



Congrès National
de Médecine &
Santé au Travail

Du 14 au 17 juin 2022
Palais de la Musique et des
Congrès de Strasbourg

Santé et innovation technologique : Mise en perspective philosophique et éthique

Ambre Davat et Thierry Ménissier

Chaire Éthique & IA

2019-2023



<https://www.ethics-ai.fr/>



Chaire Éthique & IA

11 contributrices/teurs, 6 doctorant.es

7 disciplines : philosophie, mathématiques-informatique, robotique, sociologie, psychologie sociale et clinique, sciences de l'information et de la communication, marketing, droit

« Faire dialoguer l'informatique robotique, la philosophie et les sciences humaines et sociales en vue de **la compréhension des enjeux psycho-sociaux, moraux et politiques du déploiement de l'IA**, ainsi que la détermination de **règles éthiques** en vue de solutions compatibles avec les **valeurs de la démocratie** »



Partenariats extra-académiques



life.augmented

StorAIge



États généraux de l'éthique de l'IA

Défis scientifiques / philosophie



- Quelle « intelligence » pour l'IA ?

Philosophie de l'esprit, cognition

- Quelles sont les procédures, les règles et les valeurs les plus pertinents tant pour la conception que pour les usages de l'IA ?

Méta-éthique, éthique normative et appliquée

Philosophie de l'innovation

- Quels types de société/d'humanité promeut-on avec les technologies l'IA ?

Philosophie politique, philosophie des techniques, anthropologie philosophique

- Quelle est la bonne perspective pour déterminer l'IA de manière juste et démocratique ?

Éthique publique, théories de la Justice

3 axes de la chaire

- « Ethique de l'IA en santé » :
coordonnée par Fabienne Martin-Juchat et Ambre Davat
collaboration avec Philippe Cinquin et la chaire « *Deep Care: Patient Empowerment via a Participatory Health Project* »
- « Problèmes du droit en éthique de l'IA » :
coordonné par Serge Slama et Sophie Guicher
collaboration avec Théodore Christakis et la chaire « *Legal and regulatory implications of AI* »
- « Concepts de la *computer ethics* »
coordonné par Louis Devillaine et Thierry Ménissier

Définition élargie de l'intelligence artificielle

Au sens strict : l'IA n'existe pas

(uniquement « des systèmes d'IA faibles », conçus pour fonctionner selon un cahier des charges très précis ou dans un environnement déterminé)

Au sens restrictif : disciplines scientifiques spécifiques (ex : *deep learning*)
mais frontières floues (systèmes experts ? robotique ?)

Au sens large : algorithmes / technologies numériques

Place centrale de la technique

Technique / **technologie** ?

- La science se pense comme liée aux dispositifs techniques
- Elle est enrôlée par l'industrie
- Développement exponentiel de l'automatisme
- Johann Beckmann, *Technologie*, 1777 / Jacob Bigelow, *Elements of technology*, 1829

High Technologies, « technologies de pointe »

« **Société de haute technologie** » :

- Quand le recours à l'artifice, à l'automatisme & à l'expertise scientifique et technologique devient systémique
- Quand le plus haut niveau des découvertes scientifiques et des inventions techniques est largement diffusé dans la société (usages courants / compétences des usagers)
- Quand l'économie d'une société est polarisée par le développement de valeur dans ces domaines

Transformations du travail par la technologie

Le moment du **passage** d'un système technique à l'autre engendre de très grands **bouleversements** sociaux, des tensions, des luttes et éventuellement de la violence.

ex : le *luddisme en Angleterre*

=> *13 ouvriers condamnés à mort et pendus en 1812 pour avoir détruit des métiers à tisser mécaniques*

Aujourd'hui : pas de bris de machine, mais
une augmentation des risques psycho-sociaux



Révolution technologique ?

Transformation de notre rapport au temps, à l'information, à la prise de décision, au travail, à la santé...

- Ph. Spach, « Risques humains et enjeux éthiques des innovations digitales RH », *Management & data sciences*, 21/12/20 <https://management-datascience.org/articles/14632/>

= l'outil numérique ne doit pas se substituer à l'humain, mais au contraire permettre aux professionnels de se concentrer sur leur cœur de métier

=> idéal de **remplacement des tâches fastidieuses et pénibles** (fin du « labor » d'Arendt ?)

- « Risques éthiques », Chardel & Reber 2011

= **reconfiguration de la communication interpersonnelle**

=> disparition des éléments de contexte essentiels pour se comprendre et créer du lien

=> instantanéité de la communication à distance

=> dissolution du sens des responsabilités

=> possibilité d'une surveillance constante

Propriétés / effets attendus du numérique (Pégny, 2022)

- Dépersonnalisation de la décision => équité
- Explicitation des raisons => transparence
- Uniformisation du traitement => efficacité et massification
+ contrôle centralisé

En pratique...

(Pégny, 2022)

- Dépersonnalisation de la décision ~~X~~ > équité
dispositifs socio-techniques, qui reflètent les valeurs et les rapports de forces de leur époque (« biais » algorithmiques)
- Explicitation des raisons ~~X~~ > transparence
accès au code et à sa compréhension difficile, complexité croissante des systèmes, voire abandon de l'explicabilité au profit de la performance (« boîte noire »)
- Uniformisation du traitement ~~X~~ > efficacité et massification
travail discrétionnaire des agents négligé au moment de la conception et de la mise en place de ces dispositifs => perte d'efficacité et transformation des métiers

Multiplication des chartes « éthiques »

Transparence

Justice et équité

Non-malfaisance

Responsabilité

Respect de la vie privée (Privacy)

Bienfaisance

Liberté et autonomie

Confiance

Soutenabilité

Dignité

Solidarité

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1(9), 389-399.

Principes fondateurs aux définitions variables
et parfois incompatibles

Quelle méthode d'arbitrage ?

(Tessier, 2022)

Éthique : démarche philosophique visant à hiérarchiser (et expliciter) ces valeurs

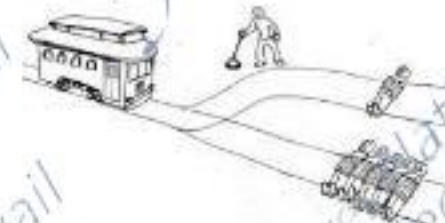
=> **non normative** (distinction éthique / morale)

=> **appliquée** : traite de cas particuliers, de situations **complexes** dans lesquelles les intérêts et les valeurs des différentes parties prenantes entrent en **conflit**

Apports de la philosophie : 1) élargir l'éthique de l'IA

Un type de raisonnement éthique particulier est le plus souvent employé pour l'IA

= Utilitarisme conséquentialiste (calcul coûts / bénéfices) :
un type de raisonnement « intéressé »



Les autres familles de l'éthique ne sont pas réellement sollicitées alors qu'elles pourraient l'être

- déontique : éthique de la loi et de la morale (« devoir moral »)
- arétaique : éthique des vertus (tient compte des intentions de l'agent)
- axiologique : éthique des valeurs et de leur hiérarchisation
- Problème de la formation en théorie morale des acteurs de l'IA à l'évaluation éthique : financeurs, concepteurs, prescripteurs

Apports de la philosophie : 2) conceptualisation

(Ménissier, 2022)

1. Éthique **informatique** (*Computer ethics*)

- Conception des systèmes d'IA (SIA) → informaticiens

2. Éthique **artificielle** (*Artificial/algorithmic/robotic ethics*)

- Conception, programmation & usages des machines intelligentes autonomes
→ roboticiens, usagers

3. Éthique **du numérique & de la donnée** (*Digital & data ethics*)

- Conception et usage des plateformes dans la dynamique de l'innovation de l'économie numérique → acteurs économiques, usagers
- Création et gestion des données → usagers, pouvoirs publics, acteurs économiques

4. Éthique **des usages de l'IA** (*UX AI ethics*)

- Relations entre les dispositifs technologiques et leur sens / valeurs en vigueur dans une société → philosophes, usagers, pouvoirs publics

Points de débat

- Certains secteurs peuvent-ils revendiquer une **primauté hiérarchique** sur les autres ?
Ex : les spécialistes d'IA sur tous les autres ?
- Le travail du **droit** accompagne chaque niveau : la **régulation** comme éthique appliquée ?
Ex : RGPD dans l'espace européen + proposition de réglementation sur les SIA
Droit et lobbying...
- Existe-t-il une valeur commune qui ne relève pas des particularismes culturels :
Ec : La durabilité = éthique environnementale de l'IA

Ethique des usages de l'IA

Vise à impliquer dans sa démarche
(« **compréhension-évaluation** » (Schnapper, 2012)) :

- Les concepteurs des dispositifs sociotechniques
- Les usagers
- Les *social scientist* : économie, sociologie, psychologie sociale, anthropologie

Cette démarche demeure **philosophique** :

- Évaluation **de** concepts & valeurs **par des** concepts & valeurs
- Ces concepts & valeurs : **polysémiques, contradictoires entre eux et problématiques**

Liens avec les méthodes participatives

Apparues depuis les années 70 dans de nombreux domaines

(informatique, éducation, urbanisme...)

(Zamenopoulos, Alexiou, 2018)

Dans le domaine de l'informatique :

conception « centrée-technologie » => « **centrée sur l'utilisateur** »

(Bødker, King, Grønbaek, 1995)

(Bødker, Ehn, Sjögren, Sundblad, 2000)

Idées clés :

Concevoir pour avec les différentes parties prenantes

Reconnaître l'expertise des utilisateurs / usagers

Place de la santé au travail ?

Face à la **révolution technologique**, il est nécessaire de réinventer les concepts, les imaginaires, les usages...

La médecine du travail a une **place privilégiée** pour participer à cette réinvention, de part ses liens étroits avec les différents acteurs du corps social (et les différentes générations), les entreprises et les institutions.



36^e

Congrès National de Médecine & Santé au Travail

Du 14 au 17 juin 2022
Palais de la Musique et des
Congrès de **Strasbourg**

Merci de votre attention !

ambre.davat@univ-grenoble-alpes.fr
thierry.menissier@univ-grenoble-alpes.fr

<https://www.ethics-ai.fr/>