

A QUEL POINT

L'INTELLIGENCE ARTIFICIELLE

EST-ELLE DIGNE DE CONFIANCE?

Que signifie une IA "digne de confiance"? Pour vous aider à répondre à cette question, certaines lignes directrices pour le développement et l'utilisation de l'IA existent. Ces lignes directrices se concentrent sur plusieurs questions éthiques et sociales liées à l'IA.

Le Knowledge Centre Data & Society résume les [7 exigences pour l'IA](#), rédigées par le « Groupe d'Experts de Haut Niveau sur l'IA » (AI HLEG) créé par la Commission européenne. Pour chaque exigence, un certain nombre de questions sont formulées et vous permettent de réfléchir à la fiabilité de votre système d'IA.

Si vous souhaitez une approche plus détaillée, le Knowledge Centre Data & Society a récemment développé [les cartes "AI Blindspots"](#). En utilisant cet outil, vous réfléchissez de manière proactive aux décisions et actions (éthiques et sociales) que vous souhaitez prendre au début

de votre projet ou pendant le développement de votre système d'IA.

Avertissement : les 7 exigences se chevauchent, nous avons donc réduit les questions redondantes et choisi les questions les plus importantes afin d'en réduire la complexité.

Knowledge Centre Data & Society (2020). How ethical is AI? brAlnfood from the Knowledge Centre Data & Society. Brussels: Knowledge Centre Data & Society

This document is available under a CC BY 4.0 license.

brAlnfood est un projet développé conjointement par D&M, CLAIRE et Inria, encourageant une vision européenne commune de l'IA pour le bien de la société et au service de tous.

#AI4Good #AI4All

INTERVENTION HUMAINE & SURVEILLANCE

- L'interaction entre votre système d'IA et un humain est-elle significative et pertinente?
- Qui prend la décision finale, la machine ou l'humain? Si c'est la machine, alors il n'y a pas de surveillance humaine.
- Votre système d'IA est-il autonome ou auto-apprenant? Si oui, existe-t-il des mécanismes de contrôle?

ROBUSTESSE TECHNIQUE ET SÉCURITÉ

- Quelles mesures allez-vous prendre si votre système d'IA est attaqué ou se comporte différemment de ce qui est prévu, ou s'il est utilisé à d'autres fins (non souhaitées)?
- Quelles sont les mesures mises en place pour remédier à l'inexactitude?

VIE PRIVÉE ET GOUVERNANCE DES DONNÉES

- Votre système d'IA utilise-t-il des données personnelles? Si oui, êtes-vous conscient des implications et des exigences liées à l'utilisation de données personnelles? Pouvez-vous prouver que vous respectez la réglementation en matière de protection des données?
- Disposez-vous de mécanismes de contrôle pour la collecte, le stockage, le traitement et l'utilisation des données?
- Qui peut accéder aux données des utilisateurs? Disposez-vous de mesures de protection de l'accès aux données?

TRANSPARENCE

- Pouvez-vous montrer comment l'algorithme est conçu et construit, et comment les décisions sont prises, à votre équipe mais aussi à toute personne en contact avec le système? Assurez-vous que tout le monde comprend bien qu'une composante de l'IA fait partie du système.

DIVERSITÉ, NON-DISCRIMINATION ET ÉQUITÉ

- Comment allez-vous éviter les préjugés injustes dans votre système d'IA? Est-il possible pour d'autres personnes de signaler un parti pris ou une discrimination?
- Avez-vous réfléchi à la manière dont votre innovation en matière d'IA modifie l'accès à votre service pour les groupes marginalisés? Si oui, proposez-vous un service alternatif pour les personnes exclues?
- Vos parties prenantes peuvent-elles participer au développement et à l'utilisation de votre système d'IA?

BIEN-ÊTRE SOCIÉTAL ET ENVIRONNEMENTAL

- Pouvez-vous mesurer l'impact environnemental du cycle de vie de votre système d'IA? Savez-vous comment le réduire?
- Les utilisateurs (finaux) sont-ils conscients des limites de l'interaction sociale avec votre système d'IA et de l'impact sociétal de votre système d'IA? D'autres groupes ou individus pourraient-ils être indirectement affectés par votre système d'IA, au-delà de votre public cible?

RESPONSABILITÉ

- Avez-vous une vue d'ensemble de toutes les décisions et de tous les compromis que vous avez faits pour créer votre système?
- Avez-vous identifié les conséquences négatives pour toutes les parties concernées?
- Avez-vous été en mesure de réduire les conséquences négatives, et ces mesures sont-elles documentées?

brAlnfood of the Knowledge Centre Data & Society



CLAIRE

Inria