

L'encadrement éthique de l'IA

Jérôme BÉRANGER

Bienvenue dans cette séquence sur l'encadrement éthique de l'IA.

Afin d'alimenter mes propos, je répondrai à sept questions majeures qui sont associées à cette thématique.

Qu'est-ce que l'éthique ?

L'éthique vient du grec *ethos*, qui signifie mœurs et habitudes. Elle désigne une réflexion sur les valeurs qui sous-tendent nos actions et qui permettent de rechercher la bonne décision dans un contexte donné.

Elle repose sur quatre fonctions principales :

- Déterminer ce qui est moral
- Comprendre pourquoi il est important de vivre moralement
- Donner du sens à nos actions
- Appliquer ces principes à la vie quotidienne.

L'éthique peut s'appuyer sur quatre principes universels :

- La bienfaisance : c'est le bienfondé d'une action.
- L'autonomie : c'est la réalisation d'une action sans influence extérieure, avec un consentement libre et éclairé.
- La non-malfaisance : c'est quand les conséquences d'une action ne nuisent pas à la personne.
- La justice : c'est une action basée sur l'égalité et/ou l'équité au niveau sociétal.

L'éthique est donc une boussole qui guide et donne du sens à nos actions, à nos comportements pour garantir une vie bonne, à la fois pour soi mais également pour les autres.

Dans quels domaines y a-t-il des enjeux, risques et limites par rapport à l'éthique de l'IA ?

L'IA pose des défis majeurs dans plusieurs domaines :

- La transparence : les décisions algorithmiques sont parfois opaques (black box), ce qui limite la compréhension et le contrôle par les utilisateurs.
- La responsabilité : qui est responsable lorsque l'IA prend une mauvaise décision ? Le concepteur, l'utilisateur ou l'algorithme lui-même ?
- La justice sociale : l'IA peut reproduire ou amplifier des discriminations en raison de biais dans les données d'entraînement.
- L'environnement : l'entraînement des modèles d'IA consomme énormément d'énergie et pose des questions de soutenabilité voire d'éco-responsabilité.
- L'autonomie humaine : l'IA peut restreindre notre libre arbitre en influençant nos décisions sans que nous en soyons pleinement conscients.
- La confidentialité et la vie privée : les données d'entrée qui alimentent l'IA peuvent faire l'objet de divulgation ou captation extérieure à des fins illicites voire malveillantes.

Quelle est la ligne de partage entre éthique et encadrement légal ?

L'encadrement légal impose des règles minimales à respecter, alors que l'éthique va plus loin en intégrant des valeurs et des principes. Légalement, on respecte la loi ; éthiquement, on vise le meilleur choix possible. Un usage peut être légal sans être éthique. Par exemple, un algorithme de recrutement basé sur l'historique des embauches passées peut être légal, mais s'il exclut systématiquement certains profils (ex. : minorités, femmes), il devient contraire aux principes éthiques de justice et d'équité.

Un cas d'usage de l'IA peut-il être légal et en même temps non éthique ?

Oui. Un exemple serait un système de notation de crédit basé sur des critères prédictifs. Si ce système évalue la solvabilité d'une personne en fonction de son lieu de résidence ou de son réseau social, cela peut être légal, mais discriminatoire et donc contraire à l'éthique. Il ne faut pas se limiter à la seule conformité légale, mais d'intégrer une vigilance éthique tout au long du cycle de vie du système d'IA.

Un agent territorial qui utilise une IA doit-il prévenir les administrés ?

Oui. La transparence est un principe fondamental de l'éthique du numérique. Il est essentiel d'avoir une communication claire sur les usages de l'IA, idéalement sous trois niveaux :

- Déclaration d'intention : informer de manière sincère et accessible sur le rôle de l'IA utilisée.
- Explication des mécanismes : détailler les principes de fonctionnement si cela impacte directement les citoyens.
- Identification des responsables : assurer un contact humain pour répondre aux questions.

Les collectivités territoriales doivent-elles mettre en place des comités d'éthique et des chartes pour encadrer l'IA ?

Oui. Cela est fortement recommandé.

Un comité d'éthique peut :

- Identifier les valeurs humaines à préserver face aux usages de l'IA
- Vérifier l'impact des algorithmes sur la société locale
- Sensibiliser les agents publics et les citoyens aux enjeux éthiques

Il est fondamental d'avoir une gouvernance éthique et adopter une approche souple et évolutive (Ethics by Evolution), qui permet d'adapter les principes éthiques en fonction des retours et des évolutions technologiques. Une charte éthique permet de constituer un socle moral composé de principes et d'engagements éthiques sur lequel se fonde la gouvernance éthique.

Comment se servir de l'éthique pour faire un usage de l'IA responsable et compatible avec le bien commun ?

L'éthique doit être intégrée dès la conception des systèmes d'IA (Ethics by Design) et évoluer avec leur usage (Ethics by Evolution).

Cela passe par :

- L'inclusion des parties prenantes : impliquer citoyens, chercheurs et experts dès le début.
- L'explicabilité des décisions algorithmiques : rendre les processus compréhensibles et auditables.
- L'évaluation continue des impacts : ajuster les usages de l'IA en fonction des retours et des risques identifiés.

- L'intégration de garde-fous éthiques : éviter la dépendance excessive aux décisions automatisées.

L'IA n'est pas qu'un simple outil technologique, c'est une transformation sociétale qui doit être encadrée par des principes éthiques solides. L'éthique permet d'aller au-delà du respect de la loi et d'assurer que les usages de l'IA soient justes, sécurisés, transparents, bénéfiques, et éco-responsables pour le bien commun.