

Formation Déjouer les biais de l'IA par la pensée critique

Présentation

L'Intelligence Artificielle Générative (IAG) ouvre de nouvelles perspectives, mais elle n'est pas exempte de biais cognitifs et algorithmiques pouvant influencer les décisions.

Cette formation propose aux participants de développer une pensée critique appliquée à l'IA, afin d'évaluer ses productions avec discernement, d'identifier les biais et de mettre en place des pratiques d'utilisation plus fiables, éthiques et stratégiques.

Durée : 7,00 heures (1 jours)

Tarif INTRA : Nous consulter

Objectifs de la formation

A l'issue de la formation, le stagiaire sera capable de renforcer son esprit critique lors de l'utilisation de l'Intelligence Artificielle (IA).

- Identifier les fondements de l'esprit critique
- Comprendre les biais cognitifs et algorithmiques
- Adopter une posture critique, éthique et stratégique face à l'Intelligence Artificielle Générative (IAG)

Prérequis

Aucun prérequis

Public

Toute personne souhaitant renforcer son esprit critique face à l'IA.

Programme de la formation

Module 1. Introduction : Pourquoi développer l'esprit critique face à l'IA

Objectif spécifique : Comprendre l'importance de la pensée critique dans l'usage professionnel de l'IA.



Contenu

- Contexte : la place croissante de l'IA dans les décisions professionnelles et personnelles
- Risques liés à la confiance excessive dans les résultats IA (illusion de fiabilité, absence de remise en question)
- Importance de développer une posture active de questionnement
- Différence entre esprit critique et méfiance systématique

Exemple d'atelier : évaluation rapide de plusieurs réponses produites par IA (texte, image, résumé) pour identifier intuitivement leurs limites, erreurs ou incohérences

Module 2. Fondements de l'esprit critique appliqué à l'IA

Objectif spécifique : Connaître les principes et outils permettant de questionner et d'évaluer les informations produites par l'IA.

Contenu

- Définition et composantes de l'esprit critique : observation, questionnement, validation, décision
- Méthodes pour interroger une information : qui l'a produite ? comment ? dans quel but ?
- Techniques de vérification : recoupement des sources, confrontation à des données fiables, utilisation d'outils de fact-checking
- Limites humaines : surcharge informationnelle, heuristiques mentales

Exemple d'atelier : analyse critique d'un rapport généré par IA – identification des affirmations discutables, des manques d'éléments factuels et propositions de reformulation

Module 3. Comprendre les biais cognitifs

Objectif spécifique : Identifier les biais mentaux qui influencent l'interprétation des résultats IA.

Contenu

- Définition d'un biais cognitif et mécanismes psychologiques sous-jacents
- Les biais les plus fréquents face à l'IA :
 - Biais de confirmation : tendance à chercher des éléments qui confirment nos idées préconçues
 - Biais de disponibilité : se baser sur les informations les plus faciles à rappeler

- Biais d'ancrage : influence excessive d'une première information reçue
- Impact des biais cognitifs sur la prise de décision et l'évaluation des résultats IA
- Exemples issus de contextes professionnels : recrutement, gestion de projet, analyse de données

Exemple d'atelier : étude de scénarios concrets où un biais cognitif a conduit à une mauvaise décision, identification du biais et reformulation de l'analyse pour corriger la conclusion

Module 4. Identifier et analyser les biais algorithmiques

Objectif spécifique : Reconnaître les biais présents dans les modèles d'IA et comprendre leur origine.

Contenu

- Définition du biais algorithmique et distinction avec les biais cognitifs
- Origines possibles :
 - Biais dans les données d'entraînement : données incomplètes ou déséquilibrées
 - Biais liés à l'architecture du modèle : paramètres influençant certains résultats
 - Biais d'utilisation : mauvaise adaptation au contexte métier
- Conséquences concrètes : discriminations involontaires, résultats inexacts, perte de confiance
- Études de cas inspirées de situations réelles :
 - RH : algorithmes de tri de CV défavorisant certains profils
 - Finance : notation de crédit influencée par des données historiques biaisées
 - Santé : diagnostic moins précis pour certaines populations

Exemple d'atelier :

- Présentation rapide d'un exemple concret par le formateur
- Travail en petits groupes :
 1. Identifier le ou les biais présents
 2. Rechercher les causes possibles
 3. Proposer 2 à 3 actions correctives simples (ex. améliorer la représentativité des données, mettre en place une validation humaine, rendre l'algorithme plus transparent)

Restitution synthétique en plénière

E-learning Découverte de l'IA générative mis à disposition post-formation et consultable pendant 1 mois pour renforcer ou réviser ses connaissances (e-learning de 30 minutes)

Organisation

Formateur

Les formateurs de Docaposte Institute sont des experts de leur domaine, disposant d'une expérience terrain qu'ils enrichissent continuellement. Leurs connaissances techniques et pédagogiques sont rigoureusement validées en amont par nos référents internes. Riches de leur expérience sur le sujet, ils sauront accompagner vos collaborateurs dans leur montée en compétences.

Moyens pédagogiques et techniques

- Alternance d'apports théoriques et de mises en pratique contextualisées.
- Progression pédagogique structurée, de la compréhension à la mise en œuvre opérationnelle.
- Pédagogie immersive favorisant une mise en application immédiate des compétences.
- Formation interactive et participative s'appuyant sur des ateliers pratiques, études de cas et échanges collectifs (+ 60 % du temps de formation.)
- Séquences pédagogiques rythmées et évaluations formatives régulières pour renforcer l'ancrage des apprentissages.
- Supports de cours, ressources documentaires et références mis à disposition par le formateur

Dispositif de suivi de l'exécution et de l'évaluation des résultats de la formation

En amont de la formation (*évaluation en ligne dans l'extranet apprenant*)

- Recueil des besoins et des attentes des participants.
- Auto-positionnement permettant d'évaluer le niveau initial des apprenants.

Tout au long de la formation

- Évaluations continues réalisées par le formateur à travers des questions orales, exercices, QCM, études de cas et mises en situation pratiques.

À l'issue de la formation (*évaluations en ligne dans l'extranet apprenant*)

- Auto-positionnement afin de mesurer l'acquisition des compétences visées.
- Validation de l'acquisition des compétences par le formateur
- Questionnaire de satisfaction « à chaud » à l'issue de la formation
- Questionnaire de satisfaction "à froid", 45 jours après la formation afin d'évaluer l'ancrage des acquis et leur mise en application en situation professionnelle.

Accessibilité

Nos formations peuvent être adaptées à certaines conditions de handicap. Nous contacter pour toute information et demande spécifique.