



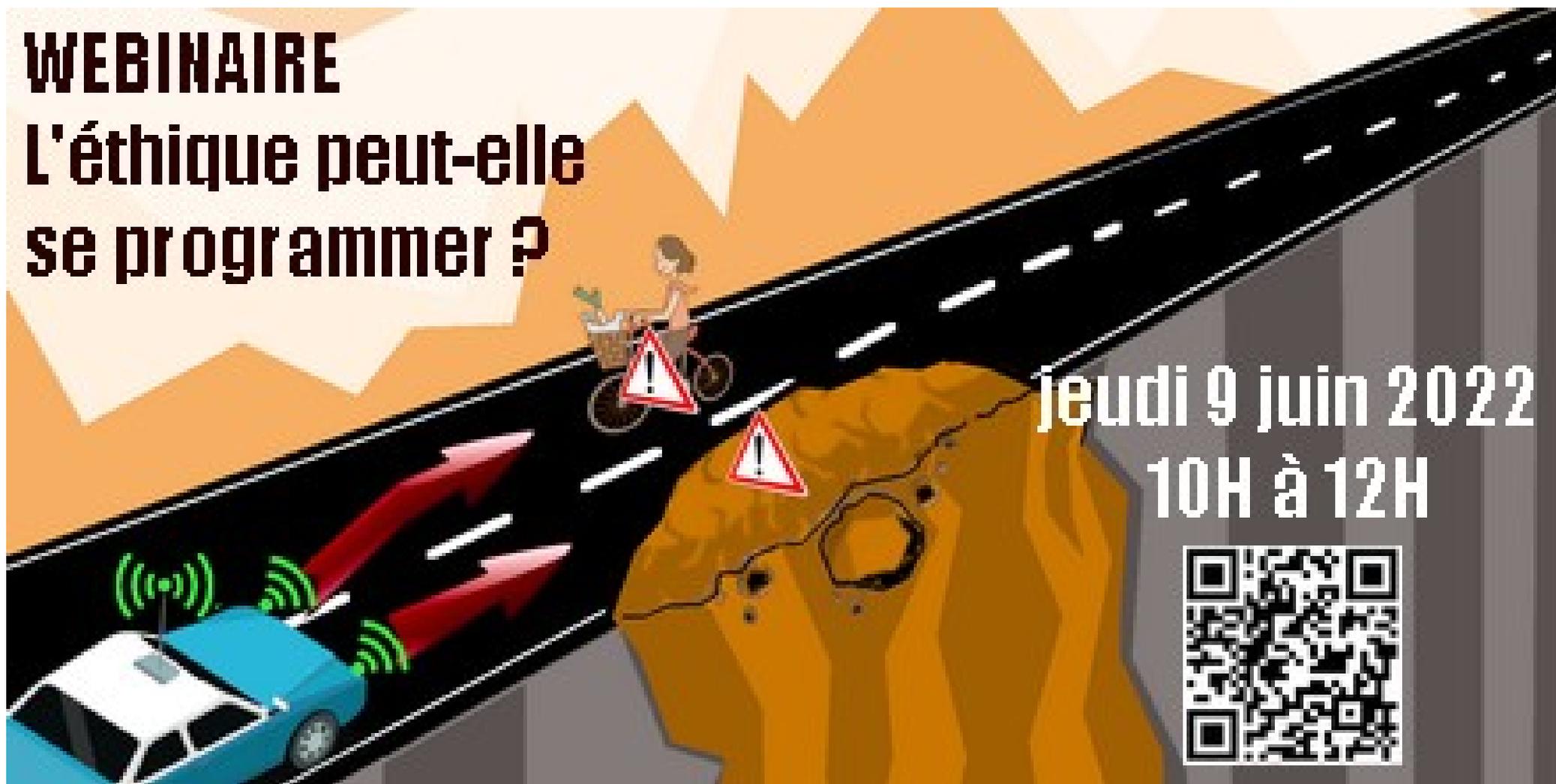
RÉGION ACADÉMIQUE
PROVENCE-ALPES-
CÔTE D'AZUR

*Liberté
Égalité
Fraternité*

Direction régionale académique
du numérique éducatif

WEBINAIRE

L'éthique peut-elle se programmer ?



jeudi 9 juin 2022
10H à 12H



Sommaire

1. quelques apports théoriques

- pour définir les contours de ce dont il s'agit
 - centrer autour de l'exemple de la programmation du véhicule autonome

2. Et en classe ?

- retour d'expérience sur des exercices déjà testés avec des collégiens
 - pistes et ressources prometteuses

I. Ethique et véhicule autonome ou « Qui écraserez- vous ? »

Morale et éthique

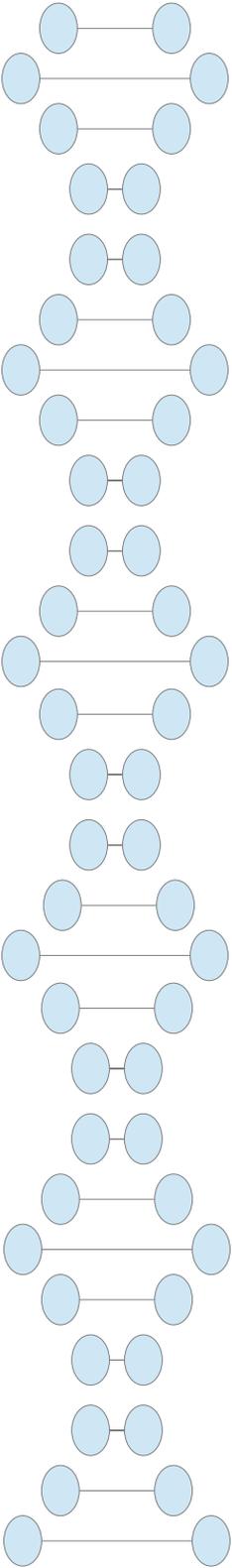
• La morale

- Impersonnelle
- Règles à suivre
- Doit / ne doit pas

• L'éthique

- Plus personnelle
- Restreinte à des groupes (médecins, samourais...)

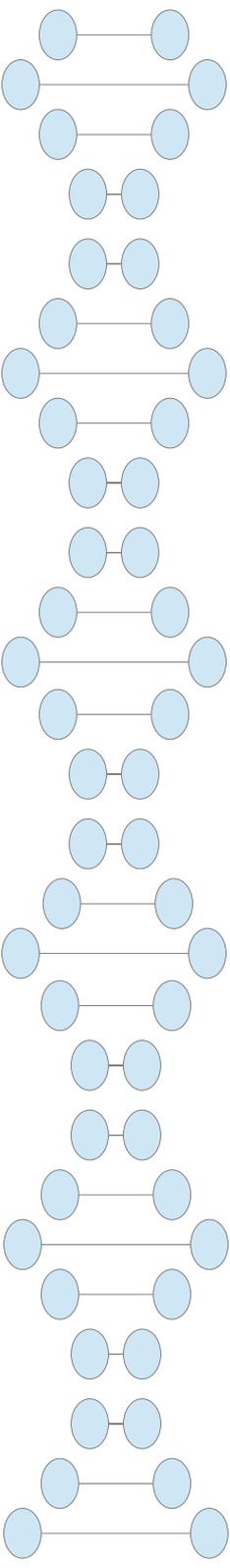
- **Permet de choisir quelle est la Bonne action et comment y parvenir**
- **Étymologie** : mœurs, caractère, règles de conduite, habitudes et justification de ces habitudes
- Viennent de principes religieux, sociaux, philosophiques, ancestrales...



Éthique appliquée

- L'éthique appliquée consiste, en gros, à évaluer, du point de vue de leur valeur politique et morale, les arguments utilisés dans les débats publics relatifs à certains domaines d'action spécifiques (rapport à la nature, aux animaux, biomédecine, relations sexuelles, justice globale, etc.).

Ruwen OGIEN



Éthique appliquée

- Forme de l'éthique « qui au point de contact avec le réel discute des cas concrets et de leur résolution si elle est possible ou bien identifie des dilemmes éthiques dans ce qu'ils ont d'impossible à résoudre, s'ils existent »
- Isabelle PARIENTE BUTTERLIN
- → comités d'éthique

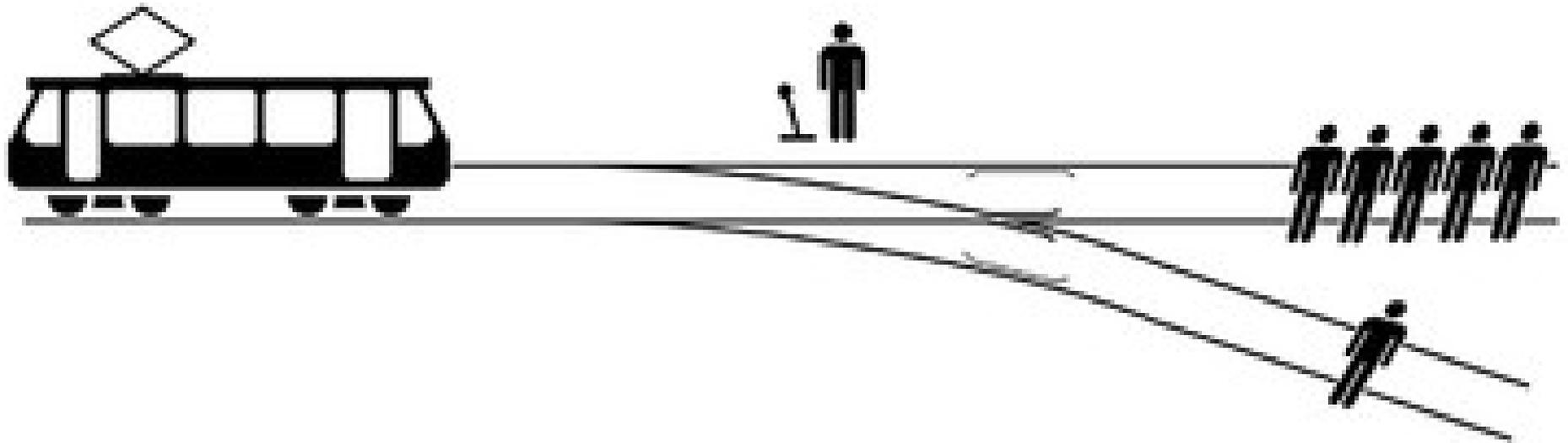
Dilemme du trolley

Dilemme : « nécessité dans laquelle se trouve une personne de devoir choisir entre les deux termes contradictoires et également insatisfaisants d'une alternative. »

- Bon vs meilleur

- Expériences de pensée
- Artificielles
- Test de nos intuitions
- Philippa Foot, 1967
- Judith Jarvis Thompson, 1976

Dilemme du trolley



Et vous, tirez-vous le levier ?

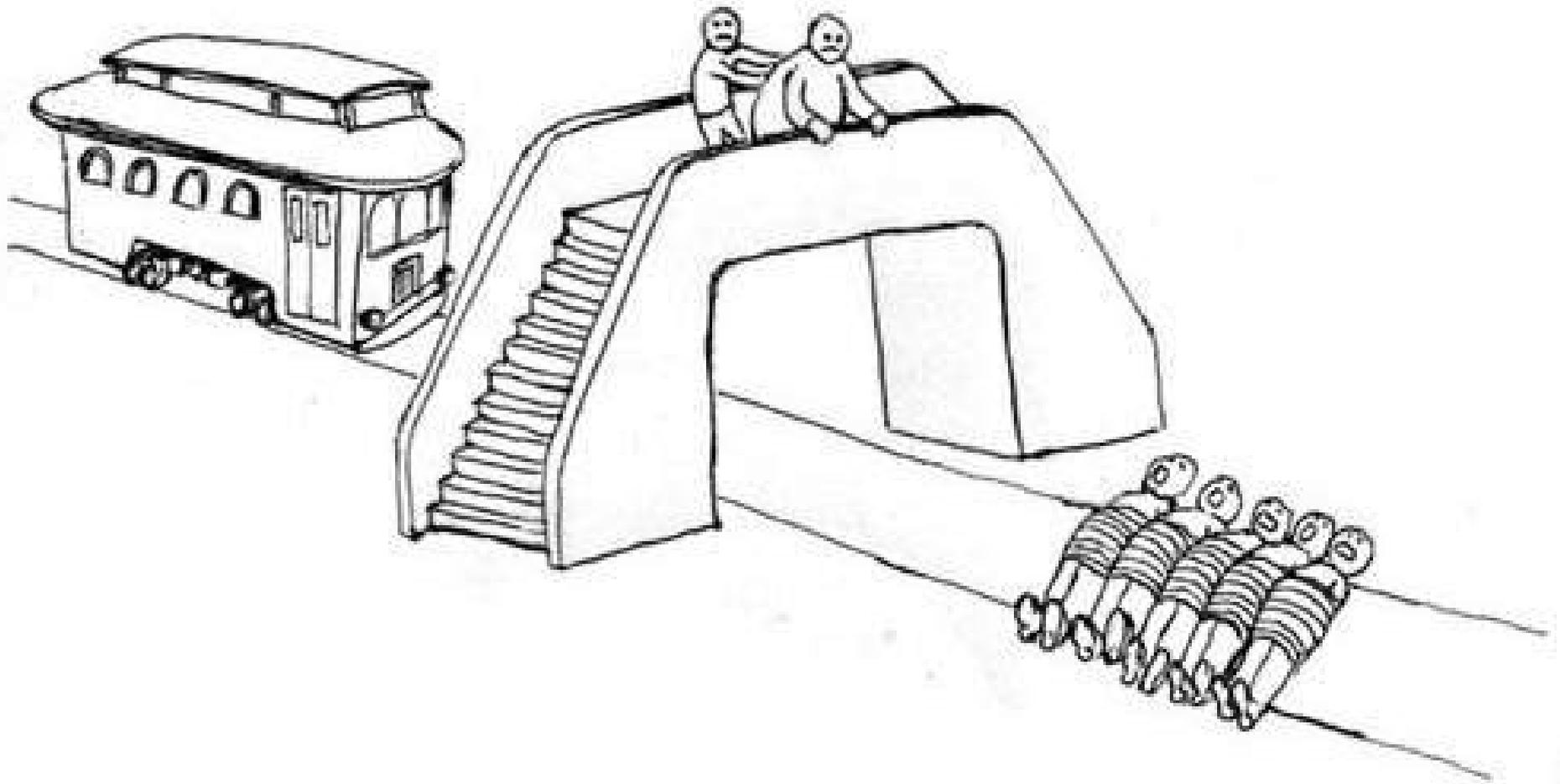
Résultats de l'expérience

- l'équipe du professeur Marc Hauser qui avait mené une large étude sur internet entre 2003 et 2004, portant sur 2600 personnes.
- 89 % : il est moral de détourner le tramway

- Et si la personne seule sur la voie était votre mère / sœur / enfant.. ?

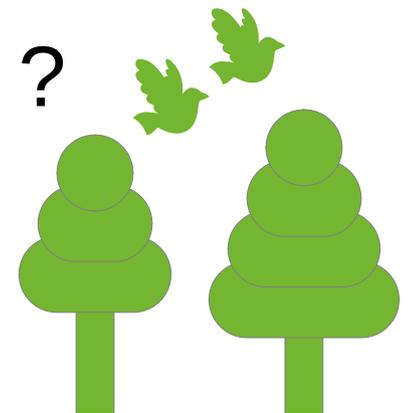


Variante du gros homme



Résultats variante du gros homme

- 11 % : il faut pousser le gros homme
- → universalité des intuitions ou incohérence ?
- → l'utilitarisme a-t-il des limites ? D'ordre déontologiste (refus de considérer le gros homme comme un moyen) ? Ou émotionnelles ?
- Et vous alors, plutôt Kant ou Bentham ?



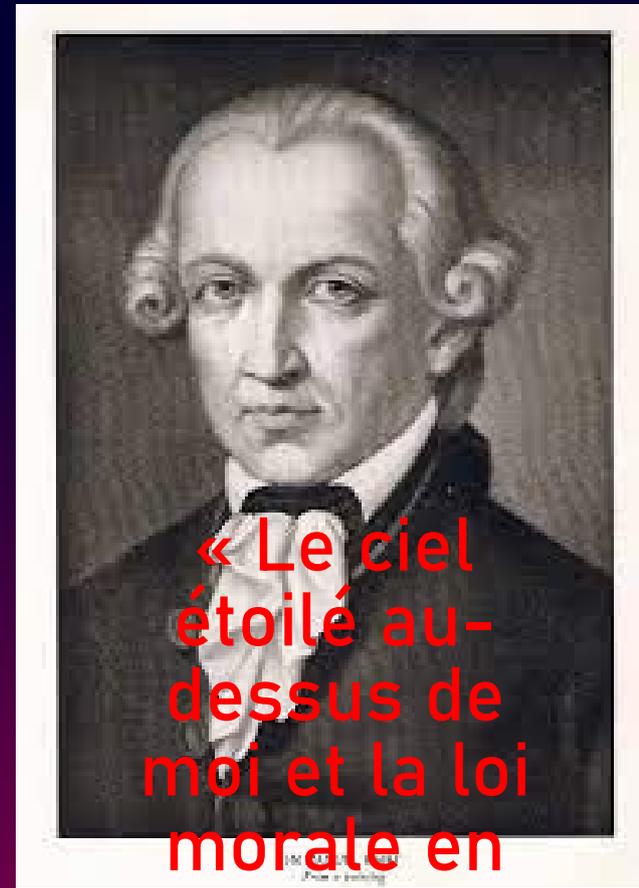
L'éthique déontologiste

Deon = le devoir

Morale = qualité de l'être rationnel
→ Pas d'expert en morale

Test d'universalité

Impératif catégorique qui peut
contraindre la volonté



« Le ciel étoilé au-dessus de moi et la loi morale en moi »

Morale basée sur l'intention

Autrui = toujours une fin,
jamais un moyen

L'éthique utilitariste

La conséquence de l'action prime : conséquentialisme

Calcul d'utilité, felicis calculus basé sur 7 critères: intensité, certitude, fécondité, durée, proximité, pureté et étendue (dernier critère qui consiste au calcul des 6 premiers sur toutes les personnes potentiellement touchées par l'action évaluée)

Mill rajoute la qualité des plaisirs



Principe d'impartialité

Sidgwick : dualisme de la raison pratique

Problèmes : les paramètres de calcul, le monstre d'utilité...

<https://www.youtube.com/watch?v=AZBDMN5wZ-8>

Jusqu'où serez-vous utilitariste ? (Ft. Science4All)



SERAS-TU

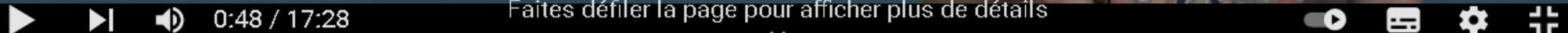
UTILITARISTE

JUSQU'AU BOUT ?



0:48 / 17:28

Faites défilez la page pour afficher plus de détails

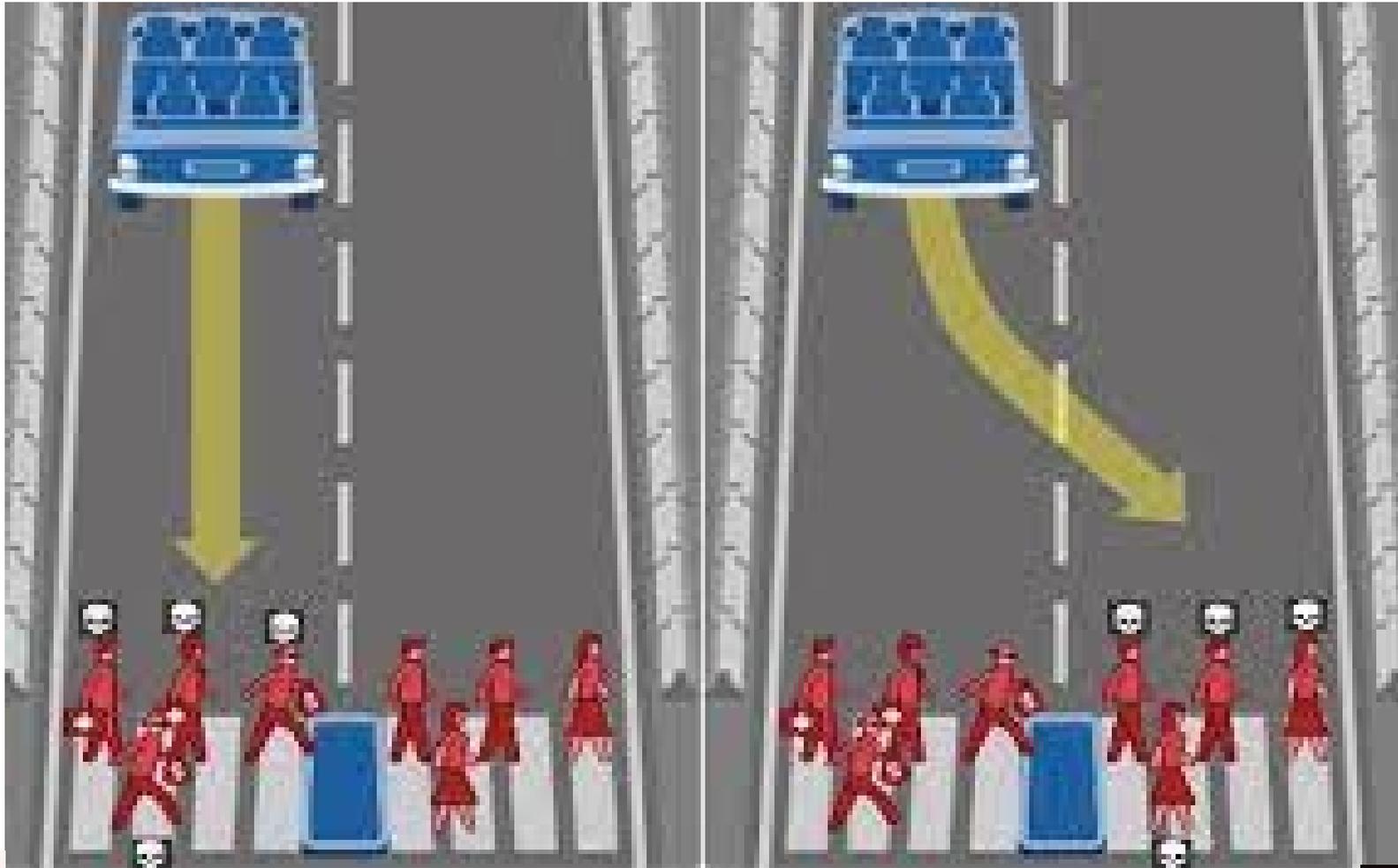


Retour à l'IA :

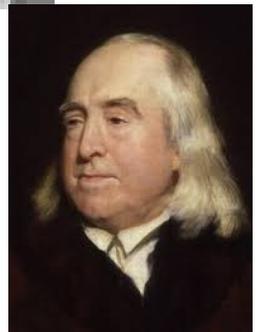
La moral machine

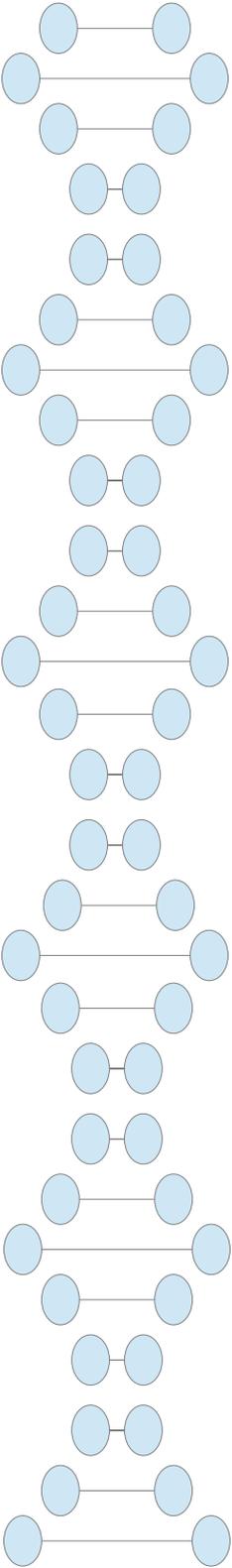
- Outil du MIT
- Créée par Iyam Rahwan, Jean-François Bonnefon
(Président du groupe d'experts de la Commission européenne de l'éthique sur la mobilité sans conducteur, Responsable du programme IA et société du centre digital, Chaire d'IA moral à l'Intelligence Artificielle et Naturelle de l'Institut de Toulouse, Chercheur invité à l'institut Technologique du Massachusetts) et Azim Sharrif
- But : construire une image de l'opinion humaine sur la façon avec laquelle les machines doivent prendre des décisions lorsqu'elles sont confrontées à des dilemmes moraux, proposer ces résultats à la puissance politique comme une base de réflexion pour la législation

<https://www.moralmachine.net/hl/fr>



Alors, qui écrasez-vous ?





Les conclusions de la *moral machine*

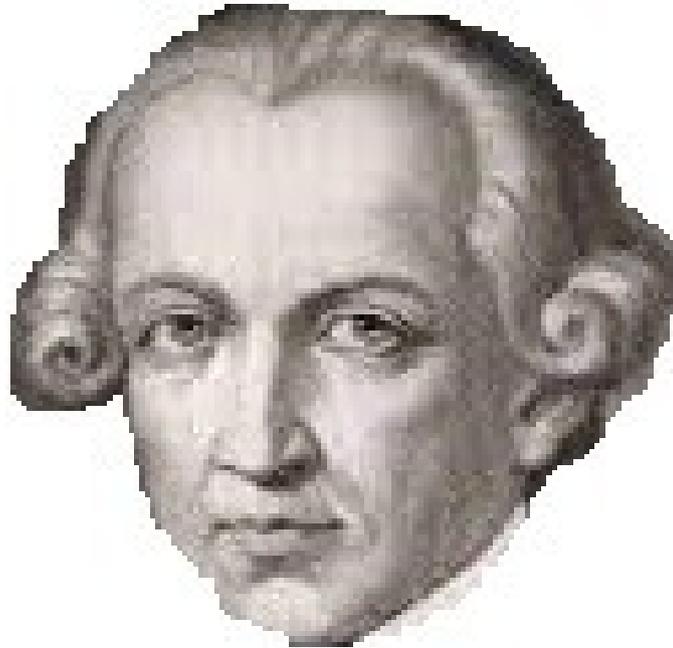
- Utilitarisme vainqueur (mais on ne peut pas ne pas choisir)
- Le nombre de morts prime
- Identification de 3 clusters (différences de culture, éducation, sentiments?)
- « Nous discutons de la manière dont ces préférences peuvent contribuer à développer des principes globaux, socialement acceptables pour l'éthique des machines. »

Peut-on hiérarchiser les vies humaines ?

1/3 des clients potentiels refuseraient d'acheter une voiture programmée de façon utilitariste

- Alexander Dobrindt, ministre allemand des transports, 2017
- → vie humaine > vie animale
- → sauver le plus grand nombre est recommandé mais non requis
- → interdiction de choisir entre les personnes
- → ne pas impliquer une personne extérieure pour en sauver d'autres (cf le gros homme)

- Kant revient par la fenêtre



Si j'avais pu, j'aurais mis Kant jaillissant du tableau de bord d'un véhicule en disant « Nein ». Mais je ne suis assez douée en informatique pour cela...

Indemnisation des compagnie d'assurance et algorithmes

- Si un véhicule autonome doit heurter un édifice pour éviter un accident mortel... que doit faire le programme ?
 - **Solution déontologiste**
 - la vie humaine a un prix supérieur à n'importe quelle infrastructure.
 - → aucun calcul
 - **Solution utilitariste**
 - Si l'infrastructure vaut plus que 9 millions de dollars (estimation du coût d'une vie humaine), la question peut se poser et le calcul se faire.
 - Voir les modalités de calcul des compagnies aux Etats-Unis

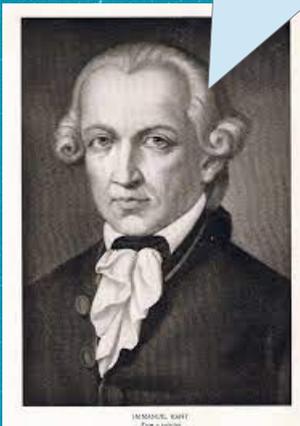
Le prix d'une vie

- Suite aux attentats du World Trade Center, un avocat cherche pour le compte de l'État à savoir comment les victimes seront indemnisées

« L'homme n'est jamais un moyen »

Ca dépend s'il a les moyens...*

* Bien sûr Bentham n'a jamais dit ni même sûrement pensé un truc aussi affreux



Quels paramètres dans l'algorithme

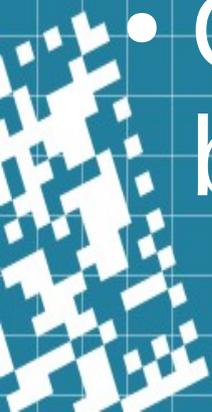
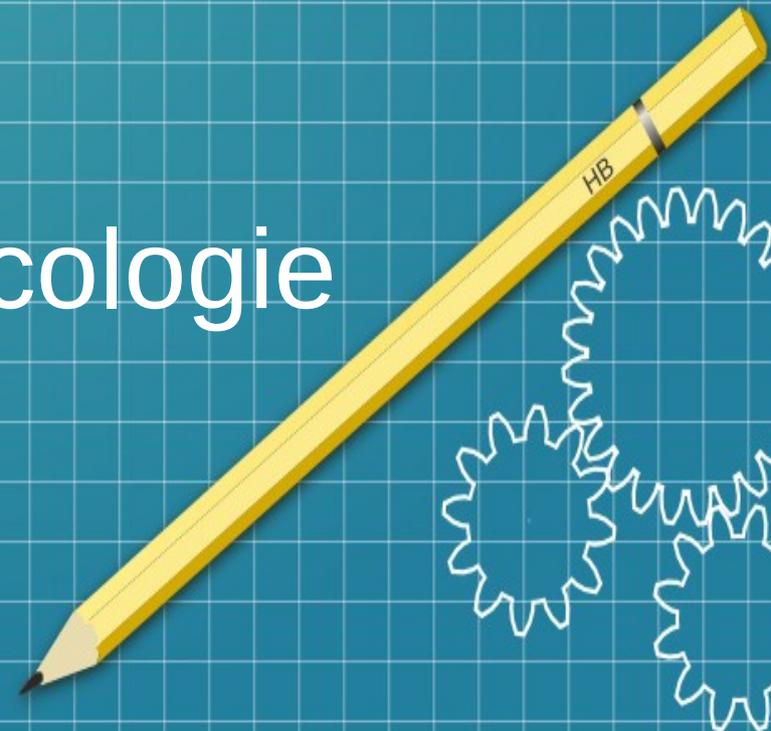
- Problème : quels type de données dans l'algorithme ?
 - Bien être vs souffrance
 - Quelle limite temporelle
 - Somme des personnes concernées ou moyenne
 - Étendre à tous les êtres sensibles ?
- Modulations de l'utilitarisme :
 - Utilitarisme de la règle (comporte des devoirs)

Alors comment programmer ?

- Programmation déontologiste ? Difficile à imaginer
 - Mix utilitariste avec des barrières déontologistes ? L'option allemande
- « Le **hasard** étant une bonne façon de résoudre de nombreux dilemmes lorsqu'on préfère ne pas avoir à trancher. Les dés peuvent incarner une certaine forme de justice. Concrètement, cela implique d'introduire une fonction aléatoire dans l'algorithme. Mobiliser le hasard, c'est dire, en somme, qu'on n'a pas de bonnes raisons morales de choisir une option plutôt qu'une autre » Martin Gibert, *Faire la morale aux robots*
- Proposition soutenue par Alexei Grinbaum
<https://www.youtube.com/watch?v=jIx1z8hW45g>

IA et éthiques : autres questions

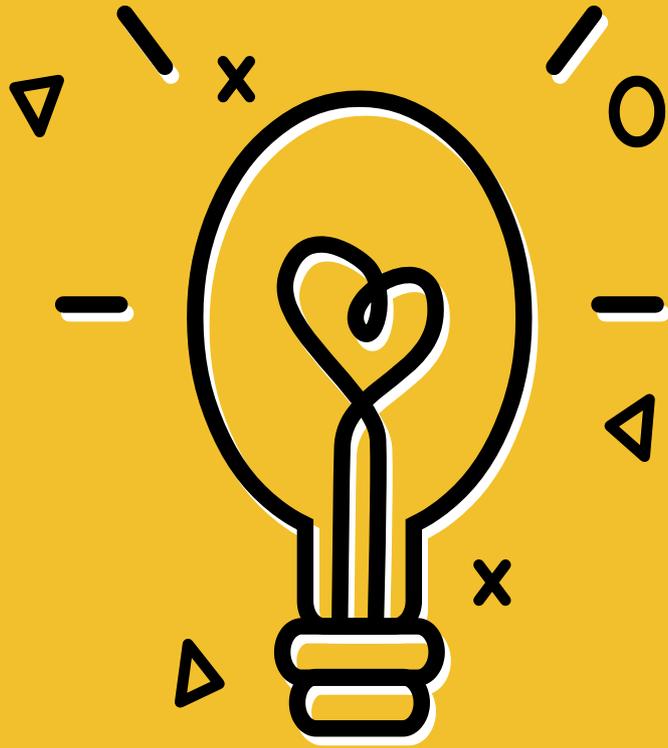
- Biais de programmation
- Préjugés des programmeurs
- Orientation des bases de données
- Protection des données
- Ecologie



Conclusion

- IA extrêmement efficace mais peut-être que tout ne se pose pas en algorithme
- Ne pas abandonner l'humain pour l'efficacité : certaines choses devraient peut-être ne pas être paramétrées → imprédictibilité
- Hypothétique accident : nombre ridicule de cas mais effet de loupe sur des questions essentielles
- Absurdité de la demande éthique qui aboutit à de l'anti-éthique
- La demande d'éthique est parfois un leurre

II. Pistes et ressources pédagogiques



- Quel objectif ?
- Auto-formation
- Jeux sérieux et applications
- Des idées pour les cours

Retour d'expérience sur « jeu débat »

- Constats
 - Admiration sans borne pour la technologie
 - Peu ou pas de conscience des enjeux éthiques et des risques
- Objectifs
 - Ouvrir des questions pour nuancer cette admiration, développer leur esprit critique
 - Les amener à réfléchir par des exemples concrets et peut-être les guider à ce stade
 - faire des expériences de pensée

Auto-formation

En amont, côté enseignant, se renseigner sur les définitions de l'éthique, ce que sont des valeurs, les réflexions menées au niveau national

Concepts, philosophie et large banque de dilemmes moraux

<https://www.enseignerlethique.be/>

<https://www.lefigaro.fr/vox/societe/2019/02/21/31003-20190221ARTFIG00141-l-intelligence-artificielle-peut-elle-etre-ethique.php>

- Rapport de la CNIL de 2017

Former à l'éthique tous les acteurs-maillons de la « chaîne algorithmique » (concepteurs, professionnels, citoyens) : l'alphabétisation au numérique doit permettre à chaque humain de comprendre les ressorts de la machine (Recommandation n°1)

Avec les élèves

Travailler sur la notion de **dilemme** à associer à un **débat**

- dilemme du trolley et ses variantes
- banque de dilemmes :
 - <https://www.enseignerlethequie.be/content/banque-de-cas-%C3%A9thiques>

https://poleia.quebec/wp-content/uploads/2020/02/Guide_IA_VF.pdf

- transition écologique, IA et territoires
- robot et amour
- santé mentale et ia
- langue minoritaire

Avec les élèves (ou pas...)

Aborder la question des **biais de programmation**

Application ludique +++

<https://pixees.fr/classcodeiai/app/tuto1/>

→ jeu / test pour constituer une base de données et créer un programme de reconnaissance

Qu'est-ce-qu'un biais ? Quels sont les vôtres ? (attention, expérience possiblement violente...)

- <https://implicit.harvard.edu/implicit/france/>
- <https://usbeketrica.com/fr/article/personnes-noires-chances-renverser-voitures-autonomes>

Avec les élèves

La question des algorithmes de **recommandations** et des intérêts qui se cachent derrière

- A partir du site
https://kidscodejeunesse.org/fr/ressources-pour-educateurs?subject=ai_algo
- <https://www.youtube.com/watch?v=2oegxbYUFAs>
- <https://algorithmliteracy.org/fr/>
- <https://docs.google.com/document/d/1pQ8D4iDnwKoiveJOZZgy6SLvgDD1nYQOPFUwyuBpEic/edit#heading=h.1et5vs39qkyh>
 - les algorithmes reflètent les opinions
 - matrice éthique
 - Introduction à l'apprentissage automatique supervisé et au biais algorithmique*

Avec les élèves

- Jeu débat
 - un jeu de rôle organisé clé en main
 - élèves mis en situation de débattre de décisions politiques en envisageant les intérêts divers des citoyens
 - Un guide assez complet à proposer avant de lancer le jeu

<http://eidos64.fr/wp-content/uploads/2021/02/A04-AIAethique2.pdf>

Avec les élèves

- **En Lettres**

La nouvelle La brousse de Bradbury, qui aborde la question de l'IA dans un récit d'anticipation. Face à une maison intelligente, quelles sont nos responsabilités ? Quelle humanité nous reste-t-il si c'est la maison qui est humaine ? (Succès garanti avec les élèves!)

- **En Education civique**

Une IA est-elle souhaitable dans les tribunaux pour décharger le travail des magistrats et pré-juger une affaire en fonction des cas similaires déjà jugés ?

<https://www.vie-publique.fr/eclairage/277098-lintelligence-artificielle-ia-dans-les-decisions-de-justice>

- **En SVT**

Domaine médical, choix des patients pendant le COVID : une IA aurait-elle été nécessaire pour faire le tri entre ceux qu'il fallait sauver face à la pénurie de lits et l'urgence des situations ? et si oui, quelles données pour cette IA, quelles valeurs (vacciné ou pas, obèse ou pas, âge.. ?)

- **En Langue**

google traduction et les biais genrés par exemple (<https://www.numerama.com/tech/446024-google-a-compris-que-son-algorithme-de-traduction-contenait-des-biais-sexistes.html>)

Google ne veut plus que Translate reproduise « des préjugés sexistes »

<https://www.presse-citron.net/google-corrige-biais-sexistes-service-translate/>

Une mine de ressources

- <https://lascientotheque.github.io/ressources-ia/Contenu/FichesPedagogiques.html#kids-code-jeunesse>

Bibliographie

Sitographie

- **BIBLIOGRAPHIE**

- *Faire la morale aux robots*, Martin GIBERT, UQAM, 2020

- *L'influence de l'odeur des croissants chauds sur la bonté humaine*, Ruwen OGIEN, Grasset, 2011

- **SITOGRAPHIE**

- Textes de référence

- <https://www.cnil.fr/fr/comment-permettre-lhomme-de-garder-la-main-rapport-sur-les-enjeux-ethiques-des-algorithmes-et-de>

- <https://www.intelligence-artificielle.gouv.fr/fr/thematiques/ethique-de-l-intelligence-artificielle>

Article

- <https://dailyscience.be/31/01/2020/integrer-du-hasard-dans-les-algorithmes-de-lia/>

- <https://www.lefigaro.fr/vox/societe/2019/02/21/31003-20190221ARTFIG00141-l-intelligence-artificielle-peut-elle-etre-ethique.php>

- → <https://usbeketrica.com/fr/article/personnes-noires-chances-renverser-voitures-autonomes>

Sitographie

- Références pédagogiques

→

<https://jeudebat.com/jeux/lintelligence-artificielle/>

→

<https://pixees.fr/classcodeiai/app/tuto1/>

→

https://kidscodejeunesse.org/fr/ressources-pour-educateurs?subject=ai_algo

→

<https://www.youtube.com/watch?v=2oegxbYUFAs>

→ <https://algorithmliteracy.org/fr/>

- Expériences diverses

→ Moral Machine

→ Implicit Project

→ vidéo Mr Phi et Lê

→ bases de données de dilemmes éthiques