et les environnements virtuels
- **Bibliothèques de data science** : Pandas, NumPy, Scikit-learn
- **Outils de visualisation** : Matplotlib, Seaborn
- **Gestion des dépendances** : Pip, Conda
- **Manipulation des LLMs** : PydanticAI
- **Détection et gestion des biais** : AI Fairness 360

### Langage de programmation

Python a été choisi comme langage de programmation principal pour ce projet en raison de ses nombreux avantages dans le domaine de la data science et de l'intelligence artificielle.

- **Richesse en bibliothèques** : Python dispose d'un écosystème riche en bibliothèques spécialisées pour le traitement des données (Pandas, NumPy), l'apprentissage automatique (Scikit-learn, TensorFlow), et la visualisation (Matplotlib, Seaborn). Ces bibliothèques permettent de développer des solutions robustes et efficaces.
- **Communauté active** : Python bénéficie d'une communauté de développeurs très active, ce qui facilite l'accès à des ressources, des tutoriels, et des solutions aux problèmes courants.

- **Facilité d'intégration** : Python s'intègre facilement avec d'autres outils et technologies, ce qui en fait un choix idéal pour des projets nécessitant une interopérabilité avec des systèmes existants.



Figure 4.1 – Logo de Python, le langage de programmation choisi pour ce projet.

## Environnement de développement

Un environnement de développement moderne et efficace est essentiel pour assurer la productivité et la qualité du code. Les outils suivants seront utilisés :

- **VSCode** :
  - Un éditeur de code léger mais puissant, avec un support natif pour Python.
  - Offre des fonctionnalités avancées comme le débogage, la complétion de code, et l'intégration avec Git.
  - Extensible via des extensions pour répondre à des besoins spécifiques (par exemple, des extensions pour la data science ou l'IA).



Figure 4.2 – Logo de VSCode, l'éditeur de code utilisé pour ce projet.

— **Git :**

- Un système de contrôle de version distribué pour suivre les modifications du code et collaborer efficacement.
- Permet de gérer les branches, les fusions, et les versions du projet de manière structurée.
- Intégré avec des plateformes comme GitHub ou GitLab pour le stockage et le partage du code.



Figure 4.3 – Logo de Git, le système de contrôle de version utilisé pour ce projet.

— **Environnement virtuel :**

- Utilisation d'environnements virtuels (par exemple, avec 'venv' ou 'conda') pour isoler les dépendances du projet.
- Permet de gérer les versions des bibliothèques et d'éviter les conflits entre projets.
- Facilite la reproductibilité des résultats en garantissant que tous les membres de l'équipe utilisent les mêmes versions des outils.

## **Manipulation des LLMs avec PydanticAI**

Pour intégrer des modèles de langage (LLMs) dans le projet, la bibliothèque PydanticAI a été choisie. Cette bibliothèque offre des fonctionnalités avancées pour interagir avec des LLMs tout en garantissant une gestion structurée des données et des réponses.

- **Structuration des données** : PydanticAI permet de définir des schémas de données robustes pour les entrées et les sorties des LLMs, ce qui facilite la validation et la manipulation des données.
- **Intégration avec les LLMs** : La bibliothèque supporte l'intégration avec des modèles de langage populaires comme GPT, BERT, et d'autres, permettant des interactions fluides et efficaces.
- **Gestion des erreurs** : PydanticAI inclut des mécanismes de gestion des erreurs pour traiter les réponses inattendues ou invalides des LLMs, améliorant ainsi la robustesse du système.



Figure 4.4 – Logo de PydanticAI

## Détection et gestion des biais avec AI Fairness 360

Pour garantir que les systèmes d'IA soient équitables et non discriminatoires, la bibliothèque AI Fairness 360 (AIF360) a été sélectionnée. Cette bibliothèque open-source, développée par IBM, fournit des outils pour détecter, mesurer, et atténuer les biais dans les modèles d'IA.

- **Détection des biais** : AI Fairness 360 inclut des algorithmes pour identifier les biais dans les données et les modèles, en analysant les disparités entre différents groupes (par exemple, genre, race, âge).
- **Atténuation des biais** : La bibliothèque propose des techniques pour réduire les biais, telles que la réévaluation des poids des données, l'ajustement des seuils de décision, et l'utilisation de méthodes de rééchantillonnage.
- **Métriques d'équité** : AI Fairness 360 fournit des métriques pour évaluer l'équité des modèles, comme l'égalité des chances, la parité démographique, et l'impact disparate.





## Ai Fairness 360

Figure 4.5 – Logo de Ai Fairness 360.

### 4.1.3 Environnement technique

Pour garantir une exécution fluide et efficace des tâches de développement et d'analyse, il est essentiel de disposer d'un environnement technique adapté. Cette section décrit la configuration matérielle utilisée pour ce projet, ainsi que les spécifications techniques qui permettent de supporter les charges de travail liées au prétraitement des données, à l'apprentissage automatique, et à la gestion des modèles de langage.

#### Configuration matérielle

- **Ordinateur portable** : HP Laptop 14s-dq2395nia
- **Processeur** : 11th Gen Intel(R) Core(TM) i5-1135G7 @ 2.40GHz (4 cœurs, 8 threads)
- **Mémoire RAM** : 16,0 Go (DDR4)
- **Système d'exploitation** : Windows 11 (64 bits)

## METHODES

### 5.1 Architecture logicielle

L'architecture de FairAutoClean a été conçue avec une approche modulaire pour garantir une maintenabilité optimale et une extensibilité future. Cette structure permet de séparer les responsabilités fonctionnelles en modules distincts, facilitant ainsi la gestion du code, les tests unitaires et l'ajout de nouvelles fonctionnalités. Chaque module est conçu pour être indépendant et réutilisable, ce qui améliore la flexibilité du système et permet une intégration aisée avec d'autres outils ou frameworks. L'architecture repose sur une interface unifiée, gérée par le module `interface.py`, qui sert de point d'entrée pour toutes les interactions avec le système.

#### Structure générale

- **Architecture en couches avec séparation claire des responsabilités :**  
L'architecture de FairAutoClean est organisée en plusieurs couches logiques, chacune ayant une responsabilité spécifique. Cette séparation permet une meilleure gestion des dépendances et une plus grande clarté dans l'organisation du code. Les couches principales incluent la gestion des données, le traitement des données, l'analyse des biais et la génération de rapports.
- **Modules indépendants et réutilisables :** Chaque fonctionnalité du système est encapsulée dans un module distinct. Par exemple, le module `DataProcessor`

est responsable du chargement et du prétraitement des données, tandis que le module `FairnessAnalyzer` se concentre sur la détection et la mitigation des biais. Cette modularité permet de réutiliser les composants dans d'autres projets ou de les mettre à jour indépendamment sans affecter l'ensemble du système.

- **Interface unifiée via le module `interface.py`** : Le module `interface.py` sert de point d'entrée unique pour toutes les interactions avec FairAutoClean. Il gère la communication entre les différents modules et expose une API simple et cohérente pour les utilisateurs. Cette interface unifiée simplifie l'intégration de FairAutoClean dans des workflows existants et permet une configuration facile via des fichiers JSON ou des paramètres en ligne de commande.

### Composants principaux

- **AutoClean** : Cette classe est le cœur du système. Elle orchestre le pipeline de nettoyage des données en coordonnant les différentes étapes du prétraitement, telles que la gestion des valeurs manquantes, la détection des valeurs aberrantes et la normalisation des données. Elle garantit que chaque étape est exécutée dans le bon ordre et que les données sont correctement transformées avant d'être analysées.
- **DataProcessor** : Ce module est responsable du chargement des données brutes et de leur préparation pour le traitement. Il inclut des fonctionnalités pour la validation des données, la conversion des formats, et la gestion des erreurs. Le `DataProcessor` assure que les données sont dans un état exploitable avant d'être transmises aux autres modules.
- **FairnessAnalyzer** : Ce composant est dédié à la détection et à la mitigation des biais dans les données. Il utilise des algorithmes avancés pour identifier les disparités entre différents groupes (par exemple, genre, race, âge) et propose des corrections pour garantir l'équité des résultats. Le `FairnessAnalyzer` est essentiel pour respecter les normes éthiques et légales en matière de traitement des données.

- **AICodeAnalyzer** : Ce module analyse le code source utilisé pour le pré-traitement des données afin de détecter d'éventuels biais algorithmiques. Il examine les opérations de filtrage, de transformation, et de sélection pour s'assurer qu'elles ne favorisent pas ou ne défavorisent pas injustement certains groupes. Le AICodeAnalyzer génère également des recommandations pour améliorer l'équité du code.
- **ReportGenerator** : Ce composant est chargé de générer des rapports détaillés sur les opérations effectuées par FairAutoClean. Il produit des métriques de performance, des indicateurs de qualité des données, et des évaluations des biais. Ces rapports sont essentiels pour assurer la transparence et la traçabilité du processus de prétraitement, permettant aux utilisateurs de comprendre et d'auditer les actions entreprises par le système.

## 5.2 Flux de données et architecture système

Le système FairAutoClean est organisé en plusieurs couches logiques, chacune ayant un rôle spécifique dans le processus de prétraitement des données. Cette architecture en couches permet une séparation claire des responsabilités, une meilleure gestion des dépendances et une extensibilité accrue. Voici une description de chaque couche et de ses composants :

1. **Couche d'entrée** : Cette couche est responsable de la réception et de la préparation des données et des configurations nécessaires au fonctionnement du système. Elle comprend deux éléments principaux :
  - **Données d'entrée** : Les données brutes sont chargées dans le système en format CSV ou Excel, qui sont des formats de données largement utilisés dans le domaine des data sciences en particulier et dans tous les secteurs en général
  - **Configuration JSON** : Un fichier de configuration au format JSON est utilisé pour paramétrer le comportement du système. Ce fichier contient des informations telles que les méthodes de prétraitement à appliquer, les seuils pour la détection des valeurs aberrantes, les paramètres de réduction de dimensionnalité, etc. La configuration JSON

permet une personnalisation fine du processus de prétraitement sans nécessiter de modifications du code source.

2. **Couche d'interface** : Cette couche fournit des interfaces pour interagir avec le système, que ce soit via une interface en ligne de commande (CLI) ou une API Python. Elle comprend :

- **Interface CLI** : Une interface en ligne de commande permet aux utilisateurs de configurer et d'exécuter FairAutoClean avec des paramètres personnalisés. Elle est idéale pour les utilisateurs techniques ou pour l'intégration dans des scripts automatisés.
- **Package Python** : FairAutoClean est également disponible sous forme de package Python, permettant une intégration facile dans des projets existants. Les utilisateurs peuvent importer les modules et les fonctions directement dans leur code.
- **Gestion de la configuration** : Ce composant permet aux utilisateurs de configurer le système via des fichiers JSON ou des paramètres en ligne de commande. Il garantit une flexibilité maximale dans la personnalisation du processus de prétraitement.

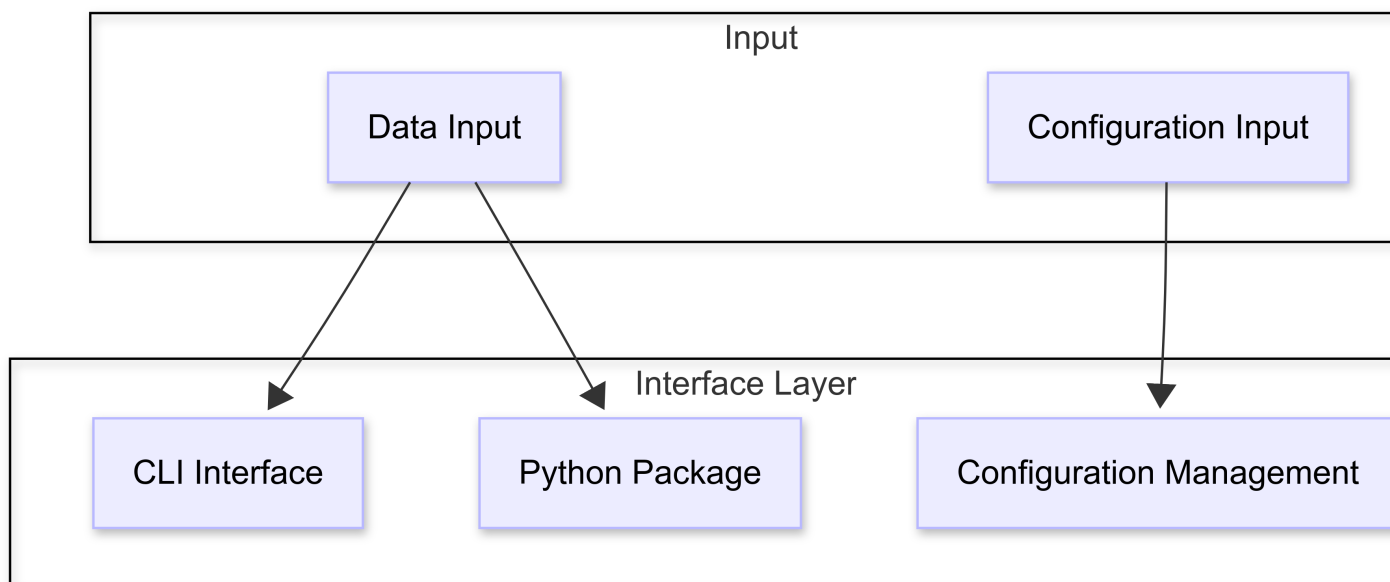


Figure 5.1 – Représentation des composants de la couche d'entrée des données

3. **Couche de traitement principal** : Cette couche est le cœur du système, où les données sont nettoyées, transformées et préparées pour l'analyse. Elle comprend plusieurs modules clés :
- **Processeur de données** : Ce module charge les données brutes, valide leur intégrité, et les prépare pour le traitement. Il gère également les erreurs de format ou les données manquantes en appliquant des stratégies prédéfinies (par exemple, l'imputation ou la suppression des valeurs manquantes).
  - **Nettoyeur de données** : Ce module effectue des opérations de nettoyage telles que la suppression des doublons, la gestion des valeurs aberrantes, et la normalisation des données. Il garantit que les données sont propres et cohérentes avant d'être transmises aux étapes suivantes.
  - **Normalisateur** : Ce composant applique des transformations pour standardiser les données, comme la normalisation des valeurs numériques ou l'encodage des variables catégorielles. Ces transformations sont essentielles pour garantir que les données sont adaptées aux algorithmes d'apprentissage automatique.
  - **Réduction de dimensionnalité** : Ce module réduit la complexité des données en supprimant les variables redondantes ou peu informatives. Des techniques telles que l'Analyse en Composantes Principales (PCA) ou les autoencodeurs sont utilisées pour conserver l'information essentielle tout en réduisant le temps de calcul.
  - **Ingénierie des caractéristiques** : Ce module crée de nouvelles variables ou caractéristiques à partir des données existantes pour améliorer la performance des modèles d'apprentissage automatique. Il peut inclure des opérations telles que la création d'interactions entre variables ou l'extraction de caractéristiques temporelles.

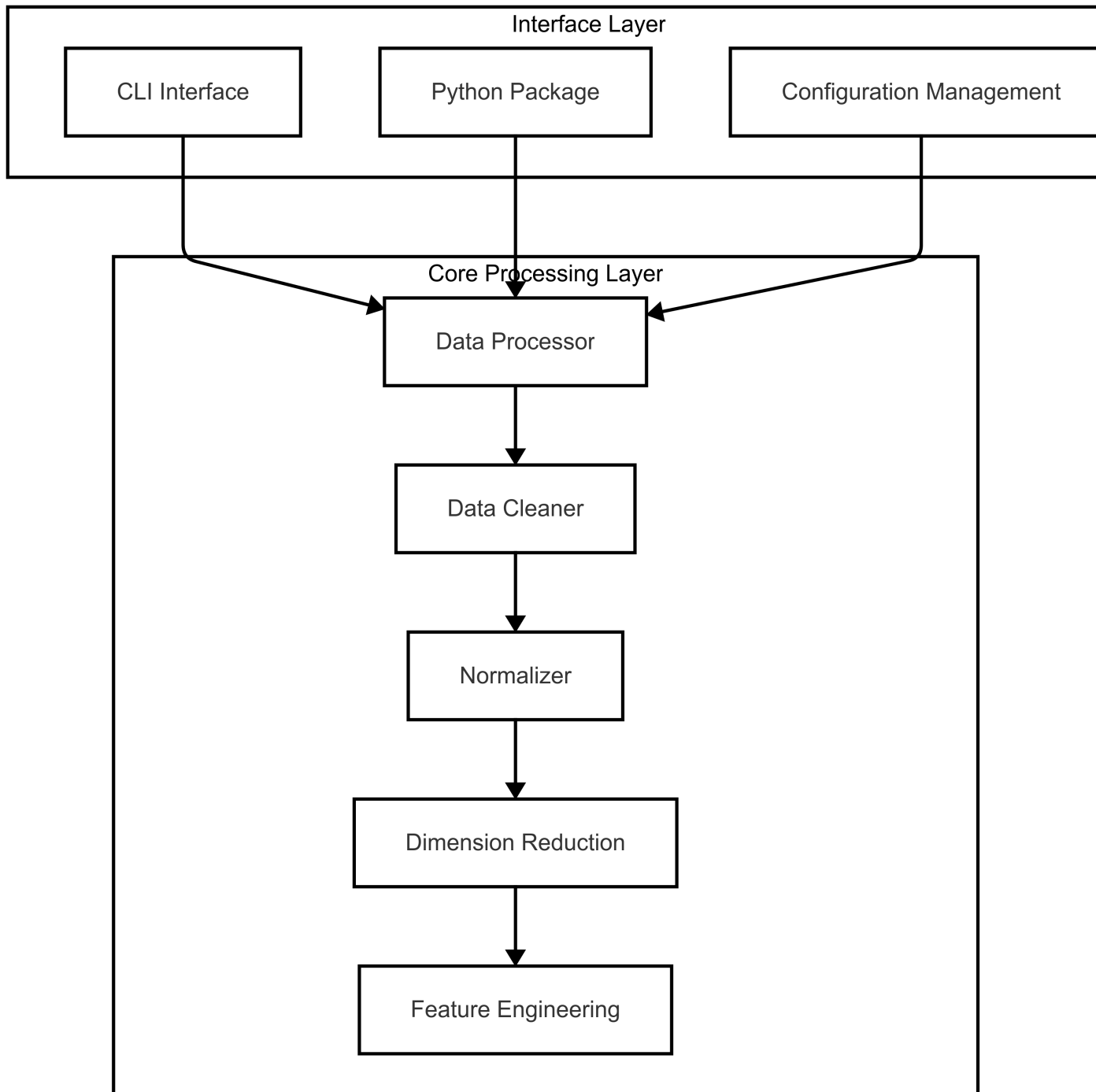


Figure 5.2 – Représentation des composants de la couche de traitement principal

4. **Couche d'équité et d'analyse** : Cette couche est dédiée à la détection et à la mitigation des biais, ainsi qu'à l'analyse éthique des données et du code. Elle comprend les composants suivants :

- **Analyseur d'équité** : Ce module utilise des algorithmes pour identifier les biais dans les données, tels que les disparités entre groupes

démographiques. Il propose des corrections pour garantir que les données sont équitables et conformes aux normes éthiques.

- **Analyseur de code** : Ce composant examine le code source utilisé pour le prétraitement des données afin de détecter d'éventuels biais algorithmiques. Il génère des rapports détaillés sur les risques éthiques et propose des recommandations pour améliorer l'équité du code.
- **Système d'audit** : Ce module enregistre toutes les opérations effectuées par le système, y compris les transformations des données, les détections de biais, et les corrections appliquées. Il garantit une transparence totale et permet aux utilisateurs de vérifier et d'auditer le processus de prétraitement.



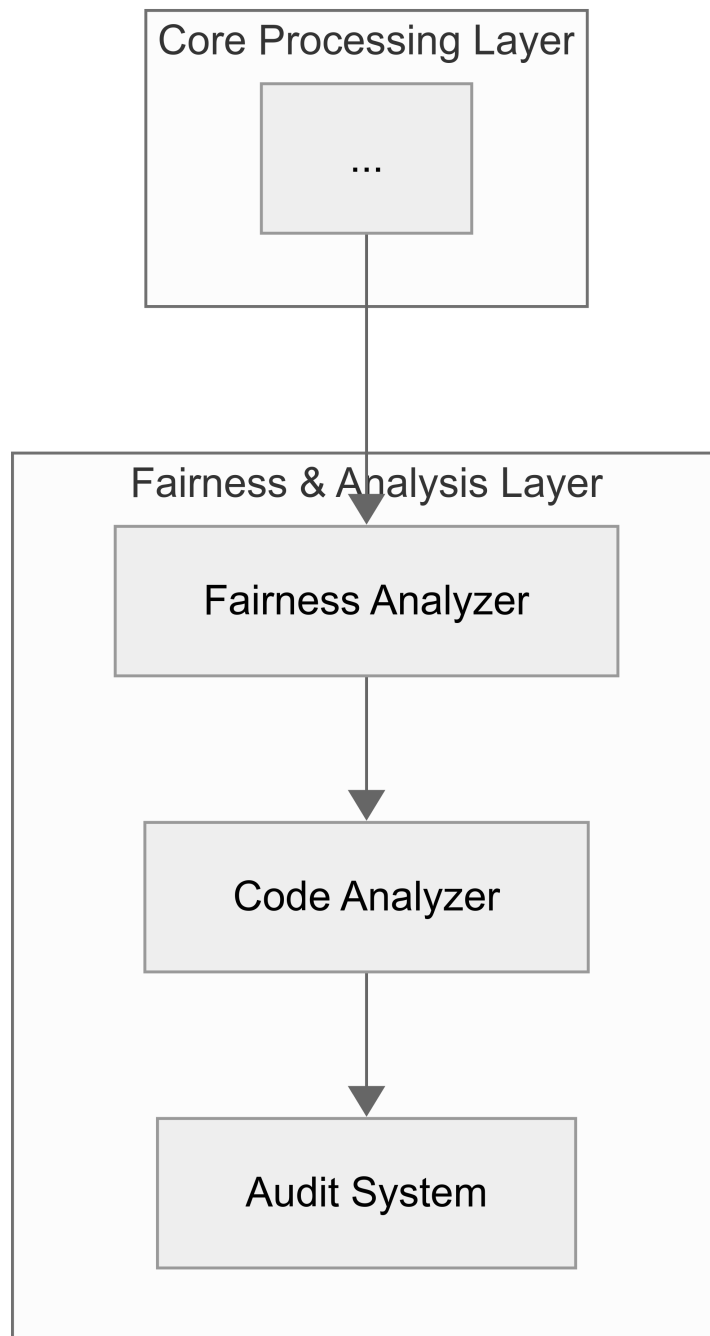


Figure 5.3 – Représentation des composants de la couche d'équité et d'analyse

5. **Système d'agent IA** : Cette couche intègre des modèles de langage (LLMs) et des agents intelligents pour améliorer les fonctionnalités du système. Elle comprend :

- **Intégration LLM** : Des modèles de langage avancés, tels que GPT, sont utilisés pour analyser le contexte des données et générer des re-

commandations éthiques. Ces modèles permettent également de détecter des biais subtils qui pourraient échapper aux algorithmes traditionnels.

- **Gestionnaire de prompts système** : Ce composant génère des prompts adaptés pour interagir avec les LLMs, garantissant des réponses précises et pertinentes. Il est essentiel pour maximiser l'efficacité des interactions avec les modèles de langage.

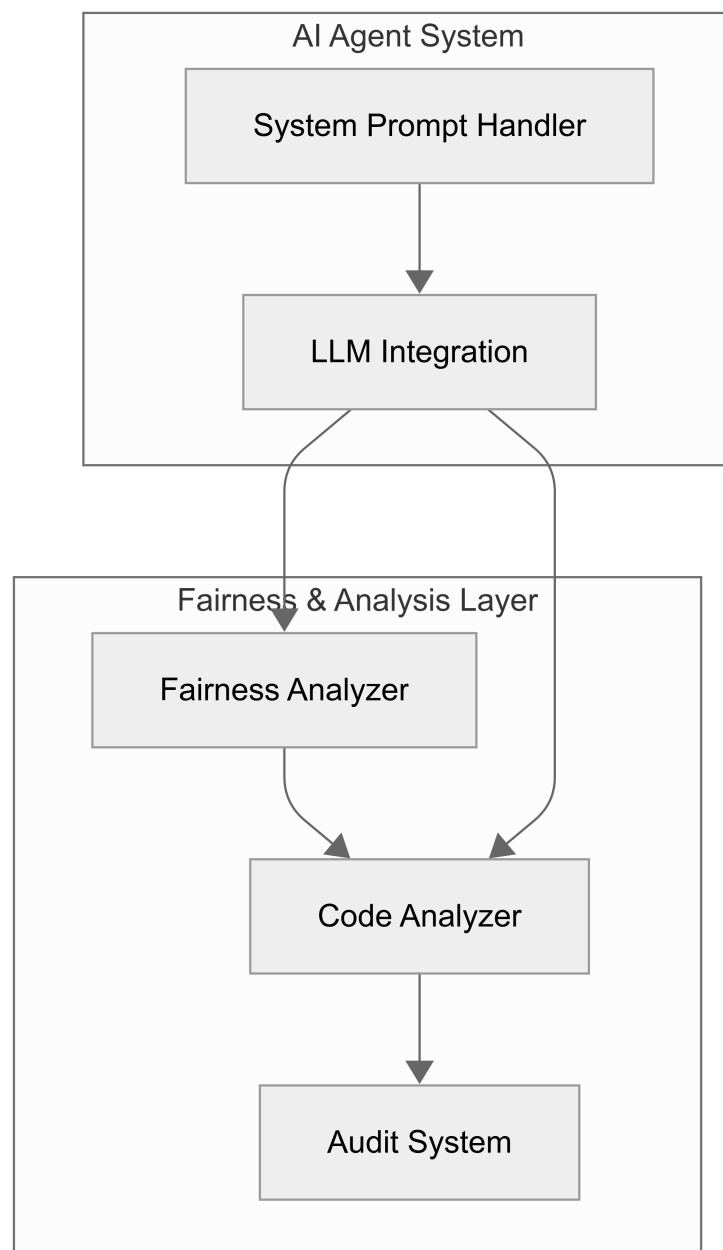


Figure 5.4 – Représentation des composants du système d'agent IA

6. **Sortie et rapports** : Cette couche est responsable de la génération des résultats finaux et des rapports détaillés. Elle comprend :

- **Générateur de rapports** : Ce module produit des rapports complets sur les opérations effectuées, les biais détectés, et les corrections appliquées. Ces rapports sont essentiels pour assurer la transparence et la traçabilité du processus.
- **Générateur de profils** : Ce composant crée des profils détaillés des données, incluant des statistiques descriptives, des visualisations, et des indicateurs de qualité. Ces profils aident les utilisateurs à comprendre les caractéristiques des données avant et après le prétraitement.
- **Générateur de logs d'audit** : Ce module enregistre toutes les actions entreprises par le système, créant une piste d'audit complète pour chaque exécution. Cette fonctionnalité est cruciale pour répondre aux exigences réglementaires et éthiques.

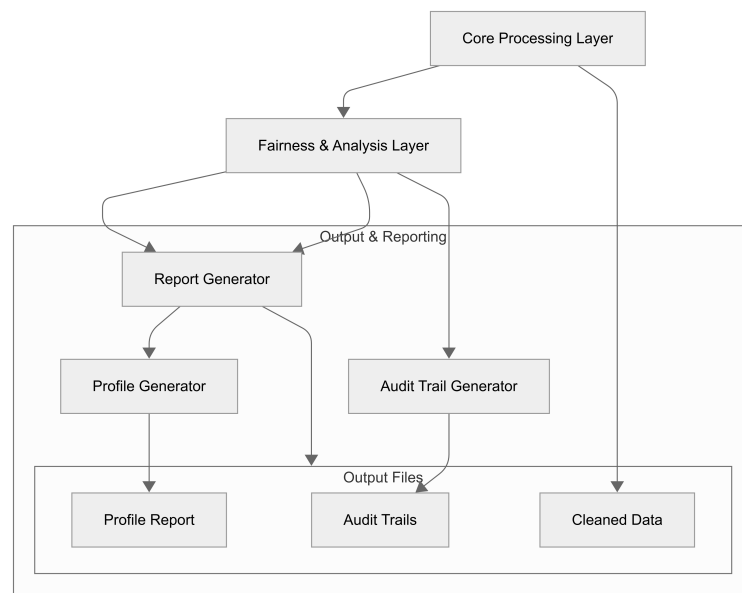


Figure 5.5 – Représentation des composants de la partie sortie et rapports

### 5.3 Configuration et utilisation de l'outil

#### Listing 5.1 – Interface CLI

```
def main() :  
    parser.add_argument( '--config' , type=str , required=True)  
    parser.add_argument( '--dataset' , type=str , required=True)  
    parser.add_argument( '--output' , type=str , required=True)
```

---

Listing 5.2 – Pipeline principal

```
def process_dataset(config_path , dataset_path , output_path) :  
    # 1. Chargement et validation de la configuration  
    # 2. Prétraitement des données  
    # 3. Analyse d'équité  
    # 4. Génération des rapports  
    # 5. Sauvegarde des résultats
```

Cette implémentation détaillée montre comment les différents composants s'articulent pour former un système cohérent de prétraitement équitable des données.

# RÉFÉRENCES BIBLIOGRAPHIQUES

- [1] M. Bilal, G. Ali, M. W. Iqbal, M. Anwar, M. S. A. Malik et R. A. Kadir, “Auto-Prep : Efficient and Automated Data Preprocessing Pipeline,” *IEEE Access*, vol. 10, p. 107 764–107 784, 2022, conference Name : IEEE Access. [En ligne]. Disponible : <https://ieeexplore.ieee.org/document/9856663/?arnumber=9856663>
- [2] □. T. Post, “Essential Open-Source Tools for Bias Detection and Mitigation,” sept. 2024. [En ligne]. Disponible : <https://www.turingpost.com/p/ai-fairness-tools>
- [3] P. I. Ouedraogo, “AUTOMATISATION DU PRE-TRAITEMENT DES DONNEES PAR L’OPTIMISATION METAHEURISTIQUE.”
- [4] A. Mumuni et F. Mumuni, “Automated data processing and feature engineering for deep learning and big data applications : A survey,” *Journal of Information and Intelligence*, janv. 2024. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S2949715924000027>
- [5] “Les défis éthiques communément rencontrés en matière d’IA - Droits Humains et Biomédecine.” [En ligne]. Disponible : <https://www.coe.int/fr/web/bioethics/common-ethical-challenges-in-ai#%7B%22123745781%22:%5B%7D>
- [6] “Advancing AI ethics beyond compliance - From principles to practice.”
- [7] “Éthique de l’intelligence artificielle | UNESCO.” [En ligne]. Disponible : <https://www.unesco.org/fr/artificial-intelligence/recommendation-ethics>

- [8] “Transparency you can trust : Transparency requirements for artificial intelligence between legal norms and contextual concerns.” [En ligne]. Disponible : <https://journals.sagepub.com/doi/epub/10.1177/2053951719860542>

# TABLE DES MATIÈRES

<b>INTRODUCTION GÉNÉRALE</b>	<b>1</b>
<b>I SYNTHÈSE BIBLIOGRAPHIQUE</b>	<b>5</b>
<b>1 AUTOMATISATION DU PRE-TRAITEMENT DES DONNÉES</b>	<b>6</b>
1.1 Clarification conceptuelle . . . . .	8
1.2 Solutions existantes . . . . .	11
1.2.1 Description et comparaison de quelques outils . . . . .	11
1.2.2 Google AutoML . . . . .	11
1.2.3 H2O.ai . . . . .	12
1.2.4 DataRobot . . . . .	12
1.2.5 Trifacta Wrangler (DataWrangler) . . . . .	13
1.2.6 IBM Watson Studio (AutoAI) . . . . .	13
1.2.7 Tableau Comparatif . . . . .	14
1.3 Analyse des Axes de Comparaison . . . . .	15
<b>2 CONSIDERATIONS ETHIQUES DANS LE TRAITEMENT AUTOMATISE DES DONNEES</b>	<b>16</b>
2.1 Clarification conceptuelle . . . . .	16
2.1.1 Responsabilité des données . . . . .	18
2.1.2 Alignement des valeurs . . . . .	19
2.1.3 Redevabilité algorithmique . . . . .	19



2.1.4	Inclusion et équité . . . . .	20
2.2	Normes éthiques dans l'IA . . . . .	20
<b>3</b>	<b>TRANSPARENCE DANS LES OUTILS AUTOMATISES</b>	<b>22</b>
<b>II</b>	<b>MATERIELS ET METHODES</b>	<b>24</b>
<b>4</b>	<b>MATERIELS</b>	<b>25</b>
4.1	Analyse des besoins fonctionnels et techniques . . . . .	25
4.1.1	Analyse des besoins fonctionnels . . . . .	25
4.1.2	Analyse des besoins techniques . . . . .	30
4.1.3	Environnement technique . . . . .	34
<b>5</b>	<b>METHODES</b>	<b>35</b>
5.1	Architecture logicielle . . . . .	35
5.2	Flux de données et architecture système . . . . .	37
5.3	Configuration et utilisation de l'outil . . . . .	44
	<b>RÉFÉRENCES BIBLIOGRAPHIQUES</b>	<b>46</b>