Structuration et documentation des données pour l'IA

INFORMATIONS GÉNÉRALES

Type de formation : Formation continue

Domaine : IA, Big Data et Bases de données

Filière : IA

Rubrique : Certification ATLAS : CISIA (Actions co.)

Éligible au CPF : Non

Action collective : Oui
Code ACO : CISIA

Code de formation : CISIA-DOC

€ Tarifs

Prix public : 1000 €

Tarif & financement :
Financement possilble via les Actions Collectives ATLAS ou le Plan de Formation.

Lieux & Horaires

Durée : 8 heures

Rythme : 9h-13h et 14h-18h

Délai d'accès :
Jusqu'a 8 jours avant le début de la formation, sous condition d'un dossier d'insciption complet

Prochaines sessions

Consultez-nous pour les prochaines sessions.

PRÉSENTATION

Objectifs & compétences

Documenter le flux de traitement des données : De la source à l'exploitation, en couvrant la chaîne d'approvisionnement (C1, C3).
Documenter le cycle de vie de la donnée : De sa création à son archivage ou suppression (C1, C8).
Comprendre les enjeux de la documentation des jeux de données pour répondre aux besoins métiers et aux cas d'usage (C1).
Identifier les risques éthiques et sociétaux liés à la collecte et à l'utilisation des jeux de données (C2).
Préparer et structurer les données afin de renforcer leur intégrité pour une utilisation dans un modèle d'IA (C3).
Élaborer une documentation technique claire pour assurer la transparence et la traçabilité des données (C8).

Public visé

Toute personne impliquée dans la gestion et la documentation des données.

Pré-requis

Notions de base en gestion de données et en traitement des données.
Compréhension des concepts fondamentaux liés aux chaînes d'approvisionnement des données.
Notions de base en gestion de données (CSV, Excel, SQL).
Compréhension de l'importance des données pour les cas d'usage IA.
Connaissance des concepts de base en intelligence artificielle

PROGRAMME

Documenter les Flux et le Cycle de Vie des Données – 4H00 – 2, 3
Documenter et Structurer les Jeux de Données pour l'IA – 4H00 – 1, 2, 3

CISIA-DOC01 : Documenter les Flux et le Cycle de Vie des Données – 4H00 – 2, 3

1. **Introduction à la documentation du flux de traitement des données (1 heure)**

- Présentation des concepts de flux de données et de chaîne d'approvisionnement (C1)
- Méthodes pour documenter chaque étape du flux de données, de la source à l'exploitation (C3)

2. Documentation du cycle de vie de la donnée (1 heure)

- Étapes du cycle de vie des données : collecte, stockage, traitement, utilisation, archivage (C1)
- Importance de la documentation pour chaque phase du cycle de vie (C8)

Informations non contractuelles - PDF généré le 29/11/2026

3. Considérations éthiques et réglementaires dans la documentation (1 heure)

- Identification des risques éthiques et sociétaux associés aux données (C2)
- Respect du cadre réglementaire lors de la documentation des données (C2)

4. Mise en pratique : Atelier de documentation (1 heure)

- Exercice pratique : Création d'une documentation pour un flux de traitement et un cycle de vie de données (C1, C8)
- Analyse et discussion des résultats (C8)

CISIA-DOC02 : Documenter et Structurer les Jeux de Données pour l'IA – 4H00 – 1, 2, 3

2. **Introduction à la documentation des jeux de données (1 heure)**

- Importance de la documentation (C1)
- Présentation des datasheets et de la documentation technique (C1, C8)

2. Structure et intégrité des données (1 heure)

- Techniques de structuration et préparation des données (C3)
- Vérification de la cohérence et de la pertinence des données pour un cas d'usage IA (C1, C3)

3. Considérations éthiques et réglementaires dans la documentation (1 heure)

- Identification des risques éthiques et sociétaux (C2)
- Cadre réglementaire et législation sur la confidentialité des données (C2)

4. Mise en pratique : Élaboration d'une documentation technique (1 heure)

- Atelier pratique : création d'une datasheet pour un jeu de données (C1, C8)
- Documentation des étapes de traitement de la donnée (C8)

MODALITÉS

Modalités

L'ensemble du parcours est accessible en présentiel, à distance ou mode hybride.

Présentation théorique en présentiel (4 heures).
Atelier pratique avec exercices en ligne et en présentiel (4 heures).

Méthode

Élaboration d'un document complet pour le flux de traitement et le cycle de vie d'un jeu de données
Élaboration et présentation d'une datasheet pour un jeu de données

Informations non contractuelles - PDF généré le 29/11/2026

CERTIFICATIONS

A l'issus du parcours (10 modules), les candidats pourrons passer le jury de certifcation ATLAS :
Concevoir et implémenter une solution d'IA

Informations non contractuelles - PDF généré le 29/11/2026