

HIVER
2025
85

A3 MAGAZINE

Rayonnement du
Association des Anciens
et des Amis du
cnrs



L'Intelligence Artificielle dans notre quotidien, qu'en est-il ?

L'IA en action

Par Malik Ghallab

p.08

Percer les mystères de l'Univers à l'aide de l'IA

Par Dominique Fouchez

p.36

Intelligence Artificielle: une opportunité à saisir pour les seniors

Par Silvio Gianinazzi p.54

LE SOMMAIRE

Numéro 85
Hiver 2025

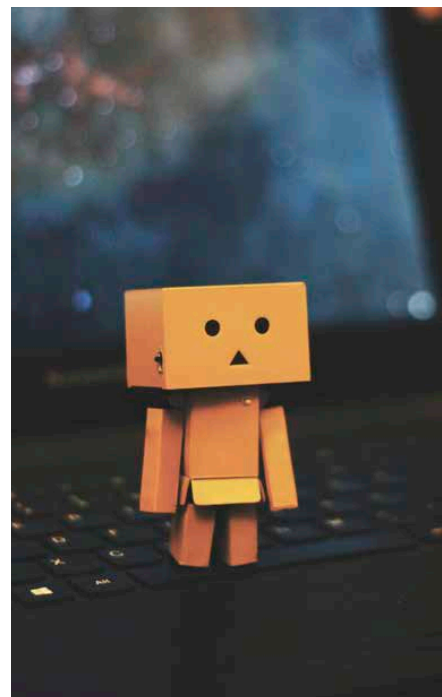
26

03 **LE MOT** de la rédaction

06 **AVANT-PROPOS**

Jean-Gabriel Ganascia

Les textes publiés dans l'A3 Magazine n'engagent que leurs auteurs comme seuls responsables de leur contenu.



10 **LES ENJEUX DE L'IA**

08 L'IA en action, *Malik Ghallab*

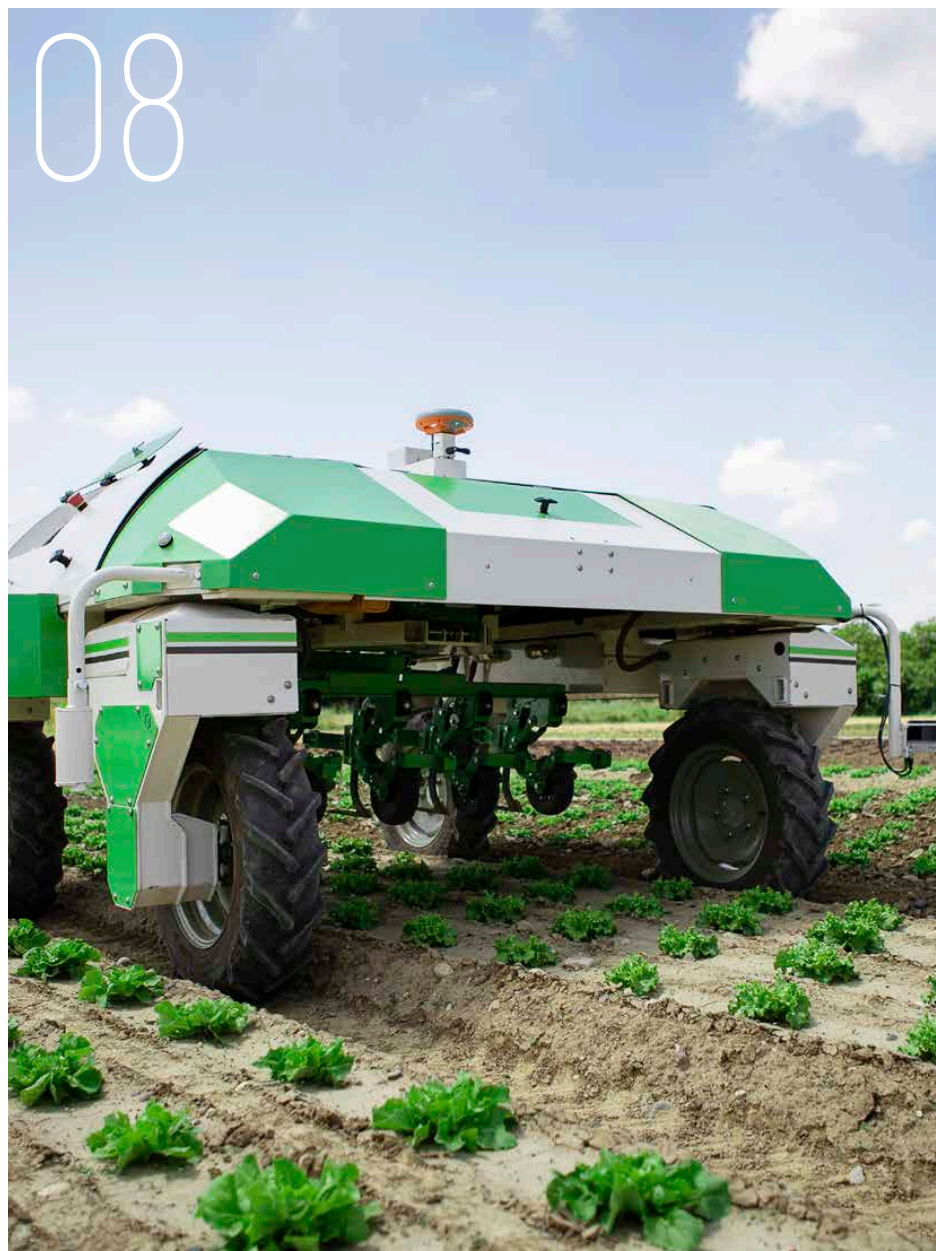
13 Pour une éthique by design de l'Intelligence Artificielle, *Vanessa Nurock*

18 IA, éducation et recherche appliquée : les actions de la DNE, *Élie Allouche*

26 IA conversationnelle : quelle place pour la France et la recherche publique ? *Thierry Poibeau*

29 L'intelligence artificielle en médecine : présent, promesses et précautions, *Jean-Patrick Lajonchère*

36 Percer les mystères de l'Univers à l'aide de l'IA, *Dominique Fouchez*



08

40 FOCUS

40 L'Institut DATAIA,
Sarah Cohen-Boulakia

43 IA, vie privée et régulation - les travaux de la CNIL, *Nicolas Berkouk*

47 L'intelligence artificielle et les sciences humaines et sociales, *Alexandre Gefen*

49 Quand l'Intelligence Artificielle explore la chimie du vin, *Tristan Pelluau, Théodore Tillement, Coraline Duroux, Augustin Tillement*

54 Intelligence Artificielle : une opportunité à saisir pour les seniors, *Silvio Gianinazzi*

56 EN SAVOIR PLUS

Des livres proposés sur le thème du dossier

☛ L'IA. expliquée aux humains, *Jean-Gabriel Ganascia*

☛ L'humain au risque de l'intelligence artificielle, *Pierre Rabhi, Juliette Duquesne*

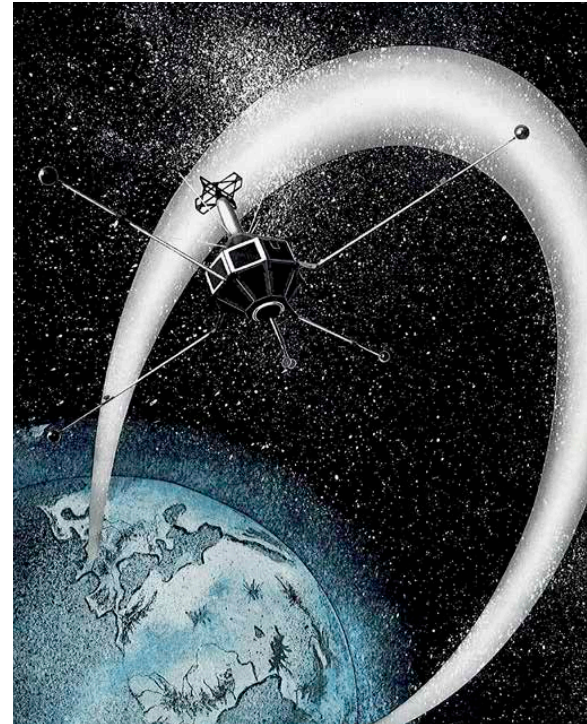
☛ Ni dieu ni IA. Une philosophie sceptique de l'intelligence artificielle, *Mathieu Corteel*

☛ IA : régression ou progrès ? *Revue #1257, Panthéon Sorbonne*

☛ Toxic Data. Comment les réseaux manipulent nos opinions, *David Chavalarias*

☛ Médecine et intelligence artificielle, *Bernard Nordlinger, Cédric Villani, Olivier de Fresnoye*

64



36

58 PASSIONS D'ADHÉRENT(E)S

☛ L'A3 au Musée de l'Air et de l'Espace du Bourget. *Hélène Kerec*

62 KIOSQUE

☛ Kolkhoze, *Emmanuel Carrère*

☛ Histoire(s) du Limousin - Faire Date, *Ouvrage collectif*

☛ Une science d'aventures. Les pérégrinations d'un cartographe au CNRS, *Michel Decobert*

64 RAYONNEMENT DU CNRS

☛ Le 60^e anniversaire de FR-1, le premier satellite scientifique : *Paul Gille*

☛ Entretien avec Céline Bellard, Prix A3 2025 : *Véronique Machelon*

☛ Le réseau Alumni du CNRS : *Daniel Charnay & Dominique Simon*



© D.R.

Pour une éthique *by design* de l'Intelligence Artificielle

Auteur

Vanessa Nurock

est Professeure de philosophie à l'Université Côte d'Azur affiliée au Centre de Recherches en Histoires des Idées (CRHI, UPR 4318, Université Côte d'Azur) et responsable de la Chaire UNESCO d'Éthique du Vivant et de l'Artificiel (EVA). Son travail se situe à l'interface des questions éthiques, politiques et scientifiques. Elle s'intéresse actuellement aux problèmes éthiques et politiques posés par l'intelligence artificielle et la robotique. Parmi ses publications *Quelle éthique pour les Nouvelles technologies?* (Vrin 2024) et *Care in an era of New Technologies and Artificial Intelligence* (Peeters 2025).



Résumé

Alors même que le développement de ces technologies soulève des enjeux éthiques essentiels, l'accélération des innovations en Intelligence Artificielle et en robotique s'accompagne trop souvent d'une forme de fatalisme technologique : tout se passe comme si l'avenir (considéré comme prometteur ou effrayant) de nos sociétés avec l'Intelligence Artificielle était déjà écrit et qu'il n'y avait pas d'alternative. Or, il est maintenant manifeste que les théories et procédures éthiques classiques apparaissent trop souvent insuffisantes pour traiter les problèmes éthiques inédits soulevés par ces technologies, ce qui vient renforcer ce fatalisme par l'impression que nous sommes condamnés à accepter passivement ce qui doit arriver sans même avoir les moyens de le réfléchir éthiquement. A rebours de cette dynamique fataliste, cet article défend la nécessité de renouveler la réflexion éthique et propose ainsi une approche processuelle permettant d'intégrer l'éthique à toutes les étapes du processus, selon une éthique « by design », dont les grandes lignes sont esquissées et mises en perspective.

●●●

Il n'est probablement pas exagéré de dire que notre monde est aujourd'hui marqué par l'omniprésence et l'ubiquité de l'Intelligence Artificielle (IA). L'IA s'est, en effet, insinuée jusque dans les moindres dimensions de notre vie, intime, privée, publique, professionnelle, individuelle ou collective. Du soin à l'éducation, des relations amoureuses à la finance et à l'art ou au politique, aucun espace humain ne semble plus lui échapper. Jamais, sans doute, une technologie n'a eu autant d'emprise sur nos vies. L'IA transforme ainsi maintenant en profondeur nos existences, comme nos sociétés, nos imaginaires comme notre réalité, jusqu'à modifier nos identités et nos relations.

Après avoir longtemps été considérée comme une question purement technologique, il apparaît maintenant avec une évidence frappante que l'IA est aussi une question sociale, éthique et politique. Pourtant, en France, dans le champ académique, la mobilisation des sciences humaines et sociales reste assez récente et se trouve malheureusement encore trop peu encouragée par les politiques scientifiques. On peut toutefois espérer un sursaut hexagonal, d'autant que l'éthique de l'IA est une question globale et qu'à l'échelle internationale, on observe un essor croissant des recherches en sciences humaines et sociales sur l'IA ainsi que la formation de centres de recherches mettant ces questions au cœur de leurs projets, ou se dotant d'équipes interdisciplinaires dédiées. De plus, après des années de développement échappant à tout outil normatif, voire parfois à toute régulation concernant l'IA, nous assistons depuis 2021 à un moment charnière, notamment depuis, sur le plan législatif, l'IA Act Européen et sur le plan éthique, la Recommandation sur l'Éthique de l'IA de l'UNESCO.

Il importe, bien sûr, de manière préalable, de veiller à distinguer ces deux plans complémentaires. D'une part, l'éthique renvoie à une réflexion normative sur ce qui devrait être, tandis que la régulation concerne l'ensemble des dispositifs juridiques et institutionnels visant à encadrer les usages de l'IA. Ces deux dimensions ne se recouvrent pas : une régulation peut être insuffisante ou inadéquate au regard des exigences éthiques ; une réflexion éthique ne peut (et ne doit) pas forcément toujours trouver de traduction réglementaire. Par ailleurs, il importe de distinguer l'acceptation sociale de l'acceptabilité morale. Alors que l'acceptation sociale mesure le degré d'adhésion ou de tolérance (parfois sous l'influence d'incitations ou de marketing) des individus face aux innovations technologiques, l'acceptabilité morale, quant à elle, interroge la légitimité éthique de ces innovations. Ce n'est pas parce qu'une technologie existe ou qu'elle est mise en place qu'elle est nécessairement éthique. Ainsi, la reconnaissance faciale, d'abord très largement mise en place aux Etats-Unis, a connu, il y a une dizaine d'années une remise en question éthique (notamment liée à la vie privée et au risque d'une surveillance généralisée) qui a conduit à l'instauration de cadres juridiques.

Cependant, face à la rapidité des évolutions technologiques, la plupart des personnes oscillent entre un sentiment de révolte et une forme de soumission fataliste, comme si l'avenir de l'IA et de la robotique était déjà écrit. Dans

ce contexte, les théories éthiques traditionnelles – qu'elles soient déontologiques, utilitaristes ou fondées sur la vertu – peinent à fournir des outils conceptuels capables de dénouer les problèmes éthiques souvent inédits que posent ces technologies. Face à ces défis, cet article propose d'expliquer pourquoi la réflexion éthique sur l'IA doit se renouveler en proposant une éthique « *by design* ».

L'ÉTHIQUE AU-DELÀ DES LISTES ET DES CHECKLIST (LISTES À COCHER) : ASSUMER L'INCERTITUDE ÉTHIQUE

Certes, l'IA soulève une série de problèmes éthiques pouvant sembler bien identifiés, dont la liste peut sembler aujourd'hui assez classique. On peut en dégager au moins cinq principaux.

» **Le premier concerne les biais et discriminations algorithmiques.** Parce que les systèmes d'IA apprennent à partir de données humaines, ils reproduisent et amplifient les préjugés présents dans ces corpus. Par exemple, certains moteurs de recherche ou systèmes de recrutement automatisés ont montré des biais de genre, associant des métiers techniques aux hommes et des métiers de soin aux femmes.

» **Le deuxième concerne la vie privée et la surveillance.** Les systèmes d'IA reposent sur la collecte massive de données personnelles, souvent sans consentement (tout court ou consentement éclairé puisqu'on ne sait pas à quoi servent ces données). Par exemple, les assistants vocaux enregistrent des fragments de conversations privées, tandis que les caméras dotées de reconnaissance faciale installées dans les espaces publics ouvrent la voie à une surveillance généralisée et, en l'absence de cadres, il ne faut parfois qu'un pas supplémentaire pour qu'elles deviennent un instrument de contrôle social.

» **Troisièmement, certains usages des IA peuvent conduire à limiter la liberté des personnes** en les influençant subrepticement. Par exemple, les algorithmes de recommandation sur les réseaux sociaux modèlent les comportements en orientant les choix de consommation ou d'information, favorisant parfois la polarisation politique et la diffusion de fausses nouvelles. Dans le domaine judiciaire, les algorithmes de justice prédictive influencent les décisions judiciaires, réduisant plus ou moins considérablement la marge de jugement humain et risquant d'installer une dépendance excessive aux calculs probabilistes.

« Ce n'est pas parce qu'une technologie existe ou qu'elle est mise en place qu'elle est nécessairement éthique. »

Quatrièmement, certains usages des IA sont susceptibles de miner la responsabilité. Par exemple, les drones militaires capables de sélectionner des cibles sans intervention humaine déplacent la décision de vie ou de mort vers une machine.

Cinquièmement, certains usages des IA sont susceptibles de modifier en profondeur nos relations et de créer des formes d'attachement inquiétantes, qui rendent les utilisateurs à la fois dépendants et de moins en moins capables de tisser des relations avec des personnes humaines réelles, que ce soit sur le plan amical, amoureux ou tout simplement social, ainsi que l'a mis en lumière le récent rapport du Comité d'Éthique du CNRS (COMETS) piloté par la roboticienne Catherine Pelachaud, intitulé *Le phénomène d'attachement aux robots dits « sociaux »*. Pour une vigilance de la recherche scientifique (1).

Si ces cinq problèmes éthiques sont fondamentaux, on peut toutefois considérer qu'ils sont souvent difficiles à anticiper dans le développement des IA dans la mesure où, par essence, il est difficile, pour ne pas dire impossible, de prévoir comment vont évoluer les IA et quels seront leurs impacts ; de sorte que dresser ce type de liste n'est probablement pas suffisant pour prendre en charge ces enjeux éthiques, encore faut-il réfléchir à la manière de le faire. En effet, ainsi que l'a bien montré J-P Dupuy (2), l'incertitude fait partie du « programme métaphysique » des technologies NBIC (Nanotechnologies, Biotechnologies, Sciences de l'Information et de la Cognition), dont fait partie l'IA. Plus encore, avec les développements récents de l'apprentissage profond et des IA génératives, cette dimension d'incertitude est devenue particulièrement patente.

Comment gérer cette incertitude ? Il apparaît d'emblée que la demande souvent faite aux éthiciens de produire des listes à cocher (*checklists*) n'est absolument pas pertinente car elle procède d'un double déni. Le premier déni est celui de cette incertitude. Le deuxième déni est celui de la signification profonde de l'éthique, qui ne vise pas tant à nous donner une liste de règles à suivre qu'à nous stimuler à aborder ces problèmes de manière réflexive. Ainsi que le souligne le rapport 2023 de l'Observatoire International sur les Impacts Sociétaux de l'IA et du Numérique (OBVIA) Québécois intitulé *L'éthique au cœur de l'IA* (3), « l'éthique est une réflexion argumentée qui est soumise à des procédures de raisonnement rationnel. C'est quand on s'interroge sur les fondements rationnels d'une règle morale (qu'elle soit religieuse ou non), sur ses conséquences logiques, sur sa place dans un système de règles cohérent qu'on entre dans le domaine de l'éthique comme discipline philosophique. » Toutefois, si l'éthique de l'IA ne saurait consister en une suite de *checklists*, comment faire pour assumer éthiquement cette incertitude et prendre en charge les enjeux éthiques du développement croissant et rapide des IA ?

AU-DELÀ DU FATALISME FACE AU ENJEUX ÉTHIQUES DE L'IA

L'un des problèmes que pose cette incertitude réside dans le sentiment de fatalité qu'elle suscite chez les individus et les collectifs. Cette impression que les trajectoires technologiques sont inéluctables, qu'elles échappent à toute maîtrise, conduit

Il est difficile, pour ne pas dire impossible, de prévoir comment vont évoluer les IA et quels seront leurs impacts.”

souvent à une résignation anticipée où il ne reste aux citoyens qu'à se soumettre ou à s'opposer. Si bien que les discussions publiques oscillent trop souvent entre l'enthousiasme excessif (*hype*) et le rejet radical (*hate*), qui semblent, dans les deux cas, prédire comme inéluctable, le remplacement de l'humanité par des IA : que ce remplacement soit perçu comme une menace ou une promesse, notre avenir semble ainsi fermé à toute autre possibilité.

À cette réduction des possibles s'ajoute ce que l'on a coutume d'appeler le « dilemme de Collingridge », qui illustre la difficulté de réguler une innovation au moment où ses effets sont encore incertains, et l'impossibilité de la modifier une fois ses impacts connus (4). Les deux branches du dilemme peuvent être ainsi formulées : au commencement d'une technologie, on ne peut en prédire les conséquences ; on est donc dans une situation de déficit des connaissances. Vers la fin, quand une technologie est avancée et intégrée dans la société, il est plus difficile, voire quasi impossible, de revenir dessus pour la modifier et transformer son intégration sociale ; c'est l'action et non plus la connaissance qui est déficitaire ici. Comme dans tout dilemme, chacune des deux branches est insatisfaisante : au début, on ne peut pas savoir même si on peut agir ; à la fin, on peut savoir mais on ne peut plus agir. Les algorithmes de recommandation déployés par les réseaux sociaux en sont une illustration frappante : introduits sans réelle anticipation possible de leurs effets tant ils pouvaient sembler anodins, ils ont eu des conséquences massives sur la polarisation politique et la désinformation. Aujourd'hui, il est extrêmement difficile de revenir en arrière ou de transformer en profondeur ces systèmes, tant ils sont intégrés dans les pratiques quotidiennes et les modèles économiques.

Or, ce dilemme de Collingridge donne donc l'impression qu'il n'y a pas d'alternative ; il nourrit l'idée que toute tentative de régulation est vouée à l'échec et restreint drastiquement l'éventail des possibles en renforçant le sentiment que notre trajectoire technologique, tout autant que sociale avec l'IA, est déjà tracée. Ainsi, on peut suggérer que les enjeux éthiques des IA ne résident pas seulement dans leurs développements et usages concrets, mais aussi dans la construction discursive et sociale d'un horizon technologique perçu comme inévitable. Ce fatalisme fragilise la capacité des sociétés à exercer un jugement moral autonome, à imaginer des trajectoires alternatives et à maintenir une pluralité de choix face à l'innovation.



Paysage surréaliste avec couple,
Felix Nussbaum c.1939



© MUSEUMSQUARTIER OSNABRÜCK

CATASTROPHE ÉCLAIRÉ OU OPTIMISME RAISONNÉ VIS-À-VIS DE L'IA ?

Toutefois, ce dilemme de Collingridge pourrait également susciter la réaction opposée en soulignant l'importance de ne pas attendre que tout ce qui est possible techniquement soit accompli pour poser sérieusement la question de l'éthique de l'IA et de la robotique tout en assumant le fait de ne pas pouvoir en prédire toutes les conséquences. C'est notamment ce que propose J-P Dupuy (5) à travers la notion de « catastrophisme éclairé » – qui concernait à l'époque, davantage les questions environnementale et nucléaire mais peut sans aucun doute être élargie à l'IA. Dupuy préconise de renverser la logique habituelle selon laquelle nous ne croyons aux choses épouvantables qu'une fois qu'elles sont advenues. Au contraire, dit-il, il faut y croire avant, tout faire comme si elles allaient arriver, précisément pour les empêcher d'arriver. Cette attitude se distingue donc du catastrophisme fataliste ou paralysant : il ne s'agit pas d'abandonner la partie mais plutôt de se servir de la catastrophe comme d'un levier éthique et politique paradoxal.

Cependant, cette posture oriente l'action en fonction de la catastrophe à venir, dont rien ne prouve cependant qu'elle soit certaine (ni qu'elle soit pour bientôt). Pour permettre la construction d'un futur alternatif, dans le catastrophisme éclairé, on agit en concentrant au moins temporairement l'énergie sur le fait qu'il n'y a pas d'alternative et l'on ne l'utilise donc pas pour explorer une pluralité de futurs possibles. Or, comme l'écrivent Catherine et Raphaël Larrère (6), « le pire n'est pas certain ». Si la posture prospective est certes une manière constructive de répondre à l'ignorance, sa tonalité catastrophiste ne va cependant pas de soi. A l'opposé du catastrophisme éclairé qui propose de construire une image de l'avenir suffisamment effrayante pour qu'elle serve de repoussoir, l'optimisme raisonné suggère de ne pas nier les risques mais d'essayer de construire une image de l'avenir suffisamment attrayante que l'on puisse désirer ardemment la réaliser. Il ne s'agit pas de nier les risques mais de ne pas les laisser nous sidérer ni laisser leur prévention orienter toute notre énergie. Il s'agit donc, dans une perspective existentialiste, d'inverser l'ordre de la construction de notre essence. Elle n'est pas définie d'emblée au sens où, selon l'adage existentialiste, « l'existence précède l'essence ». Il importe de ne pas se laisser enfermer dans l'individualisme inséparable des positions fatalistes, pour promouvoir plutôt une approche pluraliste renouvelant les liens entre éthique et politique.

L'optimisme raisonné rejoint ainsi l'éthique du care, conception développée depuis la fin du 20^e siècle. Elle propose d'élargir le champ de l'éthique au-delà de la Justice, du bien et

du mal pour intégrer notamment des questions comme celle de savoir ce qui nous importe (*What do we care about?*) ou celles de savoir qui prend soin (*take care*) de qui et comment, selon une dynamique processuelle qui intègre tant la perspective des pourvoyeurs et pourvoyeuses que des récipiendaires du soin. Cette approche féministe remet également en question l'enfermement dans des approches binaires (comme active/passif, intime/collectif etc.) et l'extrême focalisation sur la question des risques telle qu'on en perd de vue ce qui nous importe. En ce sens, comme l'écrit JC Tronto (7), il importe de ne pas se focaliser uniquement sur la logique du risque, mais d'intégrer également une logique du care. Ce faisant, nous pouvons assumer notre ignorance et l'imprévisibilité de l'IA, le fait que nous ne pouvons pas avoir une maîtrise absolue de l'IA, sans pour autant être totalement insouciants et irresponsables.

POUR UNE ÉTHIQUE BY DESIGN DE L'IA

Pour autant, même si nous refusons de nous laisser paralyser par le catastrophisme prophétique ou de mobiliser la majeure partie de notre énergie présente pour penser la catastrophe qui s'en vient, au profit d'un optimisme raisonné qui tente d'explorer les possibles, il n'en demeure pas moins que les grandes théories éthiques (ascendantes et descendantes) sont assez démunies face

au dilemme de Collingridge et au polymorphisme de l'IA. En effet, les théories dites ascendantes, de type déontologique, qui proposent l'application de grands principes vers la pratique sont parfois trop générales et peuvent être interprétées de diverses manières. Les théories dites descendantes, de type conséquentialiste, qui évoluent une action ou une technologie à l'aune des conséquences, se heurtent

“ L'éthique *by design* propose de situer l'éthique non seulement en amont (au niveau des principes) ou en aval (au niveau des conséquences) mais à toutes les étapes du processus. ”

au problème de l'imprévisibilité de l'IA et au dilemme de Collingridge puisque, une fois la technologie développée, il est difficile de la modifier (même si l'on peut certes la condamner moralement). De plus, les approches conséquentialistes articulées à une logique coûts/bénéfices sont également problématiques car elles opèrent comme si les enjeux éthiques se ramenaient uniquement à un choix entre, d'un côté, la promesse d'efficacité et de rentabilité et, de l'autre, la crainte des pertes sociales ou économiques. Ainsi, les systèmes IA de diagnostic médical automatisé sont parfois présentés comme une solution pour

réduire les dépenses de santé et accélérer les soins, tandis que les risques liés aux biais algorithmiques ou à la disparition de la relation humaine sont relégués au second plan. De même, les véhicules dits autonomes sont évalués presque exclusivement à l'aune du nombre d'accidents évités ou des emplois supprimés, sans que soient véritablement interrogées les implications éthiques d'abandonner des décisions vitales à des machines voire d'abandonner l'éthique à des IA.

C'est pourquoi il importe de compléter ces approches, lorsqu'elles ne s'avèrent pas totalement satisfaisantes, avec une éthique « *by design* » de l'IA. Qu'entend-on par là ? Ainsi qu'elle a été notamment proposée par Nurock et al. (8) et Brey et Dainow (9), l'éthique *by design* propose de situer l'éthique non seulement en amont (au niveau des principes, comme les approches descendantes) ou en aval (au niveau des conséquences, comme les approches descendantes) mais à toutes les étapes du processus. Selon Nurock et al., cette démarche s'apparente à une forme d'équilibre réfléchi inspiré du philosophe américain John Rawls, qui va des principes aux pratiques et des pratiques aux principes selon un mouvement d'aller-retour qui se stabilise lorsque l'on estime avoir suffisamment affiné les principes, mais qui est évidemment susceptible d'être modifié lorsque cela devient nécessaire. Elle s'articule également aux éthiques du *care* en divers aspects : d'une part, elle pose en son cœur la question de savoir ce qui nous importe (*what do we care about ?*) et la question de la prise en compte des plus vulnérables (*care about*) et du soin des institutions - *le care with* (10). D'autre part, c'est une approche relationnelle qui pose comme l'un de ses critères (en harmonie avec les recommandations du rapport de 2024 du COMETS du CNRS), le fait que les IA ne doivent pas être développées de sorte à remplacer les relations humaines, réciproques et empathiques, par de simples connexions (11) qui seraient capables de les feindre.

Ceci étant, l'éthique *by design* rencontre encore un certain nombre d'obstacles à surmonter. Sur le plan théorique, l'une des difficultés de l'éthique *by design* est de parvenir à intégrer dans sa dimension relationnelle, la variété culturelle de nos rapports aux technologies en général et aux IA ou robots en particulier. Cette variété est notamment marquée par une scission entre l'Occident, où l'on considère souvent que les IA vont remplacer les humains (qu'on s'en réjouisse ou s'en inquiète), et certaines cultures orientales, comme au Japon, où l'on envisage davantage une coopération et posent davantage la question en termes d'harmonie (12), qui pourrait entrer en complémentarité avec la notion de *care*. Sur le plan pratique, au niveau des laboratoires, l'éthique *by design* nécessite la possibilité d'équipes interdisciplinaires travaillant sur le temps long en coopération étroite, sans pour autant que l'éthique ne soit seulement considérée comme un appendice du projet; sans installer de rapport hiérarchique entre la dimension technologique et la dimension éthique, afin de permettre un authentique équilibre réfléchi entre théorie et pratique. L'un des enjeux de l'éthique *by design* est ainsi de pouvoir être suffisamment globale au niveau culturelle et suffisamment locale au niveau des pratiques pour permettre une réelle prise en charge des enjeux éthiques des IA. C'est certes un défi, mais un défi qu'il est urgent et nécessaire de relever aujourd'hui plus que jamais ! ■

Biblio

- (1) **COMETS (2024)** *Le phénomène d'attachement aux robots dits "sociaux". Pour une vigilance de la recherche scientifique.* Comité d'éthique du CNRS. <https://comite-ethique.cnrs.fr/avis-du-comets-le-phenomene-dattachement-aux-robots-dits-sociaux-pour-une-vigilance-de-la-recherche-scientifique/>
- (2) **Dupuy J-P (2007)** Some pitfalls in the philosophical foundations of nanoethics. *The Journal of Medicine and Philosophy* 32 : 237 261.
- (3) **Langlois L et al. (2023)** *L'éthique au cœur de l'IA.* Rapport de l'Observatoire international sur les impacts sociétaux de l'IA et du numérique (OBVIA). <https://www.obvia.ca/ressources/lethique-au-coeur-de-lia>
- (4) **Collingridge D (1980)** *The social control of technology.* Palgrave Macmillan, UK, 200 pages.
- (5) **Dupuy J-P (2002)** *Pour un catastrophisme éclairé. Quand l'impossible est certain.* Seuil, Paris, 216 pages.
- (6) **Larrère C, Larrère R (2020)** *Le pire n'est pas certain. Essai sur l'aveuglement catastrophiste.* Premier Parallèle, Paris, 208 pages.
- (7) **Tronto JC (2013)** *Caring democracy: Markets, equality, and justice.* New York University Press New York, NY, 256 pages.
- (8) **Nurock V, Chatila R, Parizeau M-H (2021)** What does "ethical by design" mean? In : Braunschweig B, Ghallab M (eds), *Reflections on Artificial Intelligence for Humanity.* Springer, Cham, pages 171-190
- (9) **Brey P, Dainow B (2023)** Ethics by design for AI (EbD-AI). *AI and Ethics* 3 : 1 15.
- (10) **Tronto JC (2012)** *Le risque ou le care ?* (F. Brugère, Trad.). Presses Universitaires de France, Paris, 64 pages.
- (11) **Nurock V (2024)** *Quelle éthique pour les nouvelles technologies ?* Vrin, Paris, 282 pages.
- (12) **Berberich N, Nishida T, Suzuki S (2020)** Harmonizing artificial intelligence for social good. *Philosophy & Technology* 33 : 613 638.