

Éthique de la recherche en robotique

GdR Robotique – Action Transverse « Robotique et Société »
8 décembre 2022

Raja Chatila, Catherine Tessier

Le programme

9h30-12h30 : Outils pour une délibération éthique

- Vocabulaire
- Cadres de philosophie morale (présentés sur une expérience de pensée)
- Principes européens relatifs à « l'éthique de l'intelligence artificielle »
- Exercice en groupes sur projets réels, restitution
- Discussions relatives à vos questions sur l'éthique de la recherche en robotique

Repas

13h30-16h00 : Expérimenter mon robot avec des participants

- Repères historiques
- Aspects réglementaires et éthiques de l'expérimentation avec des participants
- Exercice en groupes sur projets réels, restitution
- Discussion libre, en déambulation au Jardin des Plantes si la météo le permet

Le vocabulaire

- Valeur
- Morale
- Éthique
- Pourquoi parler d'éthique ?

Le vocabulaire 1/3

Valeur (morale)

Ce qui est posé comme vrai, beau, bien, d'un point de vue personnel ou selon les critères d'une société et qui est donné comme un idéal à atteindre, comme quelque chose à défendre [Larousse en ligne]

- origine : la société, soi-même
- une valeur peut être trahie
- exemples : liberté, honnêteté, bienveillance...



Image Gerd Altmann - Pixabay

Morale

Ensemble des règles, des principes selon lesquels on dirige sa vie, sa conduite, ses mœurs, considéré relativement au bien et au mal [dictionnaire de l'Académie française, sens 1]

- origine : la société (consensus social), la religion, soi-même
- normative
- exemples : ne pas mentir, ne pas voler...



Image Tumisu - Pixabay

Le vocabulaire 2/3

Éthique (nom)

Réflexion [personnelle ou collective] relative aux conduites humaines et aux valeurs qui les fondent

[Dictionnaire de l'Académie Française, sens 1]

Éthique (adjectif) : qui a rapport aux conduites humaines et aux valeurs qui les fondent. *Les valeurs éthiques. Les préceptes, les jugements éthiques.* [Dictionnaire de l'Académie Française]

Démarche en vue d'agir **au mieux**, de déterminer, de manière **argumentée**, ce qui peut être considéré comme la décision ou l'action **la plus juste** au sens d'un positionnement sur des **échelles de valeurs**.

→ **se poser des questions chemin faisant**, réfléchir, décider au cas par cas [CERNA2018]

- l'éthique recommande
- ce n'est pas le fruit d'une règle
- c'est la tension entre valeurs qui fonde la discussion éthique



[CERNA18] Commission de réflexion sur l'éthique de la recherche en sciences et technologie du numérique d'Allistene - Proposition de formation doctorale, Initiation à l'éthique de la recherche scientifique, octobre 2018

Le vocabulaire 3/3 – des contresens

→ Un objet (machine, logiciel), [une personne], ne peut pas être « éthique »

un robot qui a rapport aux conduites humaines et aux valeurs qui les fondent ???

→ Mauvaise traduction de *ethical* [MWD]

1: of or relating to ethics - *ethical theories*. 2 : involving or expressing moral approval or disapproval - *ethical judgments*. 3 : conforming to accepted standards of conduct - *ethical behavior*

La conformité d'un objet à des standards (législation, normes, etc.) ne relève pas de l'éthique
(si ce n'est de l'éthique - sens français 2. - professionnelle, de l'ingénieur, du concepteur)

L'éthique - sens français 1 - est une **réflexion**, continue, toujours en mouvement
et a à faire avec des **dilemmes** (⇒ décision où tout n'est pas « conforme »)

Éthique : méthode, repères

« Une délibération collective, **toujours à reprendre**, doit faire la part entre le **possible** et l'**acceptable** [...] »

« [...] les connaissances nouvelles, en compliquant ou en déplaçant le champ de nos questionnements, augmentent souvent nos **incertitudes**. »

« La réflexion éthique oblige donc à **s'interroger sur les principes** qui définissent la notion même de progrès [...] »

« Chaque fois [que le progrès] est invoqué, nous nous devons de soumettre à une **analyse critique les certitudes et les convictions** qui le sous-tendent, en mettant en œuvre une évaluation scientifique précise et complète de l'existant.

Cette analyse critique doit s'appuyer, d'une part, sur des **valeurs**, et d'autre part, sur l'estimation la plus lucide possible des **bénéfices** et des **risques**. »

« Mais énumérer ces valeurs pour définir un « invariant éthique » ne suffit pas. [...] toute question particulière oblige à **les articuler entre elles** pour rechercher **le meilleur point d'équilibre possible** lorsqu'elles entrent en concurrence. »

[CCNE21] Comité consultatif national d'éthique pour les sciences de la vie et de la santé – *Rapport d'activité 2019-2020*. Août 2021. [lien](#) (c'est nous qui soulignons)

Pourquoi parler d'éthique ?

Parce que nous sommes des citoyens et citoyennes
et pas seulement des scientifiques

Parce que nous sommes (majoritairement) dans des institutions publiques au service de la société

Rôles :

- pour la science : construire des connaissances, les diffuser...
- pour la société : expliquer nos travaux, éclairer des débats...

La science n'est pas (n'est plus) coupée des mœurs, de la société

Scientifiquement possible / souhaitable pour la société

« *Tout ce qui est possible sera nécessairement réalisé* »

D. Gabor (1900-1979) - Prix Nobel de physique 1971

« *Tout ce qui devient techniquement possible ne doit pas nécessairement être permis* »

Conseil d'État - *Le numérique et les droits fondamentaux*. Étude annuelle 2014

- **Éthique de la recherche**
établir des connaissances, intérêt scientifique
- **Éthique des usages**
utiliser des connaissances, impact sur la société

Considérer ma recherche d'un point de vue éthique

C'est me poser des questions sur

- mon métier de **chercheur**
- mes **responsabilités** dans le cadre de ce métier
- le **sujet** de ma recherche
- l'**impact** de cette recherche sur la société

Les réponses à ces interrogations **peuvent varier** en fonction

- de mon **employeur** (académique / entreprise, domaine d'activités...)
- de ma **fonction** (doctorant, post-doctorant, enseignant-chercheur, chercheur, ingénieur...)
- du **pays** où j'exerce
- ...

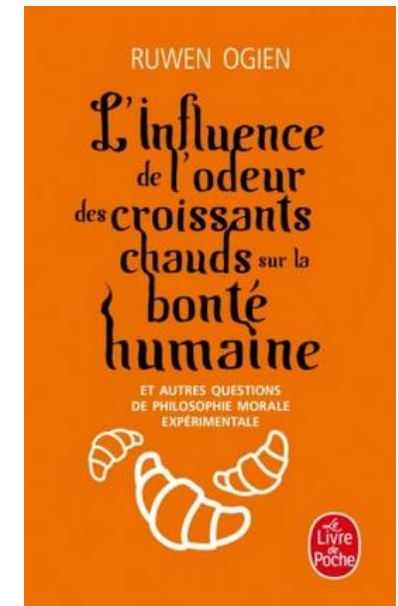
Vers la délibération éthique

- Expérience de pensée
- Illustration de cadres éthiques
- La délibération éthique

Un bon exercice : l'expérience de pensée

Par exemple les dilemmes dits « dilemmes moraux »

« donnent les moyens d'identifier plus clairement les facteurs qui influencent nos jugements moraux »



Qu'est-ce qu'un dilemme moral ?

Don Rodrigue, *Le Cid* (Corneille), acte I, scène 6

En cet affront **mon père** est l'offensé,
Et l'offenseur **le père de Chimène** !
Que je sens de rudes **combats** !
Contre mon propre **honneur** mon **amour** s'intéresse : [...]

Réduit au triste **choix** ou de **trahir** ma flamme,
Ou de vivre en **infâme**,
Des deux **côtés** mon mal est infini. [...]

Tous mes plaisirs sont morts, ou ma gloire ternie. [...]

Je dois à ma maîtresse **aussi bien** qu'à mon père ;
J'attire en me vengeant sa haine et sa colère ;
J'attire ses mépris en ne me vengeant pas. [...]

Je m'accuse déjà de trop de négligence ;
Courons à la vengeance ;
Et tout honteux d'avoir tant **balancé**,
Ne soyons plus en peine,
Puisqu'aujourd'hui mon père est l'offensé,
Si l'offenseur est père de Chimène.

→ Situation

→ Vertus, valeurs en tension

→ Trahison de ces valeurs

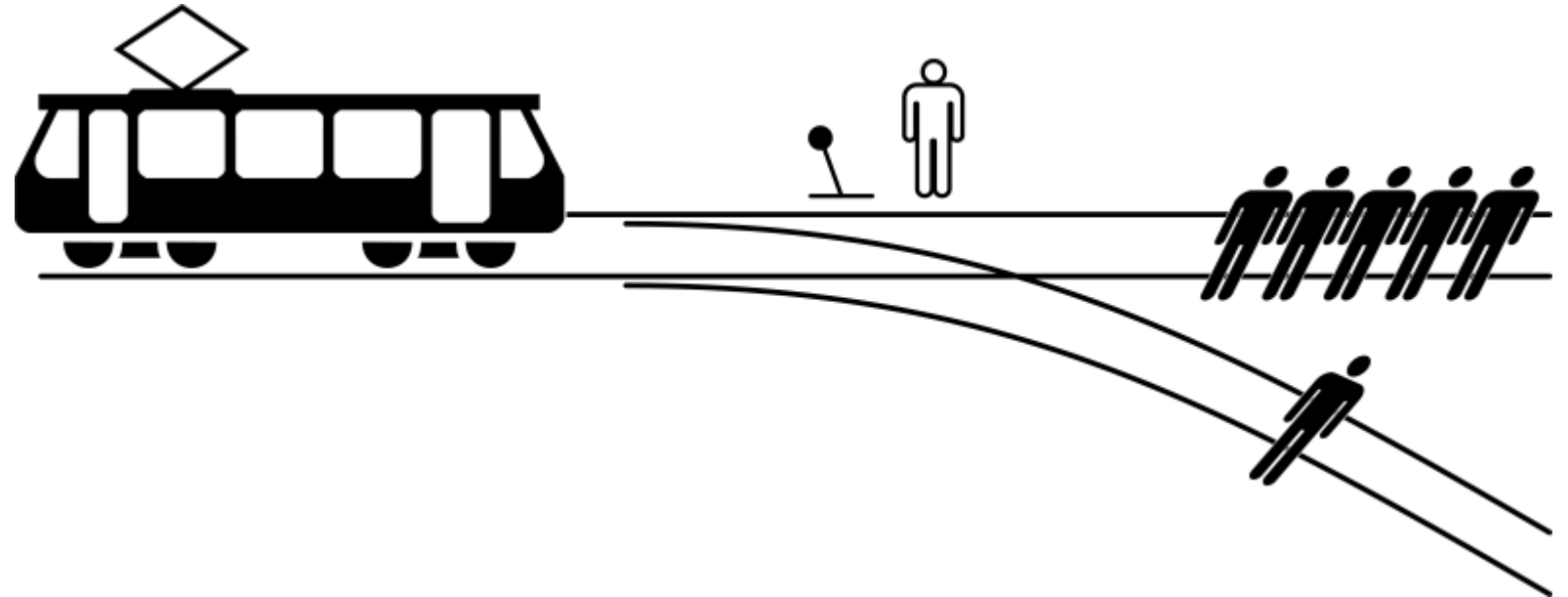
→ Conséquences négatives

→ Décision

Expérience de pensée – le dilemme du tramway [Foot02]

Que décidez-vous ?

- Je change l'aiguillage
- Je ne change pas l'aiguillage



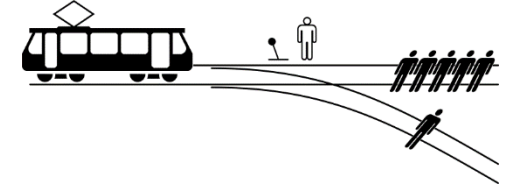
McGeddon, Wikipedia - CC BY-SA 4.0

[Foot02] Foot, Philippa. *Moral Dilemmas and Other Topics in Moral Philosophy*. Clarendon Press, Oxford University Press, 2002. DOI:10.1093/0199252866.003.0002

Tramway – Cadre conséquentialiste

Que décidez-vous ?

- Je change l'aiguillage
- Je ne change pas l'aiguillage



Action **juste** = action dont les **conséquences** sont « bonnes »

- conséquentialisme **positif** : maximiser le bien, le bonheur (5 vivants / 1 vivant)
- conséquentialisme **néгатif** : minimiser le mal, le malheur (1 mort / 5 morts)

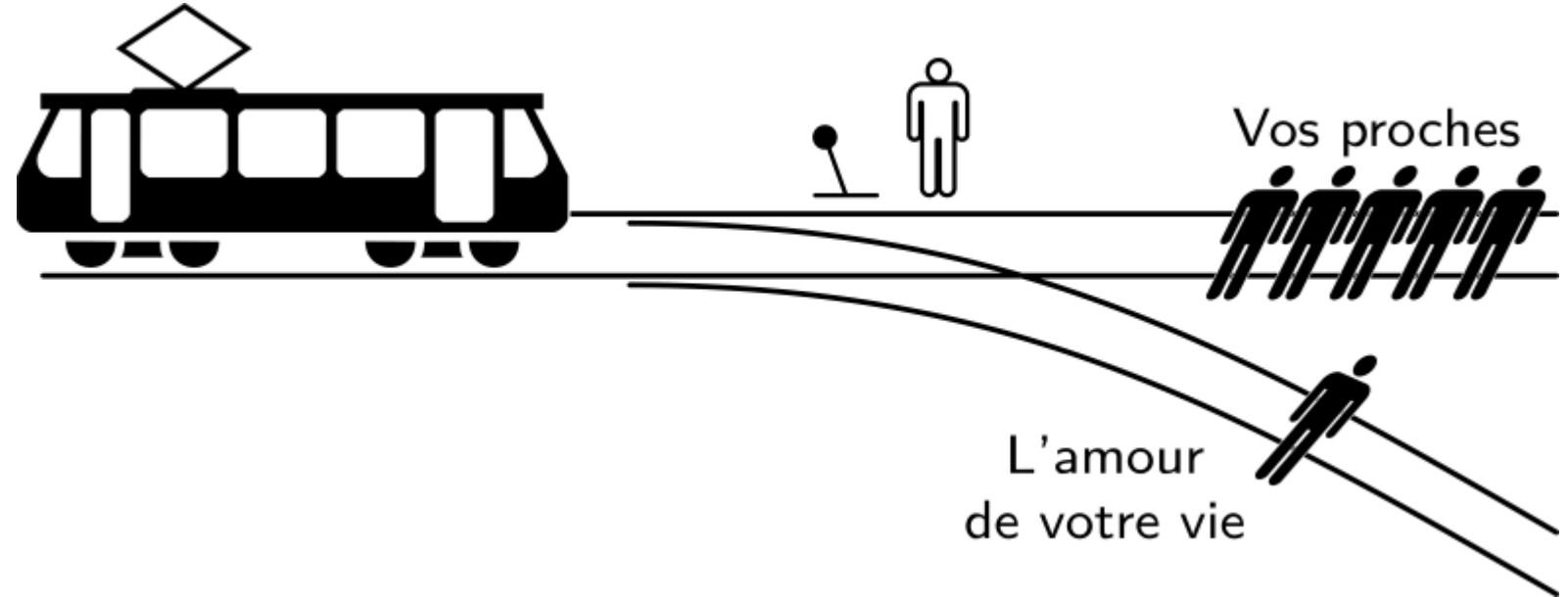
Exemples

- **Utilitarisme** : le plus grand bien pour le plus grand nombre
- **Égoïsme** : le plus grand bien pour soi-même
- **Principe de précaution** : minimisation des risques

Un autre dilemme du tramway

Que décidez-vous ?

- Je change l'aiguillage
- Je ne change pas l'aiguillage

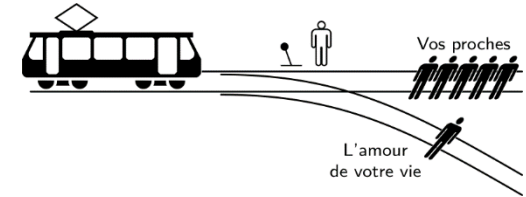


d'après McGeddon, Wikipedia - CC BY-SA 4.0

Un autre dilemme du tramway

Que décidez-vous ?

- Je change l'aiguillage
- Je ne change pas l'aiguillage



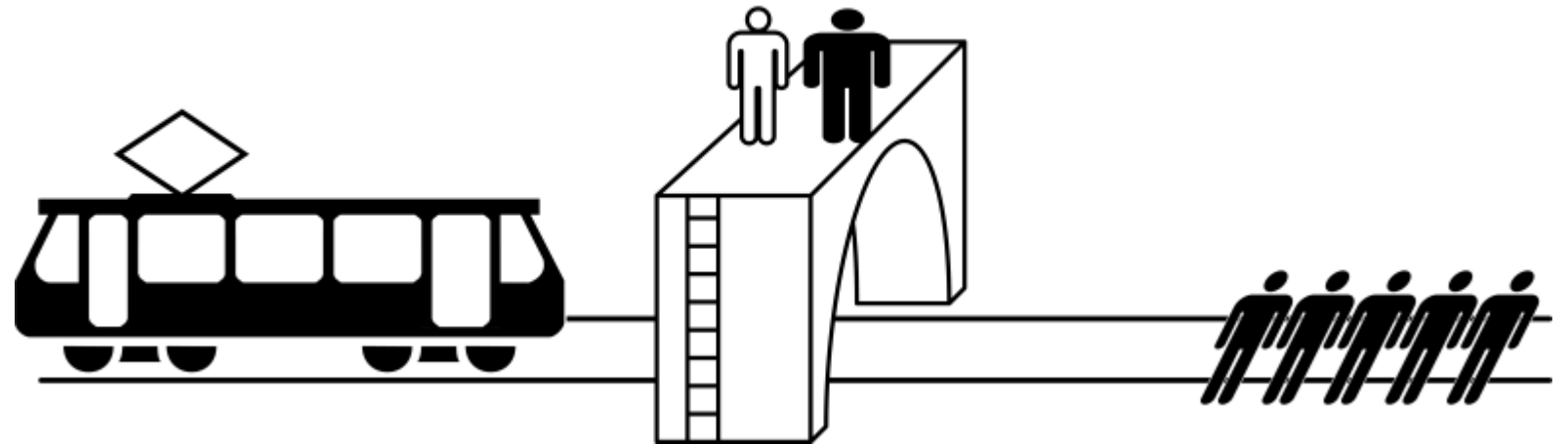
L'action **juste** n'est plus forcément celle qui correspond à un utilitarisme (numérique)
→ La réflexion éthique est **toujours en contexte**

Quelle **hiérarchie de valeurs** dans cette situation particulière ?

Expérience de pensée – le dilemme de la passerelle [Thomson76]

Que décidez-vous ?

- Je pousse mon voisin
- Je ne pousse pas mon voisin



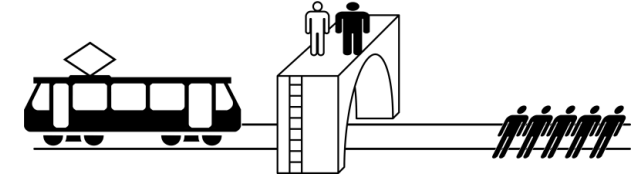
P.Gaborit, d'après McGeddon, Wikipedia - CC BY-SA 4.0

[Thomson76] Thomson, Judith Jarvis - *Killing, Letting Die, and the Trolley Problem*. *The Monist*, 59(2), 204–217, 1976. DOI:10.5840/monist197659224

Passerelle – Cadre déontologique

Que décidez-vous ?

- Je pousse mon voisin
- Je ne pousse pas mon voisin



Action **juste** = **action** qui respecte certains **principes** (au nom de quoi), et il faut suivre ces principes quelles que soient les conséquences.

Impératif catégorique (Kant)

« *Agis seulement d'après la maxime grâce à laquelle tu peux vouloir en même temps qu'elle devienne une loi universelle.* »

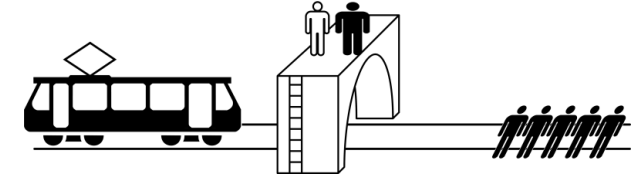
Exemples :

- toujours dire la vérité
- ne pas porter atteinte à l'intégrité d'autrui

Passerelle – Éthique des vertus

Que décidez-vous ?

- Je pousse mon voisin
- Je ne pousse pas mon voisin
- Je saute



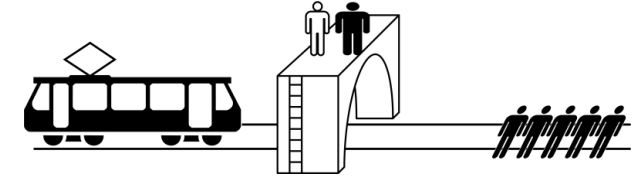
Action **juste** = l'**agent** qui la réalise met en évidence **une vertu particulière**

Exemples : honnêteté, courage, prudence, **sacrifice de soi...**

Passerelle – Doctrine du double effet

Que décidez-vous ?

- Je pousse mon voisin
- Je ne pousse pas mon voisin



Action **juste si**

- « bonne » (selon cadre déontologique) ET
- l'effet négatif n'est ni une fin ni un moyen (c'est un dommage collatéral /un effet secondaire) ET
- l'effet négatif est proportionnel à l'effet positif (proportionnalité)

Exemples concrets :

- médicament et effets secondaires
- lois de la guerre



Principes européens

relatifs à « l'éthique de l'intelligence artificielle »

Textes du HLEG IA - Europe



Seven requirements

- 1. Human agency and oversight*
- 2. Technical robustness and safety*
- 3. Privacy and data governance*
- 4. Transparency*
- 5. Diversity, non-discrimination and fairness*
- 6. Environmental and societal well-being*
- 7. Accountability*

→ *Ethics Self Assessment* pour les projets Horizon Europe

<https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>

<https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment>



Et les robots ?



Lorsque le système d'IA comporte un robot ou système conversationnel, les utilisateurs humains sont-ils informés du fait qu'ils interagissent avec un agent virtuel ?

Sont visés ici les systèmes d'IA qui communiquent et interagissent avec les êtres humains en simulant un comportement social dans les interactions entre robots et humains (IA embarquée) ou comme avatars dans la réalité virtuelle. Ce faisant, ces systèmes ont le potentiel de modifier nos pratiques socioculturelles et le tissu de notre vie sociale.



Le développement de robots humanoïdes doit par conséquent faire l'objet d'une évaluation éthique minutieuse.

L'auto-évaluation éthique

Table of contents

1. Human embryonic stem cells (hESCs) and human embryos (hEs) (HE, DEP, EU4H and EDF)	4
2. Humans (all EU Programmes)	8
3. Human cells or tissues (all EU Programmes)	15
4. Personal data (all EU Programmes)	20
5. Animals (all EU Programmes)	27
6. Non-EU countries (all EU Programmes)	31
7. Environment, health and safety (all EU Programmes)	35
8. Artificial intelligence (all EU Programmes)	39
9. Other ethics issues (all EU Programmes).....	46
10. Crosscutting issue: potential misuse of results (all EU Programmes)	48



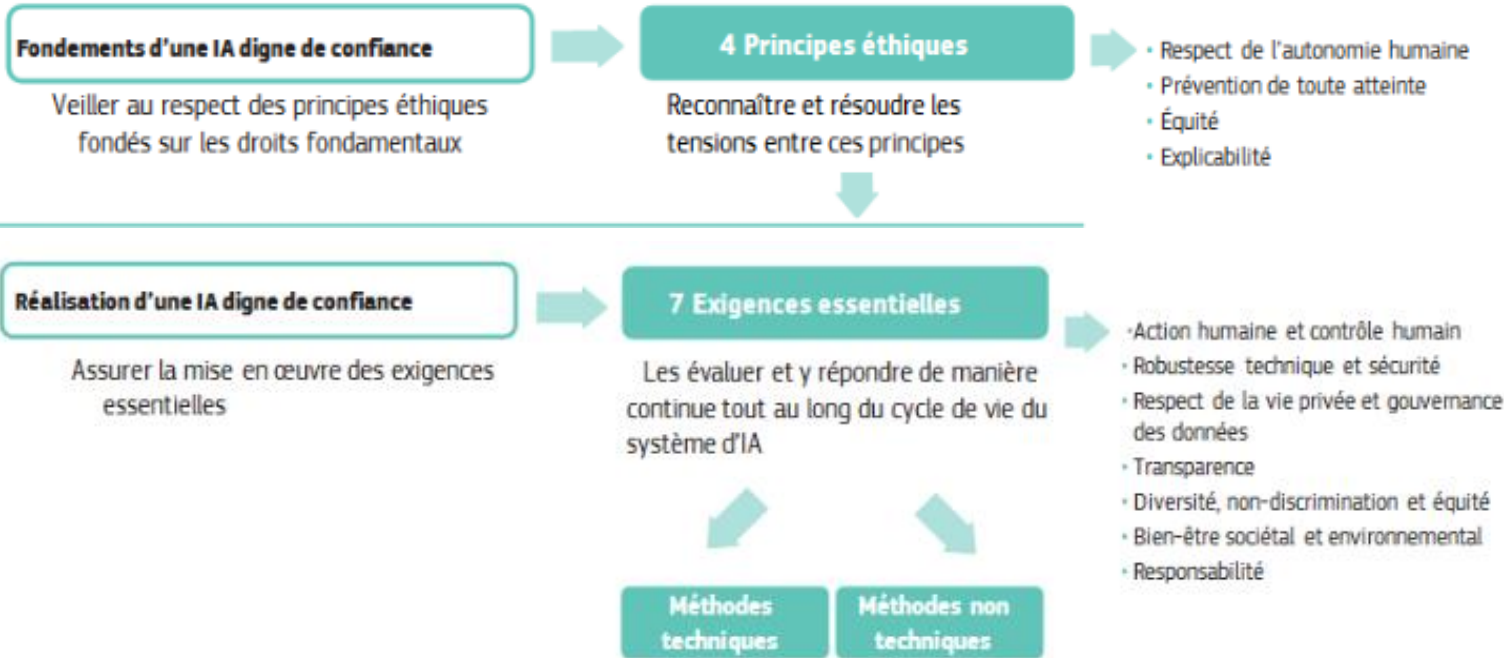
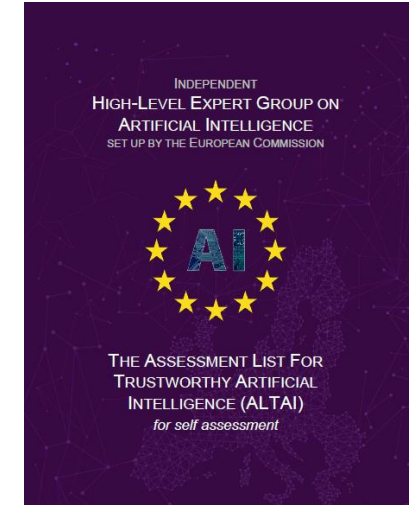
EU Grants

How to complete your ethics self-assessment

Version 2.0
13 July 2021

Chapitre 8 : intelligence artificielle (nouveau Horizon Europe)

<p>Does this activity involve the development, deployment and/or use of Artificial Intelligence-based systems?</p>	<input type="checkbox"/>	<input type="checkbox"/> <p>1) Explanation as to how the participants and/or end-users will be informed about:</p> <ul style="list-style-type: none"> - their interaction with an AI system/technology (if relevant); - the abilities, limitations, risks and benefits of the proposed AI system/technique; - the manner in which decisions are taken and the logic 	<p>1) Detailed risk assessment accompanied by a risk mitigation plan (if relevant). These must cover the development, deployment and post-deployment phases.</p> <p>2) Copies of ethics approvals (if relevant).</p>
--	--------------------------	--	--



Exercice

Deux situations – deux groupes
pour pratiquer la délibération éthique au sujet de projets en robotique

Chacun des deux sous-groupes étudie une situation

- Réflexion individuelle
→ faits, valeurs, principes, tensions, arguments selon différents cadres
- Mise en commun et discussion
- Proposition d'une argumentation étayée
- Présentation à l'autre groupe

La réflexion éthique

« J'ai coutume à rappeler que la réflexion éthique ne peut se complaire seulement dans le monde tel que l'on aimerait qu'il fût au risque d'être insignifiante. Elle doit pour acquérir de la consistance affronter le monde tel qu'il est : injuste, violent, cupide, égoïste. »

Axel Kahn – Éthique en temps de crise. Communication personnelle 12 avril 2020

Quelques règles



- Refus de l'intimidation en tant que façon d'écarter la participation de l'autre (jeu d'autorité, jargon technique...)
- Refus de la manipulation par la séduction
- Refus du mensonge, même dans ses formes subtiles (exagération, non-divulgarion de certaines informations...)
- **Écoute, ouverture d'esprit, prise en considération des arguments de l'autre**
- **S'exprimer, dire son opinion.** Encourager les plus timides
- Chercher à considérer **tous les facteurs** (exhaustivité, données pertinentes)
- Interpeller les exclus (s'interroger sur le point de vue des absents)
- **Mettre en relief les divergences** et analyser la nature et les causes des désaccords
- Aider le groupe à progresser : dégager convergences et divergences, les **conflits de valeurs** en jeu, les **dilemmes** éventuels, « ramasser les éléments susceptibles de faire consensus et d'entrer dans la rédaction d'un éventuel avis ».

J. Habermas - *Éthique de la discussion*, Durand 1999

Vos questions

Indépendance ?

D'où je parle

Impacts ?

Utilisations

? Financement

Simplifications

Responsabilité

? Dérives

Contraintes ?

Liberté

Biais ?