

Hoe AI “morele” keuzes maakt

Hoe vertalen we menselijke waarden
naar een programmeerbare ethiek?

AI in balance: Artificiële intelligentie, samenleving en duurzame ontwikkeling

6 oktober 2026

FRDO/CFDD

Carl MÖRCH, FARI

An Initiative of



Supported by



THE BRUSSELS CAPITAL REGION AND EUROPE INVEST IN YOUR FUTURE



Een antwoord in drie stappen

1. AI en moraal, waar beginnen we?
2. Menselijke waarden vertalen, maar welke?
3. Programmeerbare ethiek, welke vooruitzichten?

AI en moraal

Waar beginnen we?

Maatschappelijke risico's

Dit jaar zien we de diepe bezorgdheid over de risico's die AI-systemen inhouden voor de mensheid, onze samenleving enz. in sneltempo toenemen.



Een vakgebied

- Om al deze vragen aan te pakken, is AI-ethiek essentieel geworden.
- AI als academisch vakgebied dateert van de jaren 50, en de eerste ethische vragen doken al op in de jaren 60.
- AI-ethiek sluit aan bij een fundamentele vraag: als iets technisch mogelijk is (bv. dankzij AI), moeten we het dan ook doen?
- Het is een deelgebied van de technologie-ethiek.

“Oude nieuwe” problemen begrijpen



Onze interacties met AI begrijpen

Welcome to

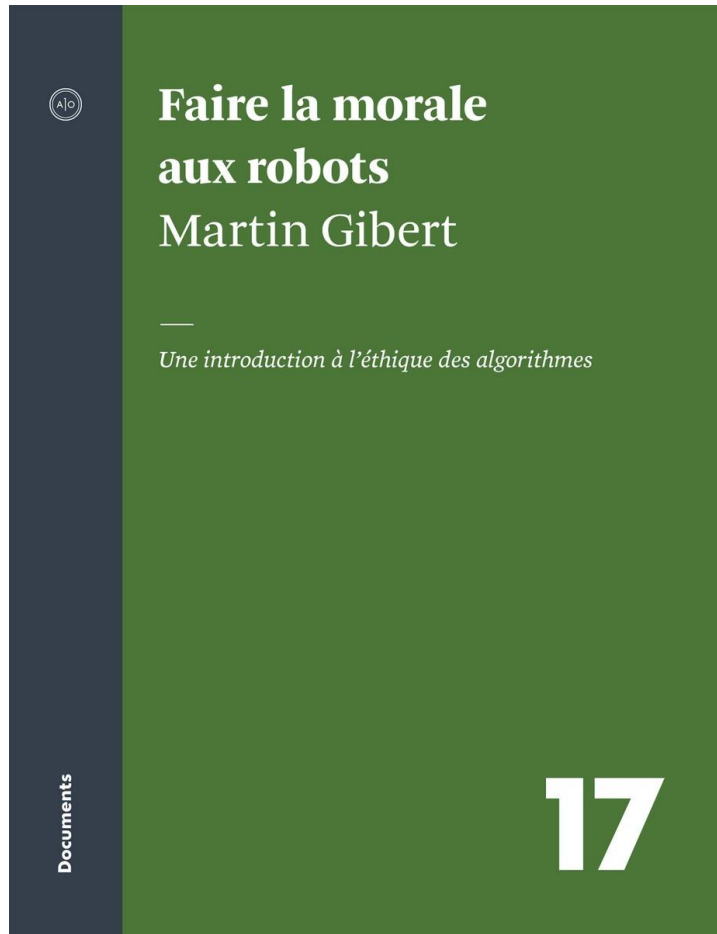
```
EEEEEE LL      IIII  ZZZZZZ  AAAAA
EE      LL      II    ZZ    AA  AA
EEEEEE LL      II    ZZ    AAAAAA
EE      LL      II    ZZ    AA  AA
EEEEEE LLLLLL  IIII  ZZZZZZ  AA  AA
```

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.
```



Een korte definitie



"Ethiek toegepast op artificiële intelligentie is het domein van de ethiek dat zich afvraagt wat goed, rechtvaardig of deugdzaam is bij de inzet van systemen van artificiële intelligentie"

(Gibert, 2019)

Opmerking: hij heeft het hier over de inzet (niet over AI op zich)

AI is het resultaat van menselijke keuzes

- AI-systemen zijn immers het resultaat van menselijke keuzes (van de infrastructuur erachter tot de gebruikte data).
- Geen enkel technologisch instrument is neutraal, ook AI niet.
- AI is ingebed in technische en sociale dynamieken en in machtsverhoudingen.
- “Morele keuzes” worden al vooraf bepaald en zijn volstrekt relatief.

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 Georgia Marino

RENEWABLE MATTER#48 - 2024 AI & Tech



16 JAN 2024 10:00
AI & Tech

Tal van vragen

- Dragen ontwikkelaars (en technologiespelers) bij aan de ontwikkeling van verantwoorde en “morele” AI?
- Zijn AI-systemen moreel, of beschikken ze over morele competentie?
- Hoe integreren we waarden in AI-systemen om ze moreel te maken?



Morele competentie: bestaat dat?

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

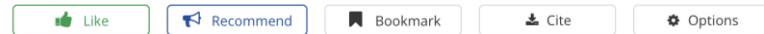
Dominic Martin, Carl-Maria Mörch & Emmanuelle Figoli

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.

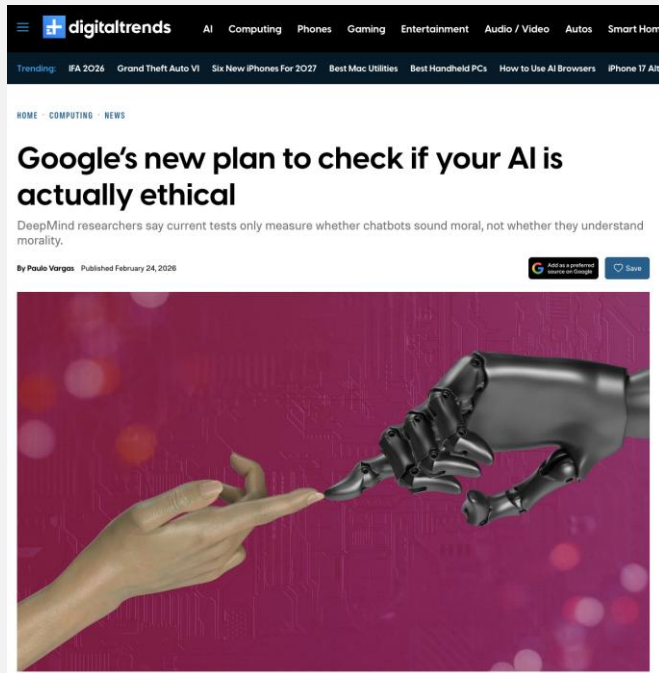


Voor mensen is het al heel moeilijk om hun morele competentie te beoordelen en te meten.

We stelden eerder vast dat het begrip “morele competentie” een “thin concept” is, dat wil zeggen een paraplubegrip dat verschillende realiteiten dekt.

Dat is op zich geen probleem, maar het is geen vastomlijnd, meetbaar idee dat verankerd is in gedeelde theorieën.

En de competentie van “machines”?



“De vraag of grote taalmodellen (LLM’s) morele capaciteiten kunnen vertonen, wekt steeds meer interesse en urgentie, aangezien deze systemen worden ingezet in gevoelige rollen zoals begeleiding en medisch advies, en steeds vaker beslissingen zullen moeten nemen en handelen namens mensen”

(Haas et al., 2026)

Een duizelingwekkende complexiteit

“We moesten een terminologie hanteren die op dezelfde manier van toepassing is op alle entiteiten die het bevolken (van dieren tot organisaties), zonder vooruit te lopen op onze conclusies. We moesten gedrag analyseren vanuit een niet-antropocentrische benadering.”

(Floridi, Sanders, 2004)

On the morality of artificial agents

Luciano Floridi & J. W. Sanders

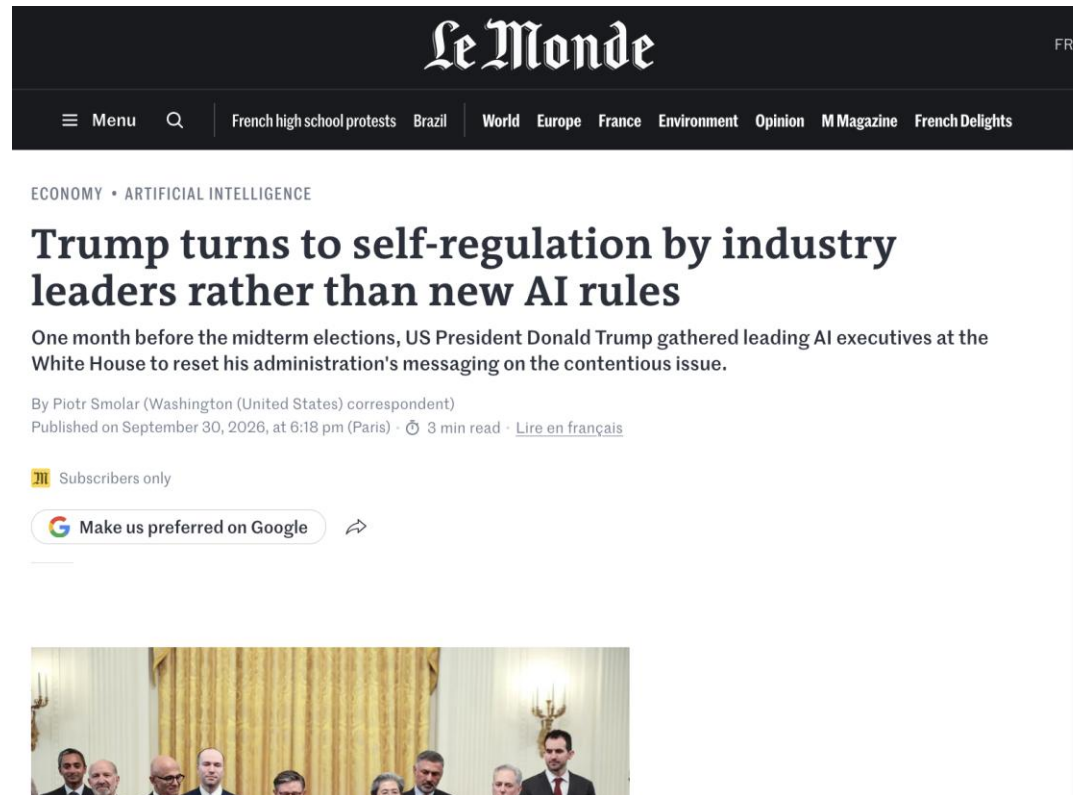
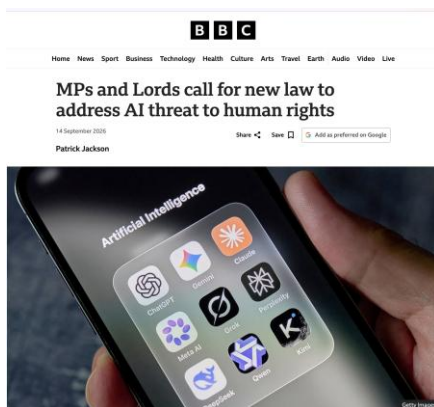
Minds and Machines 14 (3):349-379 (2004) [Copy](#) [BibTeX](#)

2. Menselijke waarden vertalen

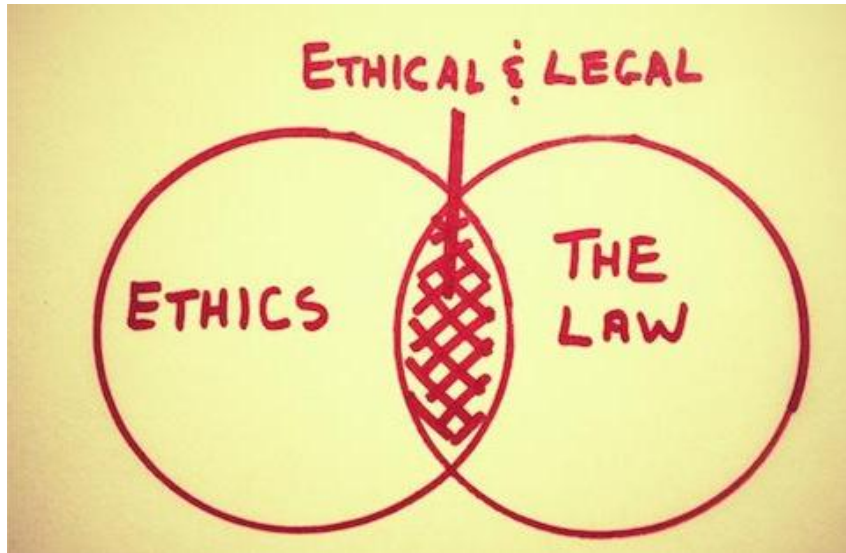
Ja, maar welke?

Waarden vertaald in wetten

AI wordt steeds meer gereguleerd. Nog altijd onderwerp van veel internationaal debat.



Volstaat dat?

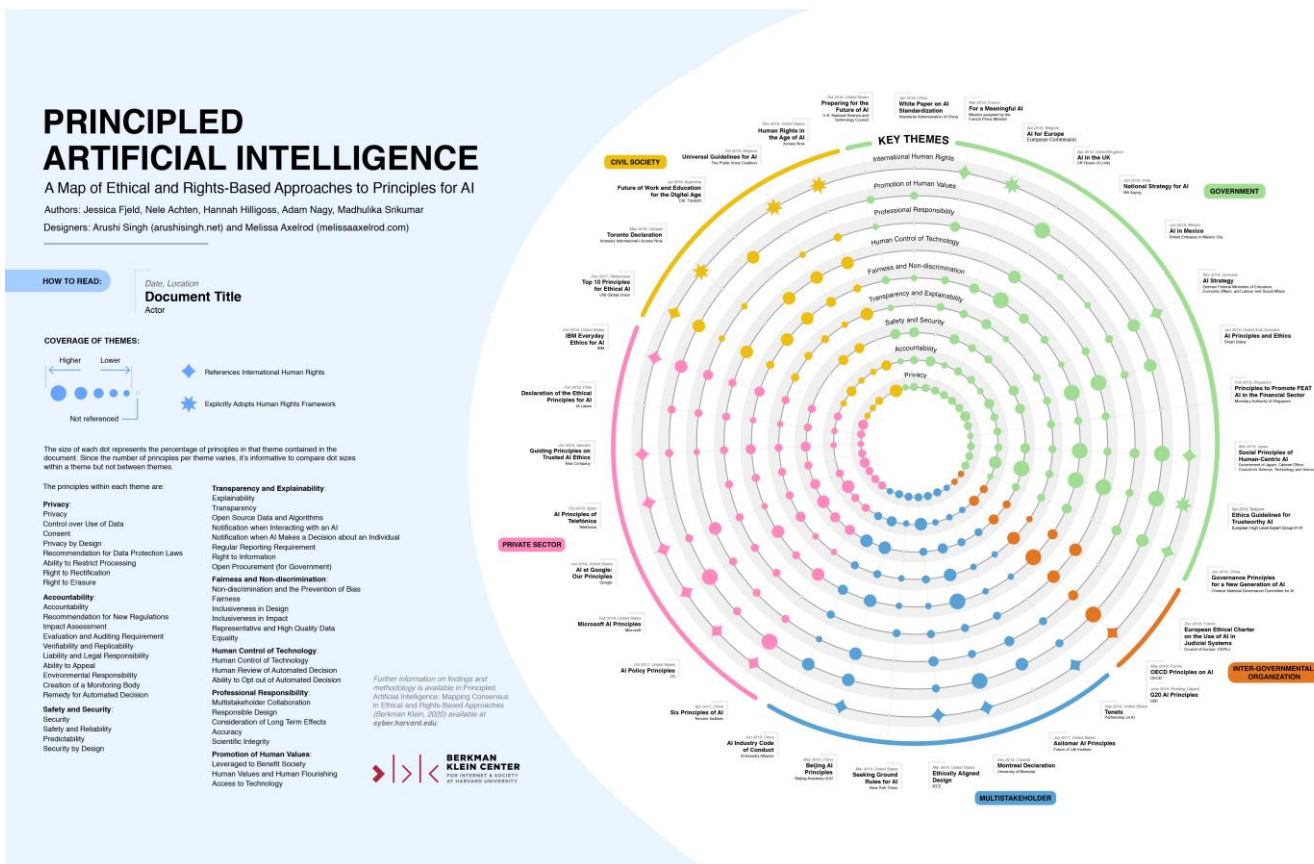


Zoals we weten, dekt het recht niet alles.

Ethiek laat ons verder kijken en opent bredere en specifiekere vragen.

Bovendien evolueren samenlevingen en verschillen regels en situaties. Hoe houden we rekening met al die diversiteit?

Er bestaan ontzettend veel regels en waarden

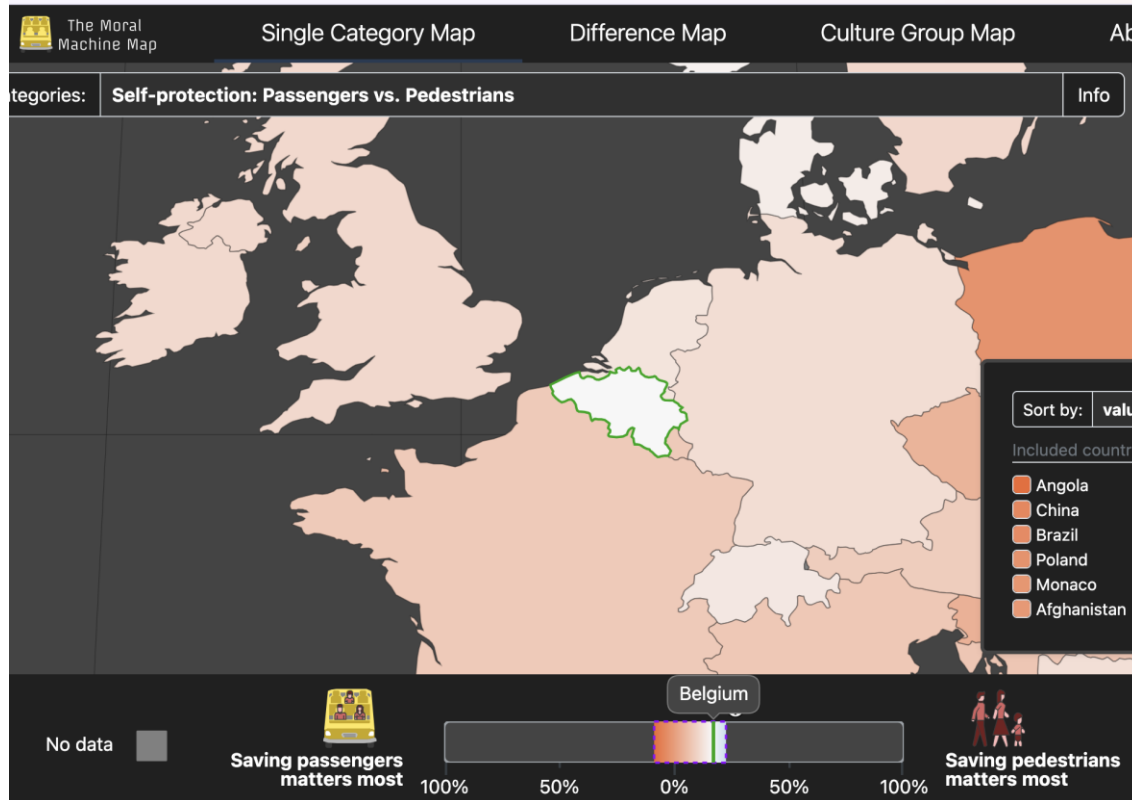


Er verschijnen honderden ethische verklaringen (naast wetten).

De uitdaging is enorm:

- Ethische principes en waarden worden anders verwoord naargelang wie ze opstelt
- Ze convergeren vaak, maar zijn nooit identiek

Universele waarden?



In de studie *Moral Machine* (Awad et al., 2018) werden bijna 40 miljoen beslissingen verzameld bij respondenten uit 233 landen en gebieden, om te onderzoeken hoe een zelfrijdend voertuig zich moet gedragen bij een dilemma zoals het “trolleyprobleem”.

De mate van consensus varieert vaak, bijvoorbeeld naargelang het beschouwde criterium (zoals de leeftijd of de sociale status van de betrokkenen).

AI aanpassen

Dat is de grote uitdaging voor onze samenlevingen: AI moet aangepast worden aan culturen, landen, praktijkdomeinen enz.



CENTRE DE RECHERCHE EN ETHIQUE

FUNDAMENTAL
ETHICS

ETHICS AND
POLITICS



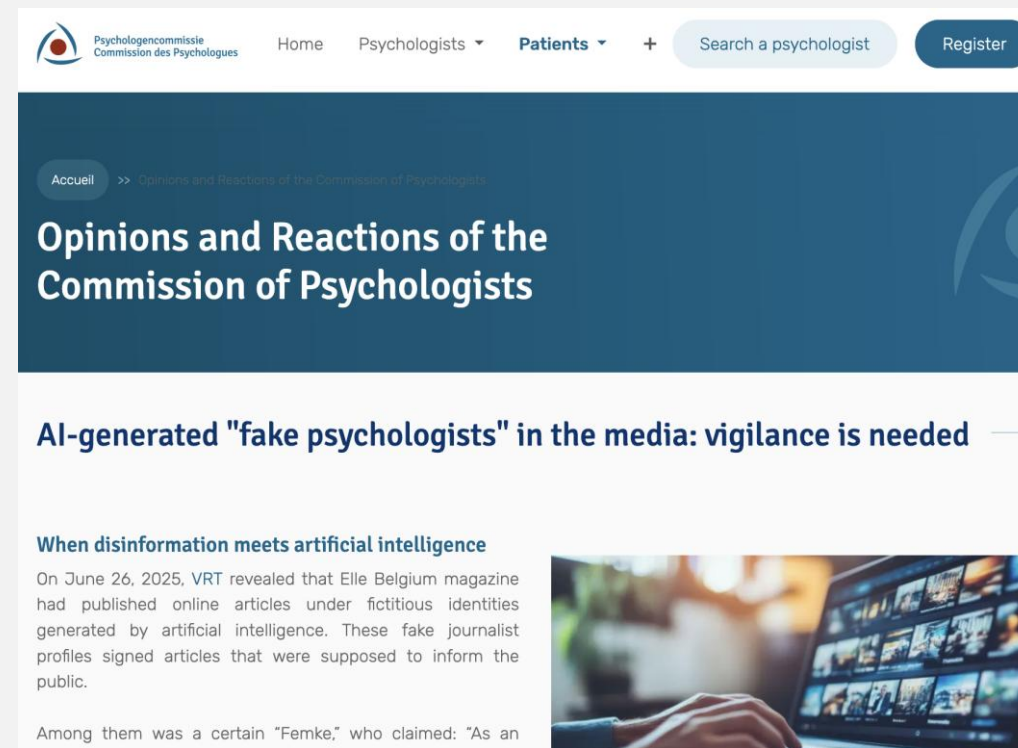
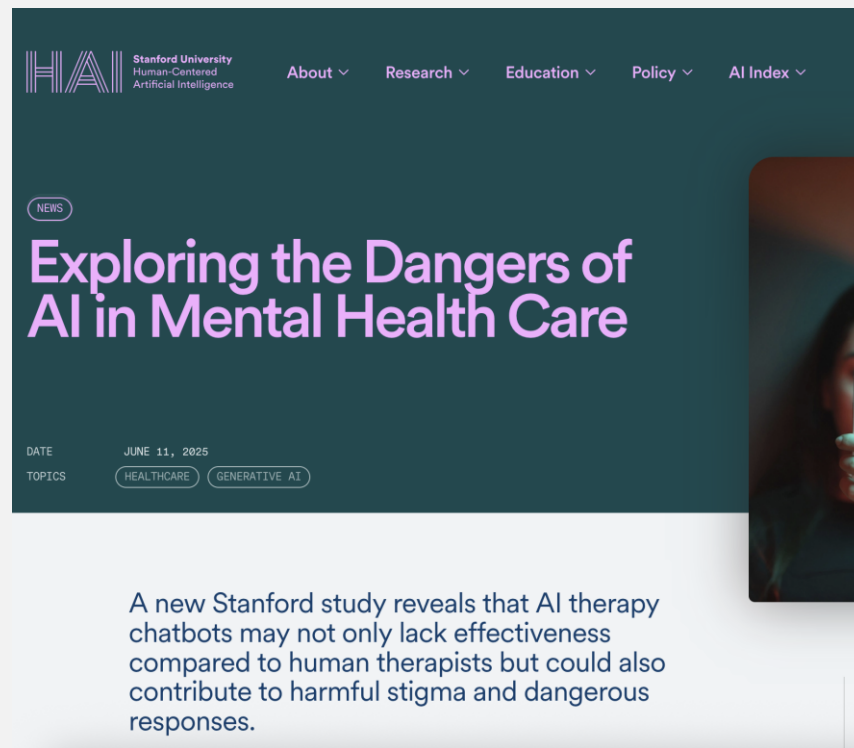
The Open Dialogue on
AI Ethics. Contribution
to the UNESCO
Recommendation on
the Ethics of Artificial

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.

Een voorbeeld: geestelijke gezondheid

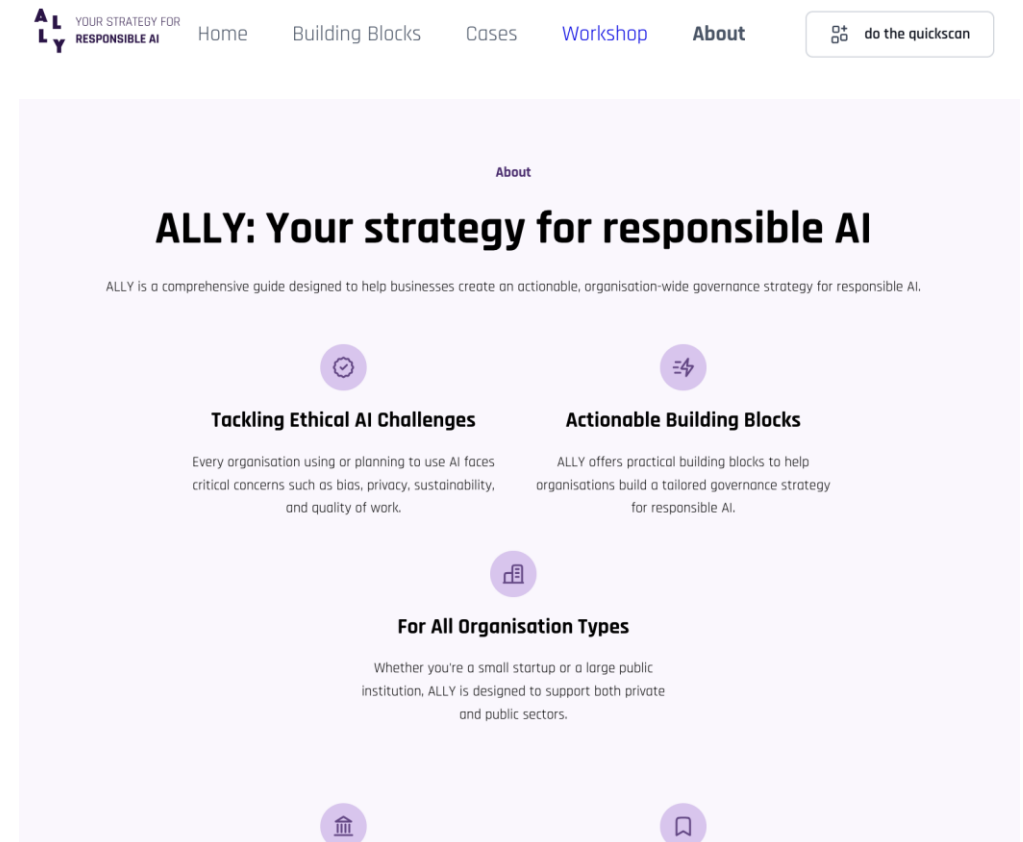


3. Programmeerbare ethiek

Welke vooruitzichten?

Van principes naar praktijk

- AI-ethiek blijft soms op een erg “hoog niveau” en laat zich niet eenvoudig vertalen naar de praktijk.
- Zo proberen tal van domeinen (vooral de gezondheidszorg) tools te ontwikkelen om ethiek te integreren.
- Op het terrein zien we nood aan “vertaling” en aan pedagogische “stap-voor-stap”-benaderingen.

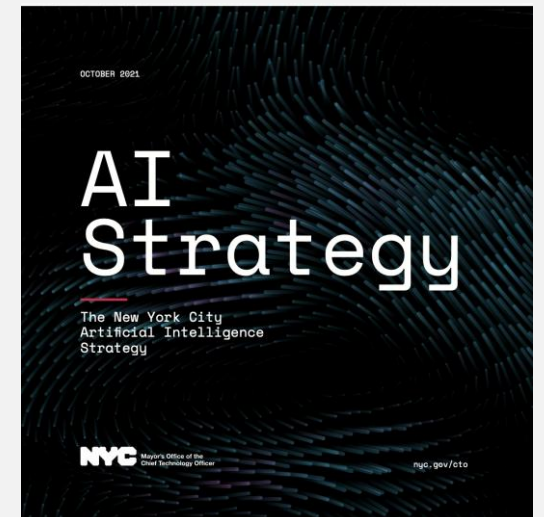


Hoe we “ethiek programmeren”

- We kunnen normen en standaarden gebruiken
 - Bv. ISO/IEC 42001: legt een auditeerbaar managementsysteem voor artificiële intelligentie (AIMS) vast, voor governance en risicobeheersing binnen organisaties.
- We kunnen ethische protocollen gebruiken
 - Bv. Z-Inspection
- We kunnen benaderingen gebruiken die zich richten op individuen of organisaties
 - Bv. AI Blindspot of Ally (KDM/FARI)

Strategieën en charters

- Om de grote principes te helpen “vertalen”, zien we complementaire benaderingen opduiken.
- Sommige organisaties stellen een charter op (zoals het charter dat we samen met AI4Belgium begeleidden voor de federale overheid).
- Andere (zoals veel steden) werken strategieën uit die de concepten toepasbaar maken.



Een voorbeeld: AI in het onderwijs

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

Recherche

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-TRIOMPH : QUAND L'INTELLIGENCE ARTIFICIELLE OPTIMISE LES ESSAIS THÉRAPEUTIQUES

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

Naar nieuwe beroepen?

AI Ethicists in Europe

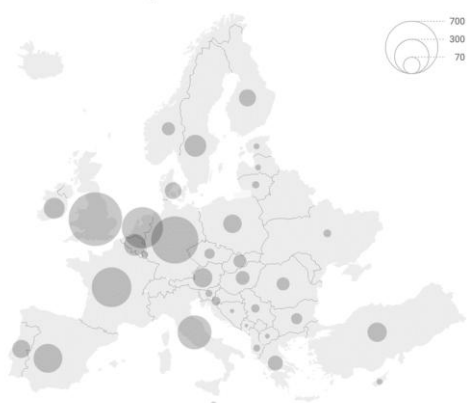


Figure 1 - Distribution of AI Ethicists on LinkedIn in Europe in April 2024

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

We zien een beroep ontstaan (AI-ethicus), en/of een titel die aan populariteit wint (en er ontstaan ook beroepsverenigingen).

Er is onmiskenbaar sprake van een juridisch vacuüm en van verwarring. *“Verschillende kritische stemmen hebben erop gewezen dat AI-ethiek nu al kwetsbaar is voor de greep van de industrie (Saltelli et al., 2022). Dat risico is bijzonder groot in praktische en professionele toepassingen. Om AI-ethici te beschermen, mogen organisaties hun rol niet herleiden tot loutere virtue signalling of ethics washing. AI-ethici moeten vrij zijn om kritische standpunten te uiten en mogen niet gedwongen worden de belangen van hun werkgever te verdedigen.”*

Een droom voor morgen

Welke vooruitzichten

Waarvoor moeten we ijveren?

- Voor meer burgerparticipatie bij het ontwerpen van de algoritmes van morgen
- Voor meer dialoog tussen wetenschappers, burgers, publieke en private actoren.



Een droom voor morgen

- Het wordt noodzakelijk (een mooie droom) om de algoritmes die ons besturen collectief te ontwerpen.
- Ruimtes creëren (en beschermen) waar die gesprekken kunnen plaatsvinden.



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

Écriture 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.

Burgers betrekken

- Met het project BrusselAir (EFRO) zetten we fora van “burgercoders” op, waarin we gebruikers betrekken bij het ontwerp van technologische oplossingen.
- In het verleden organiseerden we tal van burgerraadplegingen (over energie, over de stad en data).
- Ons doel: beter begrijpen hoe we het heft weer in handen kunnen nemen.



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

AUTEUR — Léa Rogliano, Ana Pop Stefanija

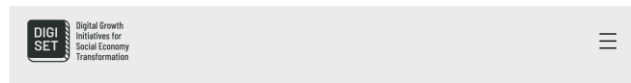
Quand les chercheurs et la société civile imaginent l'avenir de nos villes

En 2023, [FARI](#) – l'Institut de l'IA pour le Bien Commun, en collaboration avec [Paradiplom](#), 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

Werk aan de winkel (AI-educatie)



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.

Waar beginnen we?



Bedankt

Contact information



fari.brussels



info@fari.brussels



Cantersteen 16, 1000 Bruxelles
Kantersteen 16, 1000 Brussel

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

