



Principes éthiques pour la recherche sur l'AI

Version 08.2019

1. Principes éthiques pour la recherche sur l'intelligence artificielle

- Développer, déployer et utiliser les systèmes d'intelligence artificielle (IA) en respectant les principes éthiques du respect de l'autonomie humaine, de la prévention des dommages, de l'équité et de l'explicabilité/clarté. Reconnaître et corriger les éventuelles tensions potentielles entre ces principes.
- Accorder une attention particulière aux situations impliquant des groupes plus vulnérables tels que les enfants, les personnes handicapées et autres personnes traditionnellement défavorisées ou menacées d'exclusion, et aux situations caractérisées par des asymétries de pouvoir ou d'informations, telles qu'entre employeurs et travailleurs, ou entre les entreprises et les consommateurs.
- Reconnaître que, tout en apportant des avantages substantiels aux individus et à la société, les systèmes d'IA posent également certains risques et peuvent avoir un impact négatif, notamment des impacts difficiles à anticiper, identifier ou mesurer (par exemple sur la démocratie, l'état de droit et la justice distributive, ou sur l'esprit humain lui-même.) Adopter des mesures appropriées pour atténuer ces risques, le cas échéant, et proportionnellement à l'ampleur du risque.

2. Exigences de l'IA :

- S'assurer que le développement, le déploiement et l'utilisation de systèmes d'intelligence artificielle répondent aux sept exigences clés de l'IA fiable : (1) gestion et supervision humaines, (2) robustesse technique et sécuritaire, (3) confidentialité et gouvernance des données, (4) transparence, (5) diversité, non-discrimination et équité, (6) bien-être environnemental et sociétal et (7) responsabilité.
- Envisager des méthodes techniques et non techniques pour assurer la mise en œuvre de ces exigences.
- Encourager la recherche et l'innovation pour aider à évaluer les systèmes d'IA et à poursuivre la réalisation des exigences; diffuser les résultats et les questions ouvertes au grand public et former systématiquement une nouvelle génération d'experts en éthique de l'IA.
- Communiquer de manière claire et proactive aux parties prenantes des informations sur les capacités et les limites du système d'IA, permettant ainsi de définir des attentes réalistes, et sur la manière dont les exigences sont mises en œuvre. Soyez transparent sur le fait qu'ils ont affaire à un système d'IA.
- Faciliter la traçabilité et l'auditabilité des systèmes d'IA, en particulier dans des contextes ou des situations critiques.

- Impliquez les parties prenantes tout au long du cycle de vie du système d'IA. Encourager la formation et l'éducation afin que toutes les parties prenantes soient sensibilisées et formées à l'IA de confiance.
- N'oubliez pas qu'il peut exister des tensions fondamentales entre différents principes et exigences. Identifiez, évaluez, documentez et communiquez en permanence ces compromis et leurs solutions.

Source / plus d'information :

[Guide d'éthique pour une intelligence artificielle digne de confiance](#) par le groupe d'experts de la Commission européenne [AI HLEG](#).