



L'IA éthique - Enjeux, risques et bonnes pratiques

IA003

Durée: 1 jour

Public :

Professionnels utilisant l'IA ou implémentant des solutions à base d'IA : chefs de projet, développeurs, data scientists

Objectifs :

Identifier les enjeux éthiques de l'IA dans son contexte professionnel, analyser les sources de biais et leurs impacts, appliquer les principes de transparence et d'explicabilité, mettre en œuvre les bonnes pratiques de gouvernance et anticiper les évolutions réglementaires

Connaissances préalables nécessaires :

Aucune connaissance technique préalable, notions sur les enjeux du numérique

Programme :

Introduction aux principes éthiques de l'IA

Valeurs fondamentales et cadre réglementaire
Principes éthiques européens de l'IA
Panorama des cadres réglementaires nationaux et internationaux
Responsabilité algorithmique

Atelier théâtral : "Le tribunal des algorithmes" (procès fictif où différents algorithmes sont jugés selon leurs impacts sociaux)

Biais et discrimination

Sources et types de biais : historiques, de représentation, d'évaluation, de confirmation
Sources multiples : données d'entraînement, conception, déploiement
Impact critique sur les services publics, RH, santé, justice
Métriques d'équité et techniques de mitigation

Atelier interactif collaboratif : "La machine à préjugés" - Jeu de rôle où les équipes conçoivent intentionnellement un système de recommandation biaisé, puis appliquent des techniques de "debugging éthique" pour identifier et corriger les discriminations



— Phirio —

Transparence, explicabilité et droit à l'explication

Différence entre interprétabilité et explicabilité des modèles

Techniques d'XAI (eXplainable AI) : LIME, SHAP, arbres de décision, règles

Obligation d'explication selon le RGPD article 22

Communication adaptée selon les profils d'usagers

Atelier théâtral : "L'entretien impossible" - Simulation d'explications à donner à différents profils d'usagers (senior, adolescent, expert technique) sur des décisions algorithmiques complexes (refus de prêt, diagnostic médical, recommandation)

Protection des données et vie privée

Impact du RGPD sur les projets d'IA et contraintes spécifiques

Techniques d'anonymisation, pseudonymisation et préservation de la confidentialité

Minimisation des données et finalité déterminée

Consentement éclairé et droit de retrait dans l'ère de l'IA

Privacy by design et security by design

Jeu de rôle décisionnel : "Les gardiens des données" - Simulation d'un comité d'éthique devant arbitrer des cas d'usage controversés : surveillance urbaine intelligente, IA médicale prédictive, hyperpersonnalisation publicitaire, score social

Gouvernance, bonnes pratiques et perspectives réglementaires

Structures de gouvernance de l'IA : comités d'éthique, audits algorithmiques

Présentation détaillée de l'AI Act européen : obligations, calendrier, sanctions

Rôle et missions du Délégué à l'Ethique Numérique (DEN) / Data Ethics Officer (DEO)

Outils opérationnels : grilles d'évaluation éthique, check-lists projet, matrice de risques

Evolutions prospectives et standards internationaux émergents

Atelier : élaboration d'un plan d'action personnalisé pour son organisation