

Comment l'IA fait-elle des choix « moraux »

Comment traduire les valeurs humaines
en une éthique programmable ?

AI in balance: Intelligence artificielle, société et développement durable

6 Octobre 2026

FRDO/CFDD

Carl MÖRCH, FARI

An Initiative of



Supported by



Co-funded by
the European Union

THE BRUSSELS CAPITAL REGION AND EUROPE INVEST IN YOUR FUTURE



Répondre en trois temps

1. IA et moralité, par où commencer?
2. Traduire des valeurs humaines, oui mais lesquelles?
3. Une éthique programmable, quel horizon?

IA et moralité

Par où commencer?

Des risques sociétaux

On voit une accélération cette année des craintes profondes sur les risques posés par des systèmes d'intelligence artificielle sur l'humanité, notre société, etc.



Un domaine

- Pour aborder toutes ces questions, l'éthique de l'IA est devenue essentielle.
- L'IA comme champ académique date des années 50, et les questionnements éthiques ont commencé dès les années 60.
- L'éthique de l'IA fait écho à une question fondamentale: si quelque chose est techniquement possible (ex. grâce à l'IA), doit-on le faire?
- C'est un sous-ensemble de l'éthique de la technologie.

Pour comprendre de “vieux nouveaux” problèmes



Pour comprendre nos interactions avec l'IA

Welcome to

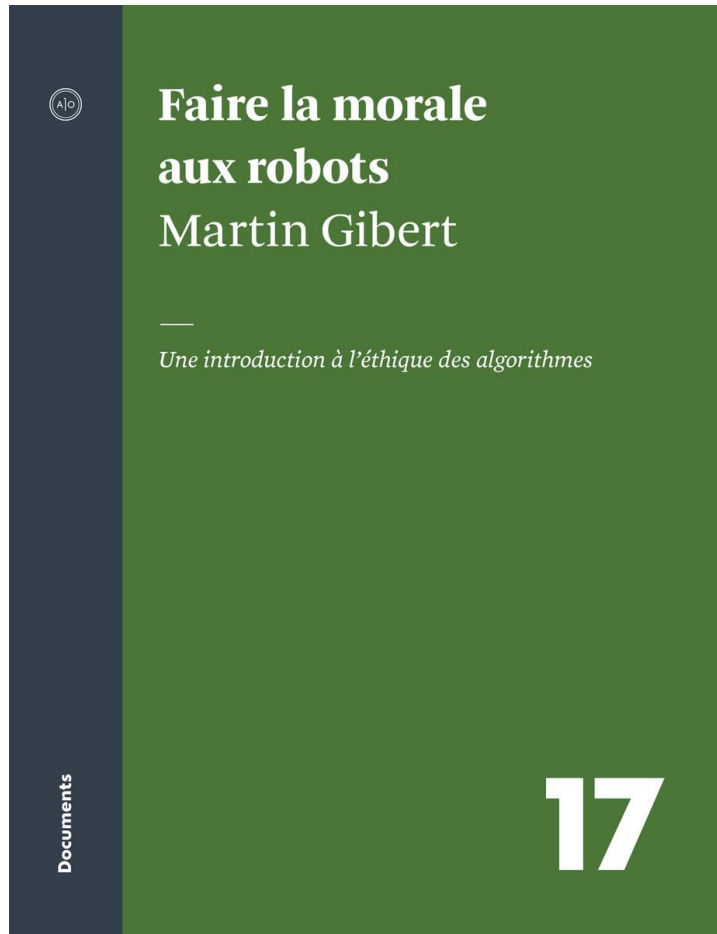
```
EEEEEE LL      IIII  ZZZZZZ  AAAAA
EE      LL      II    ZZ    AA  AA
EEEEEE LL      II    ZZ    AAAAAA
EE      LL      II    ZZ    AA  AA
EEEEEE LLLLLL  IIII  ZZZZZZ  AA  AA
```

Eliza is a mock Rogerian psychotherapist.
The original program was described by Joseph Weizenbaum in 1966.
This implementation by Norbert Landsteiner 2005.

```
ELIZA: Is something troubling you ?
YOU:   Men are all alike.
ELIZA: What is the connection, do you suppose ?
YOU:   They're always bugging us about something or other.
ELIZA: Can you think of a specific example ?
YOU:   Well, my boyfriend made me come here.
ELIZA: Is it important to you that your boyfriend made you come here ?
YOU:   He says I'm depressed much of the time.
ELIZA: I am sorry to hear that you are depressed.
```



Une petite définition



"L'éthique appliquée à l'intelligence artificielle est le domaine de l'éthique qui se demande ce qui est bon, juste ou vertueux dans la mise en œuvre des systèmes d'intelligence artificielle"

(Gibert, 2019)

Note: On voit qu'il parle ici de mise en œuvre (pas des IA en elles-mêmes)

Une IA est le fruit de choix humains

- Les systèmes d'IA sont en effet le fruit de choix humains (de l'infrastructure derrière, aux données utilisées).
- Aucun outil technologique n'est neutre, et l'IA n'y fait pas exception.
- Elle est inscrite dans des dynamiques techniques, sociales et de pouvoir.
- Les "choix moraux" sont conditionnés en amont, et tout relatifs.

INSIGHTS

Calculating and Powers: Interview with Kate Crawford

It is urgent today to look at AI for what it is: a system based on chains of resource exploitation that are not always transparent

by Giorgia Marino

RENEWABLE MATTER#48 - 2024 AI & Tech



16 JAN 2024 10:00

De nombreuses questions soulevées

- Est-ce que les développeurs (et acteurs technologiques) contribuent à développer une IA responsable et “morale”?
- Est-ce que les IA sont morales ou douées de compétence morale?
- Comment intégrer des valeurs dans des systèmes d'IA pour les rendre moraux?



La compétence morale: ça existe?

The notion of moral competence in the scientific literature: a critical review of a thin concept

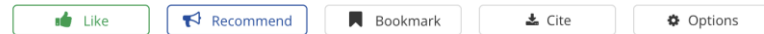
Dominic Martin, Carl-Maria Mörch & Emmanuelle Figoli

Ethics and Behavior 33 (6):461-489 (2023) [Copy](#) [BibTeX](#)



Abstract

This critical review accomplished two main tasks: first, the article provides scope for identifying the most common conceptions of moral competence in the scientific literature, as well as the different ways to measure this type of competence. Having moral judgment is the most popular element of moral competence, but the literature introduces many other elements. The review also shows there is a plethora of ways to measure moral competence, either in standardized tests providing scores or other non-standardized tests. As a second task, the article reflects critically on the general use of the idea of moral competence. Results suggest that this idea functions as what philosopher Bernard Williams would have called a thin ethical concept. Thin concepts are not problematic in and of themselves if they are used as linguistic shorthand, however, there may be shortcomings in the literature on moral competence because the idea is not rooted in more substantive views or theories in a way that is both clear and coherent.

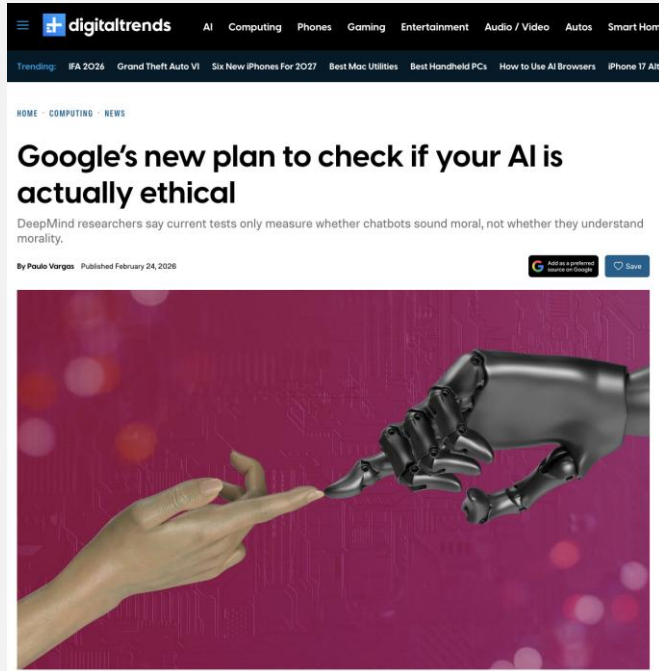


C'est déjà très difficile pour des humains d'évaluer et de mesurer leur compétence morale.

Nous avons évalué que le concept de “compétence morale” était un “thin concept”, c'est-à-dire un terme parapluie qui recouvre plusieurs réalités.

Ce n'est pas un problème en soi, mais cela n'est pas une idée stabilisée, mesurable et ancrée dans des théories communes.

Et la compétence des “machines”?



“La question de savoir si les grands modèles de langage (LLM) peuvent faire preuve de capacités morales suscite un intérêt et une urgence croissants, car ces systèmes sont déployés dans des rôles sensibles tels que l'accompagnement et le conseil médical, et seront de plus en plus chargés de prendre des décisions et d'agir au nom des humains”

(Haas et al., 2026)

Une complexité vertigineuse

“Nous avons dû adopter une terminologie qui s’applique de la même manière à toutes les entités qui le peuplent (des animaux aux organisations), sans pour autant préjuger de nos conclusions. Nous avons dû analyser les comportements selon une approche non anthropocentrique.”

(Floridi, Sanders, 2004)

On the morality of artificial agents

Luciano Floridi & J. W. Sanders

Minds and Machines 14 (3):349-379 (2004) [Copy](#) [BIBTeX](#)

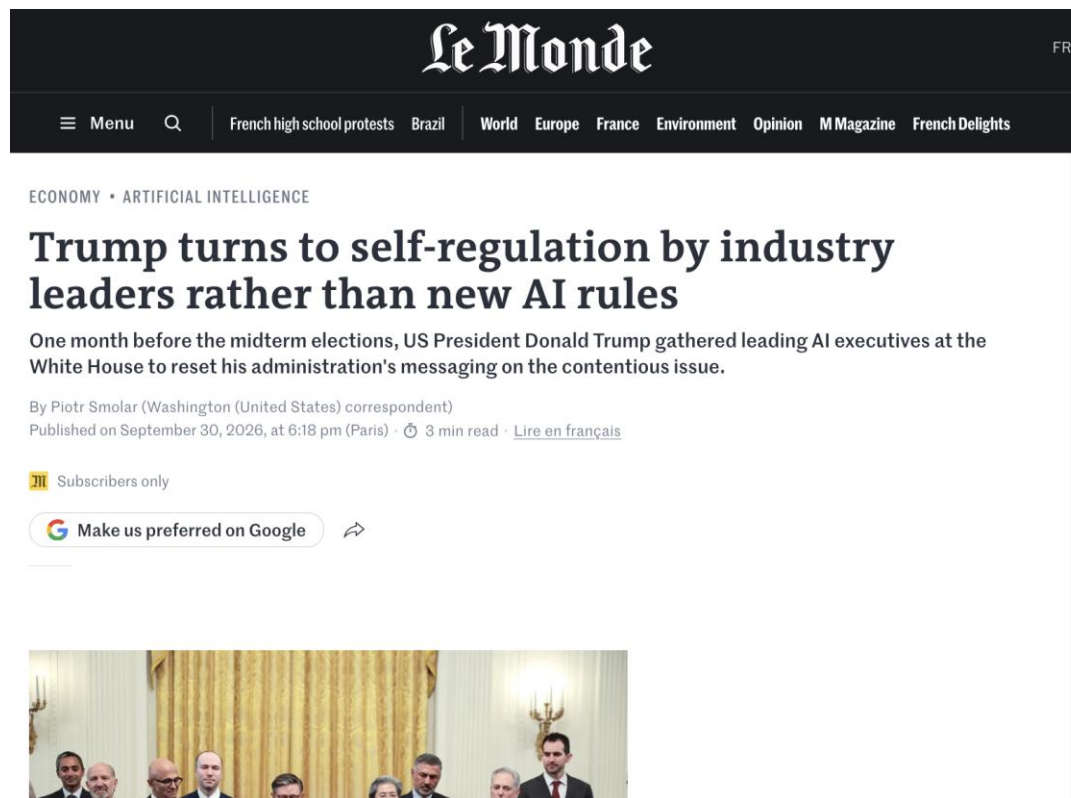
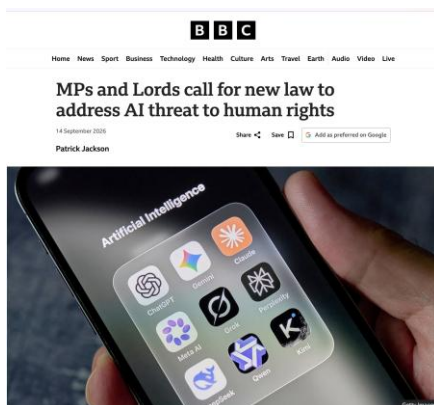
2. Traduire des valeurs humaines

Oui mais lesquelles?

Des valeurs traduites en lois

L'IA est de plus en plus régulée.

Suscitant encore de nombreux débats internationaux.



Est-ce suffisant?



On le sait, le Droit ne couvre pas tout.

L'éthique permet de voir au-delà, d'ouvrir des questions plus larges et spécifiques.

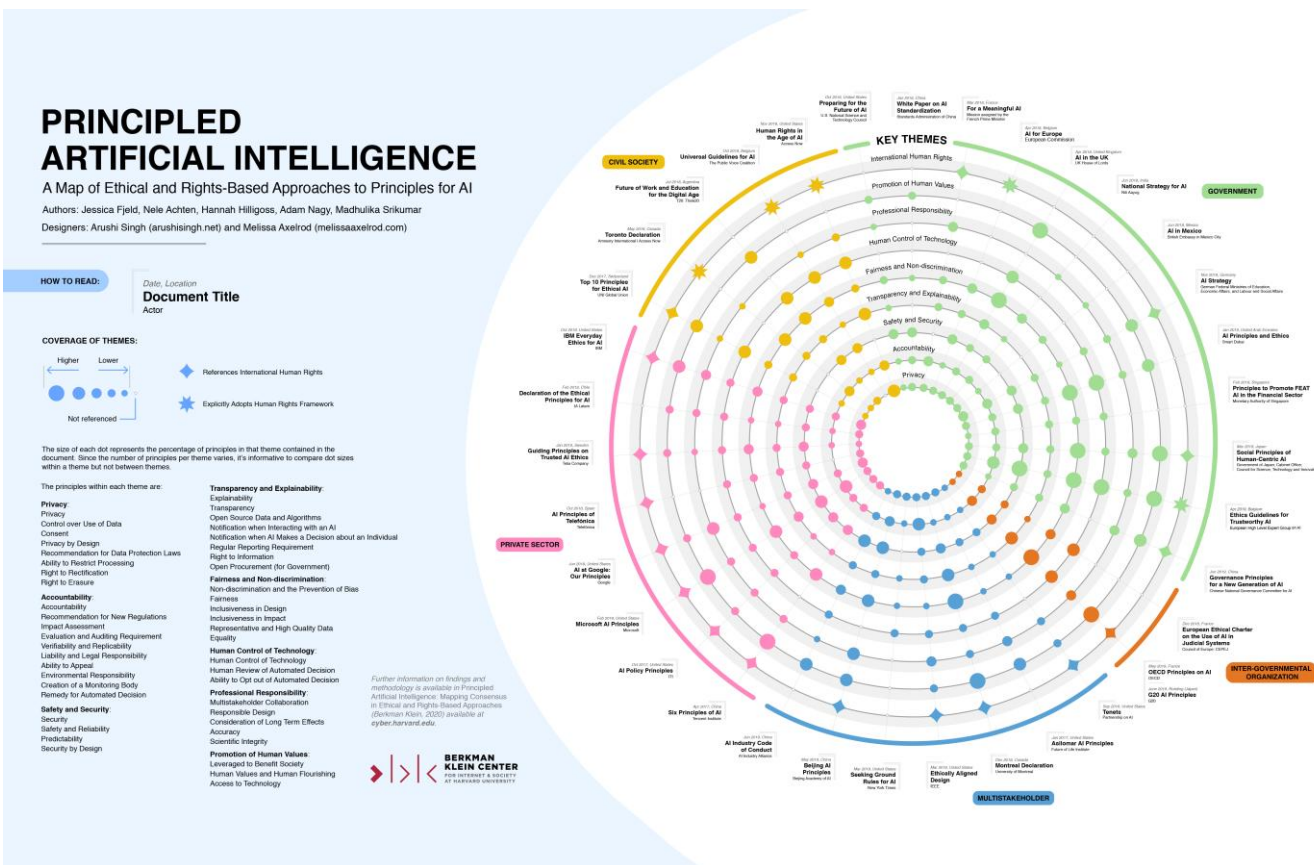
Et plus encore, les sociétés évoluent, les règles et les situations varient. Comment prendre en compte toute cette diversité?

Les règles et les valeurs sont très nombreuses

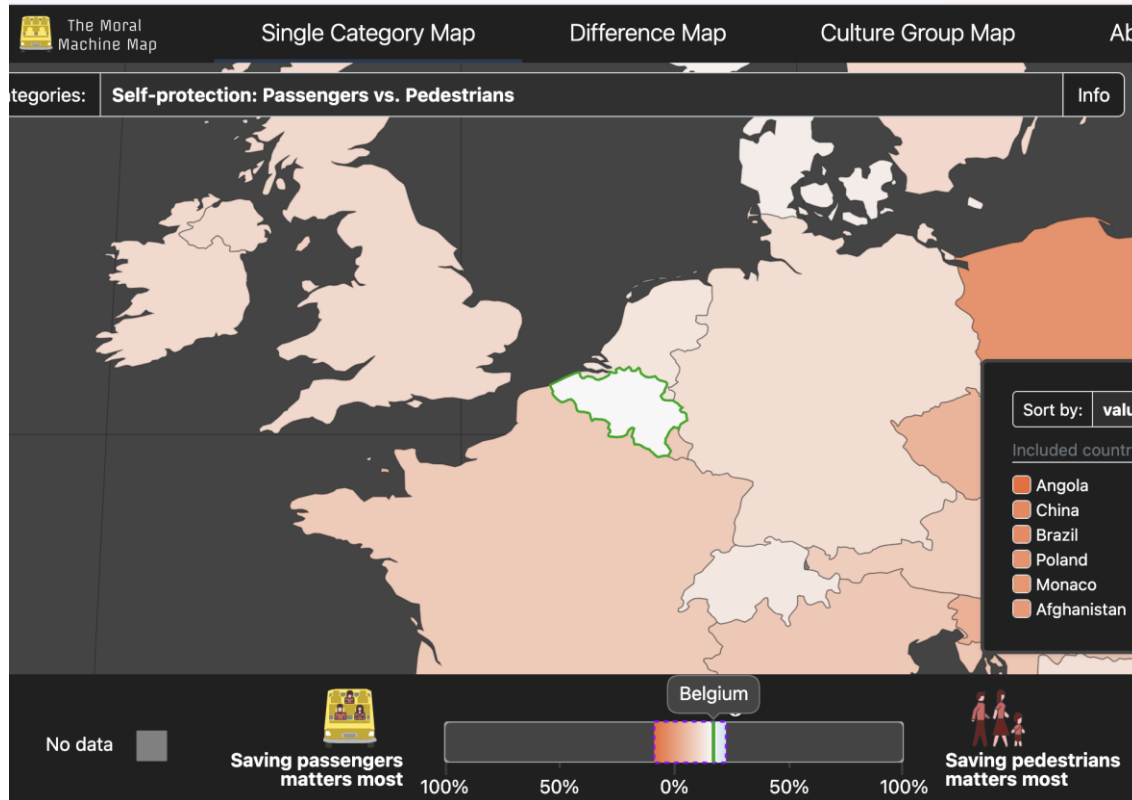
Des centaines de
déclarations éthiques (en
plus des lois) ont vu le jour.

Le défi est immense:

- L'expression des principes et valeurs éthiques varie selon qui les édicte
- Elles ont tendance à converger, mais elles ne sont jamais identiques



Des valeurs universelles?



Dans l'étude *Moral Machine* (Awad et al., 2018), près de 40 millions de décisions ont été recueillies auprès de répondants issus de 233 pays et territoires, afin d'examiner comment un véhicule autonome devrait se comporter face à un dilemme de type « problème du tramway ».

Le degré de consensus varie fréquemment, par exemple selon le critère considéré (par exemple l'âge ou le statut social des personnes impliquées).

Adapter l'IA

C'est le grand défi pour nos sociétés: il faut adapter l'IA aux cultures, aux pays, aux domaines de pratique, etc.



CENTRE DE RECHERCHE EN ETHIQUE

FUNDAMENTAL
ETHICS

ETHICS AND
POLITICS



The Open Dialogue on AI Ethics. Contribution to the UNESCO Recommendation on the Ethics of Artificial Intelligence

Posted on 27 August 2020 by [valery](#) in [Artificial intelligence](#), [News of CRE](#)

[Rapport](#) résultant du processus délibératoire qui a porté sur les recommandations de l'UNESCO sur l'éthique de l'IA, produit par [Marc-Antoine Dilhac](#) et son équipe.

Un exemple : la santé mentale

La Libre

Guerre en Ukraine Hantavirus Belgique Libre ECO International

Belgique Société

"Un précédent grave qui doit être pris très au sérieux" : après le suicide d'un Belge, Mathieu Michel veut mieux protéger les utilisateurs de l'IA

Le secrétaire d'État à la digitalisation compte prendre ce drame au sérieux pour éviter que ça ne se reproduise.

La Rédaction

Publié le 28-03-2023 à 11h00

Enregistrer



Mathieu Michel réagit après le suicide d'un Belge lié à une discussion avec un chatbot. ©Marie Russillo

HAI Stanford University Human-Centered Artificial Intelligence

About Research Education Policy AI Index

NEWS

Exploring the Dangers of AI in Mental Health Care

DATE: JUNE 11, 2025

TOPICS: HEALTHCARE GENERATIVE AI

A new Stanford study reveals that AI therapy chatbots may not only lack effectiveness compared to human therapists but could also contribute to harmful stigma and dangerous responses.

Psychologencommissie Commission des Psychologues

Home Psychologists Patients + Search a psychologist Register

Accueil >> Opinions and Reactions of the Commission of Psychologists

Opinions and Reactions of the Commission of Psychologists

AI-generated "fake psychologists" in the media: vigilance is needed

When disinformation meets artificial intelligence

On June 26, 2025, VRT revealed that Elle Belgium magazine had published online articles under fictitious identities generated by artificial intelligence. These fake journalist profiles signed articles that were supposed to inform the public.

Among them was a certain "Femke," who claimed: "As an

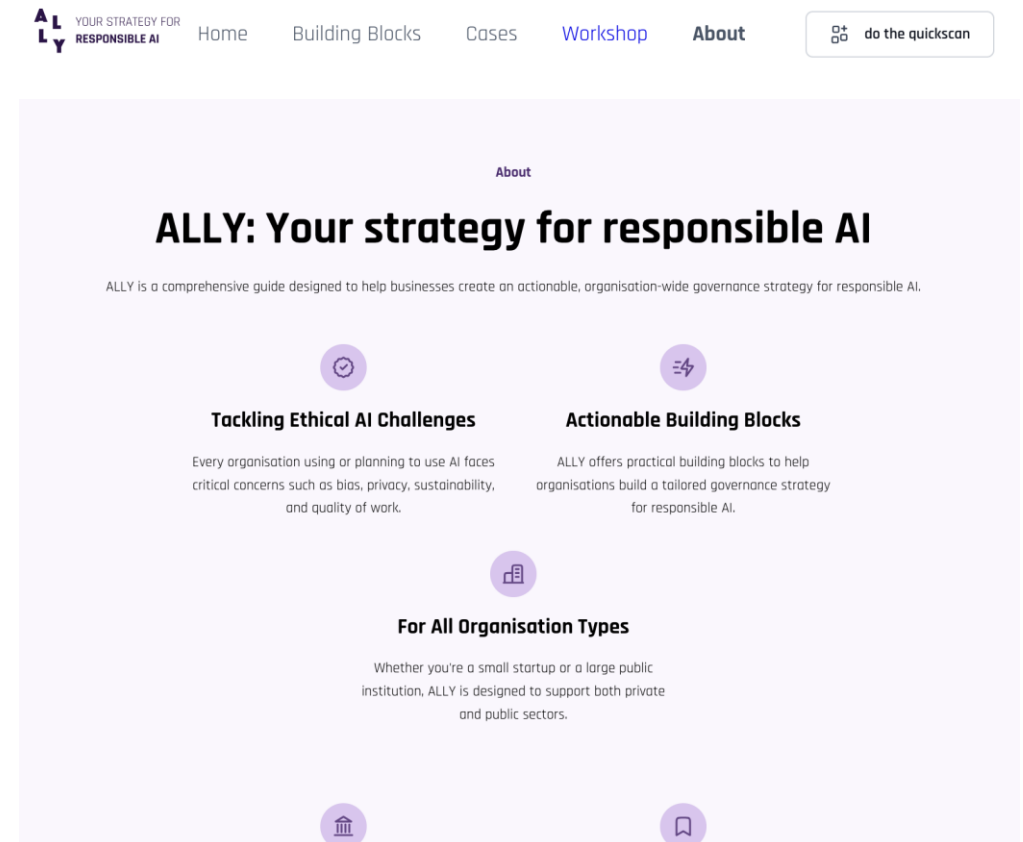


3. Une éthique programmable

Quel horizon?

Des principes à la pratique

- L'éthique de l'IA est parfois très “haut niveau” et ne se transfère pas aisément vers les milieux de pratique.
- Pour preuve, de nombreux domaines (en santé en particulier) essaient de développer des outils pour les aider à intégrer l'éthique.
- Nous voyons sur le terrain un besoin de “traduction”, de démarches pédagogiques “pas à pas”.

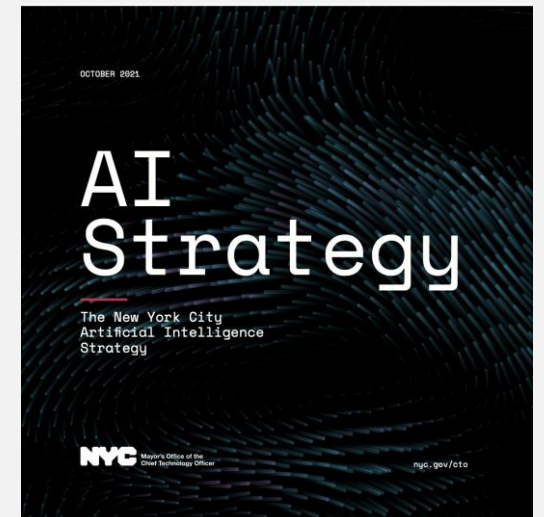


Comment on "programme de l'éthique"

- On peut utiliser des normes et standards
 - Ex. ISO/IEC 42001: Establishes an auditable Artificial Intelligence Management System (AIMS) for organizational governance and risk control.
- On peut utiliser des protocoles éthiques
 - Ex. Le Z-Inspection
- On peut utiliser des approches ciblant des individus ou des organisations
 - Ex. AI Blindspot ou Ally (KDM/FARI)

Stratégies et chartes

- Pour aider à “traduire” les grands principes, on voit des approches complémentaires émerger.
- Des organisations se dotent de chartes (comme celle que nous avons accompagnée avec AI4Belgium pour la fonction publique fédérale).
- D'autres se dotent (comme de nombreuses villes) de stratégies qui rendent les concepts applicables.



Un exemple: l'IA dans l'éducation

Accueil → FR → L'Université → intelligence artificielle → Academ-IA - Plan de formation

Academ-IA - Plan de formation à l'Intelligence Artificielle



“ **Explorez et maîtrisez les possibilités de l'IA à l'ULB**

Academ-IA est le **plan de formation institutionnel** de l'ULB, visant à accompagner la communauté universitaire dans les transformations et les évolutions provoquées par les outils d'Intelligence Artificielle Générative (IAG) dans l'enseignement supérieur. Ce plan a pour objectif de sensibiliser et d'outiller les étudiant.es, les enseignant.es, le personnel administratif (PATGS) et les chercheur.es, grâce à des formations et des ressources adaptées, dans le but de développer une culture responsable et éthique de l'IAG à l'ULB.

SORBONNE UNIVERSITÉ

FORMATION ET VIE ÉTUDIANTE RECHERCHE ET INNOVATION EUROPE ET INTERNATIONAL CULTURE UNIVERSITÉ

Accueil * Dossiers * Intelligence artificielle * L'IA dans l'éducation : entre opportunités et défis

L'IA dans l'éducation : entre opportunités et défis

Entretien avec Céline Gainet

Publié le 7/02/2025 - Mis à jour le 10/02/2025 Lecture 6 min.



2. L'IA DANS L'ÉDUCATION : ENTRE OPPORTUNITÉS ET DÉFIS

- 1. « AVEC POSTGENAI@PARIS, SORBONNE UNIVERSITÉ RENFORCE SA VISIBILITÉ EN IA »
- 2. PRÉSERVER SON ATTENTION À L'HEURE DU NUMÉRIQUE
- 9. "COMMUNS DÉMOCRATIQUES" : L'IA ÉTHIQUE AU SERVICE DES DÉBATS CITOYENS
- 10. AI-TRIOMPHE : QUAND L'INTELLIGENCE ARTIFICIELLE OPTIMISE LES ESSAIS THÉRAPEUTIQUES

<https://www.ulb.be/fr/intelligence-artificielle/academ%C2%B7ia-plan-de-formation>

Vers de nouvelles professions?

AI Ethicists in Europe

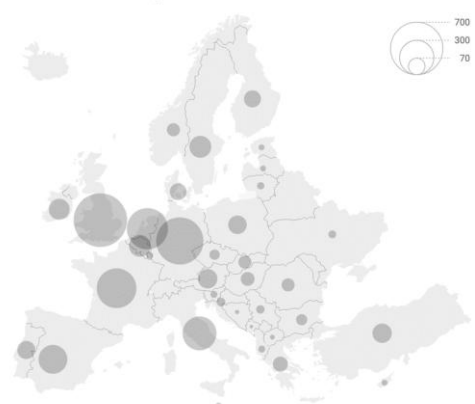


Figure 1 - Distribution of AI Ethicists on LinkedIn in Europe in April 2024

Who is an AI Ethicist? An Empirical Study of Expertise, Skills, and Profiles to Build a Competency Framework

25 Pages • Posted:

[Marianela Zoe Cocchiari](#)
Jagiellonian University in Krakow - Faculty of Philosophy; Alma Mater Studiorum University of Bologna

[Jessica Morley](#)
Yale University - Digital Ethics Center

[Claudio Novelli](#)
University of Bologna- Department of Legal Studies; Yale University - Digital Ethics Center

[Enrico Panai](#)
Università Cattolica di Milano

[Alessio Tartaro](#)
Università degli Studi di Sassari

[Luciano Floridi](#)
Yale University - Digital Ethics Center; University of Bologna- Department of Legal Studies
Date Written: July 10, 2024

Abstract

Over the last decade the figure of the AI Ethicist has seen significant growth in the ICT market. However, only a few studies have taken an interest in this professional profile, and they have yet to provide a normative discussion of its expertise and skills. The goal of this article is to initiate such discussion. We argue that AI Ethicists should be experts and use a heuristic to identify them. Then, we focus on their specific kind of moral expertise, drawing on a parallel with the expertise of Ethics Consultants in clinical settings and on the bioethics literature on the topic. Finally, we highlight the differences between Health Care Ethics Consultants and AI Ethicists and derive the expertise and skills of the latter from the roles that AI Ethicists should have in an organisation.

Keywords: AI Ethicist, ethics of AI, expertise, practical moral expertise, skills

On voit l'émergence d'une profession (Éthicien de l'IA), et/ou d'un titre qui gagne en popularité (et des associations professionnelles voient aussi le jour).

Il existe indéniablement un vide juridique et de la confusion. *“Plusieurs voix critiques ont souligné que l'éthique de l'IA est déjà vulnérable à la mainmise de l'industrie (Saltelli et al., 2022). Ce risque est particulièrement aigu dans les applications pratiques et professionnelles. Pour protéger les éthiciens de l'IA, les organisations ne devraient pas réduire leur rôle à un simple signe de vertu ou à un lavage éthique. Les éthiciens de l'IA doivent être libres d'exprimer des points de vue critiques et ne pas être contraints de défendre les intérêts de leur employeur.”*

Un rêve pour demain

Quel horizon

Pour quoi doit-on militer?

- Pour une plus grande participation des citoyens dans la création des algorithmes de demain
- Pour une plus grande discussion entre acteurs scientifiques, citoyens, publics et privés.



Une vision au coeur de la ville



Un rêve pour demain

- Il sera impératif (doux rêve) de concevoir collectivement les algorithmes qui nous gouvernent.
- Créer (et protéger) des espaces où ces discussions peuvent avoir lieu.



[Accueil](#) > [Paroles de](#) > Ecriture participative du code, un enjeu démocratique. Échange avec Hugues Bersini

Ecriture participative du code, un enjeu démocratique. Échange avec Hugues Bersini

Paroles de | 12 novembre 2021

Hugues Bersini revient sur les enjeux démocratiques de la participation citoyenne à l'écriture du code. Il est professeur d'informatique et directeur du laboratoire d'intelligence artificielle à l'Université libre de Bruxelles et responsable académique de l'initiative FARI.

Intégrer les citoyens

- Nous allons avec le projet BrusselAir (FEDER) créer des instances de “citoyens-codeurs”, où nous intégrerons des usagers dans le design de solutions technologiques.
- Par le passé, nous avons organisé de nombreuses consultations citoyennes (sur l'énergie, sur la ville et les données).
- Notre but: mieux comprendre comment nous pouvons reprendre la main.



Ana Pop Stefanija: A citizen jury for socially acceptable AI systems

AUTEUR — Léa Rogliano, Ana Pop Stefanija

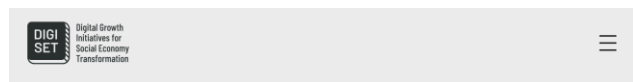
Quand les chercheurs et la société civile imaginent l'avenir de nos villes

En 2023, FARI – l'Institut de l'IA pour le Bien Commun, en collaboration avec Paradigm, a lancé un appel à projets auprès de ses centres de recherche partenaires afin de stimuler l'engagement citoyen autour de l'Intelligence Artificielle (IA), des données et de la robotique. Sept projets ont été soutenus.



Ana Pop Stefanija, chercheuse postdoctorante au sein de l'imec-SMIT, Vrije Universiteit Brussel, a expérimenté la méthode du jury citoyen pour explorer les usages éthiquement et socialement acceptables de représentants d'IA dans la

Du travail devant nous (l'éducation à l'IA)



Digital Growth Initiatives for Social Economy Transformation

The project aims to empower social economy organisations, particularly social SMEs, in their digitalization journey, with a focus on AI, data, and innovative business models. Through collaborative efforts between social and digital SMEs and enabling organisations, the project fosters the development of digital skills and knowledge exchange.



LIBRA.I. (Erasmus+ Project)

URBAN & PUBLIC AI

MAR 2025

LIBRA.I. (Erasmus+ Project)

On November 1, 2024, [FARI - AI for the Common Good Institute \(VUB & ULB\)](#) became a part of the [LIBRA.I. project](#), which will run until October 31, 2026. This initiative, funded by [Erasmus+](#), unites [4 partners](#) from [Belgium](#), [Germany](#), [Poland](#), and [Latvia](#) to integrate Artificial Intelligence (AI) into media and learning literacy trainings in public libraries. The project aims to [upskill librarians by providing AI-related training](#), enabling them to guide the public in understanding digital media, combating disinformation, addressing the ethical implications of AI, and teaching how to use AI in everyday life.

Algo → Lit
Algorithmic literacy
for effective transparency
in the eu

En Fr Nl Menu

Citizens accross Europe are struggling with opaque algorithms.

We aim to train digital inclusion workers to adress this issue.

Par où commencer?



Merci

Contact information



fari.brussels



info@fari.brussels



Cantersteen 16, 1000 Bruxelles
Kantersteen 16, 1000 Bruxelles

Our social networks

FARI - AI for the Common Good Institute



An Initiative of



Supported by



Co-funded by
the European Union

THE BRUSSELS CAPITAL REGION AND EUROPE INVEST IN YOUR FUTURE

