

Développement et usage responsable de l'Intelligence Artificielle

Éléments méthodologiques et perspectives européennes

Raja Chatila

Institut des Systèmes Intelligents et de Robotique (ISIR)

Sorbonne Université, Paris, France

Raja.Chatila@sorbonne-universite.fr

Applications de l'IA et de la Robotique en Santé

- Stockage, traitement et analyse des données médicales
- Aide au diagnostic
- Médecine PPPP
- Télémédecine
- Aide à l'autonomie et assistance aux personnes âgées, handicapées
- Chirurgie
- Réparation, Réadaptation, ... Augmentation
- Modèles épidémiologiques
- Conception de médicaments et de vaccins
- ...

Groupe d'experts de haut niveau sur l'IA (Commission Européenne)

Orientations éthiques pour une IA de confiance (2019)



Une IA centrée sur l'être humain et non sur le seul développement technologique

- **IA licite**: conforme au droit
- **IA Ethique**: alignement avec les principes et normes éthiques.
- **IA Robuste**: technique: sûreté de fonctionnement, sécurité, fiabilité; et non-technique: contexte d'application approprié, limitation des conséquences non voulues...

Exigences pour une IA de confiance



1. **Autonomie humaine et contrôle humain**
2. **Robustesse technique et sécurité**
3. **Respect de la vie privée et gouvernance des données**
4. **Transparence:** traçabilité, explicabilité, communication
5. **Diversité, non-discrimination et équité:** absence de biais injustes, accessibilité, conception universelle, participation des parties prenantes
6. **Bien-être sociétal et environnemental:** durabilité, impact social
7. **Responsabilité :** auditabilité, réduction des incidences négatives, recours.

Réalisation de l'IA de confiance: Aspects techniques



- Architectures pour l'IA de confiance
- Ethique et conformité par conception (X-by-design)
- Méthodes d'explication
- Tests et validation
- Indicateurs de qualité de service

Réalisation de l'IA de confiance: Aspects non techniques



- Réglementation
- Codes de conduite
- Standards et normalisation
- Certification
- Responsabilité et structures de gouvernance
- Education et sensibilisation à un état d'esprit éthique
- Participation des parties prenantes et dialogue social
- Diversité and inclusivité des équipes de conception

Exemple: approche de conception orientée valeurs

- **Initialisation du projet:** Analyse des besoins techniques, analyse normative, initialisation d'une réflexion éthique.
- **Identification des parties prenantes** (personnes concernées, entourage, soignants, ...) et découverte de leurs valeurs en lien avec les **7 exigences** (p. ex: bien-être, intimité, autonomie, relations sociales, transparence, données personnelles, accessibilité, équité, ...)
- Analyse de l'**impact du projet sur les valeurs** (bénéfices, risques) et identification des tensions entre elles.
- **Identification des priorités des valeurs** avec les parties prenantes
- **Choix techniques en conformité avec les priorités des valeurs** pendant les étapes du projet (conception, développement, déploiement)