

# « Éthique de l'intelligence artificielle » comme signifiant flottant : considérations théoriques et analyse critique des discours de presse

« Ethics of artificial intelligence » as a floating signifier: theoretical considerations and critical analysis of press discourses

< **Alberto ROMELE** <sup>1</sup> >

1. *Université Sorbonne Nouvelle, laboratoire IRMÉCCEN*  
*alberto.romele@sorbonne-nouvelle.fr*

DOI : [10.25965/interfaces-numeriques.5229](https://doi.org/10.25965/interfaces-numeriques.5229)

## < RÉSUMÉ >

Dans cet article, nous explorons l'utilisation de l'expression « éthique de l'IA ». Nous avançons l'hypothèse que « éthique de l'IA » constitue un signifiant flottant, une notion empruntée à Ernesto Laclau. L'article se divise en deux parties. Dans la première, nous soulignons la nécessité d'intégrer une réflexion sur l'éthique de la communication scientifique concernant l'IA à l'éthique de l'IA elle-même. Nous introduisons également le concept de signifiant flottant, illustré par ses applications récentes dans le domaine des études médiatiques. Dans la seconde partie, nous proposons une analyse critique des discours qui mobilisent l'expression « éthique de l'IA » dans la presse quotidienne. Nous examinons des articles parus dans huit journaux de quatre pays européens (France, Italie, Royaume-Uni, et Allemagne) sur une période de trois mois (de mai à août 2023). Nous identifions trois usages discursifs de l'expression « éthique de l'IA » : la normativité des institutions, la critique des chercheurs et chercheuses, et le techno-solutionnisme des entreprises.

## < MOTS-CLÉS >

éthique de l'IA, signifiants flottants, lutte pour l'hégémonie, discours de presse

**< ABSTRACT >**

In this article, we explore the use of the term «AI ethics». We propose that «AI ethics» is a floating signifier, a concept borrowed from Ernesto Laclau. The article is structured in two parts. In the first part, we emphasize the need to include reflections on the ethics of scientific communication regarding AI within AI ethics itself. We also introduce the concept of a floating signifier, illustrated by its recent applications in the field of media studies. In the second part, we offer a critical analysis of discourses that mobilize the term «AI ethics» in the daily press. We examine articles published in eight newspapers from four European countries (France, Italy, the United Kingdom, and Germany) over three months (May to August 2023). We identify three discursive uses of «AI ethics»: the normativity of institutions, the critique from researchers, and the techno-solutionism of corporations.

**< KEYWORDS >**

AI ethics, floating signifiers, struggle for hegemony, press discourses

---

**Introduction**

L'éthique de l'intelligence artificielle (IA) suscite actuellement de nombreux discours. L'objectif de cet article n'est pas tant de se pencher sur le contenu de ces discours, mais plutôt sur leur raison d'être profonde. Notre hypothèse est que derrière cette diversité de discours se trouve une lutte symbolique pour l'hégémonie dans le domaine de l'IA. Autrement dit, nous soutenons que l'expression «éthique de l'IA» fonctionne comme un signifiant flottant. Le concept de signifiant flottant, introduit par Laclau (2008), désigne des termes et concepts suffisamment polysémiques pour être interprétés, compris et utilisés stratégiquement de différentes manières par divers groupes sociaux à des fins hégémoniques.

L'article est structuré en deux parties. La première aborde la nécessité d'intégrer à l'éthique de l'IA une réflexion sur l'éthique de la communication scientifique concernant l'IA. Nous y présentons également le concept de signifiant flottant et une de ses applications dans le domaine des études médiatiques. La seconde partie propose une analyse critique des discours utilisant «éthique de l'IA» dans la presse quotidienne, considérée ici comme une arène de débat et de lutte entre des prétentions hégémoniques rivales. Nous analysons les articles parus dans huit quotidiens de quatre pays européens (France, Italie, Royaume-

Uni et Allemagne) sur une période de trois mois (entre mai et août 2023). Ces journaux ont été sélectionnés pour leur popularité et leur réputation. Une période de trois mois a été choisie pour obtenir un corpus représentatif, notamment parce que le débat sur l'IA et ses implications éthiques était particulièrement animé, en raison de discussions au niveau européen autour de l'IA Act, proposé en avril 2021 et approuvé dans sa version finale en mars 2024, ainsi que de la grève de la Writers Guild of America de mars à septembre 2023, centrée notamment sur l'utilisation de systèmes d'IA générative dans la création de scénarios. Nous mettrons en lumière trois usages discursifs de l'expression « éthique de l'IA » : la normativité des institutions, la critique des chercheurs et des chercheuses, et le techno-solutionnisme des entreprises. En conclusion, tout en reconnaissant la valeur des travaux théoriques et empiriques sur l'éthique de l'IA, nous plaiderons pour une politique des éthiques de l'IA.

### 1. « Éthique de l'IA », en théorie

Il pourrait sembler que l'analyse des discours sur l'éthique de l'IA n'ait que peu de rapport avec l'éthique de l'IA elle-même. Au mieux, cela pourrait être considéré comme relevant de l'éthique de la communication, appliquée spécifiquement à la communication scientifique sur l'IA – pour plus sur l'éthique de la communication scientifique, voir Priest, Goodwin & Dahlstrom (2018). Cependant, nous soutenons que cet oubli de la communication autour de l'IA, fréquente chez les philosophes et les éthiciens de l'IA, est le résultat d'un biais empiriste.

Avec le tournant empirique (Brey, 2010), les philosophes et éthiciens des techniques ont orienté leur attention vers les structures et les itérations matérielles des technologies, ainsi que sur les activités réelles menées dans les laboratoires de recherche et développement<sup>1</sup>. Pour eux,

---

<sup>1</sup> Parler de philosophes et d'éthiciens des techniques en termes généraux peut sembler imprécis. Cependant, le tournant empirique se manifeste principalement dans deux courants de la philosophie des techniques contemporaine : la post-phénoménologie, représentée par des auteurs comme Don Ihde et Peter-Paul Verbeek, et le programme sur la « nature dualiste des artefacts techniques », associé à des penseurs tels que Peter Kroes et Anthonie Meijers. En général, ces deux programmes se distancient du pessimisme et du

les questions éthiques ne sont jamais générales ou transcendantes, mais dépendent directement de la matérialité de l'objet technique concerné. Toute solution éthiquement appropriée résulte donc d'un réagencement des composants matériels de l'objet. À titre d'exemple, l'éthique d'un algorithme d'IA dépendra de sa robustesse, de sa transparence, de sa fiabilité à produire des résultats exempts de biais sociaux et culturels, etc. De là aussi découle un certain engouement pour le *nudge* et l'*ethics by design* en tant que pratiques éthiques. Toutefois, ces approches pourraient être critiquées pour la conception de l'humain qu'elles véhiculent, souvent vue comme dénué de capacité de réflexivité – pour une critique de ces deux approches, voir Floridi (2016).

En cherchant à se rapprocher des « choses mêmes », les philosophes et éthiciens des techniques ont parfois négligé les contextes dans lesquels ces choses sont toujours-déjà insérées. Ces contextes peuvent être de nature technologique, économique, documentaire, socioculturelle, entre autres<sup>2</sup>. Chacune de ces conditions de possibilité porte en elle des implications éthiques qui, d'un point de vue empiriste, risquent d'être sous-estimées. Dans le domaine de l'IA, il s'avère que ce sont souvent les entreprises qui, grâce à leur puissance économique et leur expertise technique, produisent les algorithmes matériellement les plus robustes, mais qui les utilisent ensuite de la manière la moins éthique possible. Autrement dit, l'éthique de l'IA ne dépend pas uniquement du *design* de ces systèmes, mais également de leurs contextes potentiels d'utilisation et de compréhension.

Les philosophes et éthiciens de l'IA, à quelques exceptions près – par exemple Coeckelbergh (2021) –, ne prêtent généralement pas beaucoup d'attention aux discours et à la communication autour de l'IA. En revanche, nous soutenons que les discours et la communication sur l'IA jouent un rôle fondamental dans les processus d'innovation dans ce

---

déterminisme caractéristiques d'auteurs classiques tels que Martin Heidegger et Jacques Ellul. Pour une discussion critique sur le tournant empirique et ses représentants, voir Romele (2024), notamment les chapitres 2 et 4.

<sup>2</sup> À ce propos, Smith (2018) a introduit le concept de mapping comme méthode pour étudier les multiples « transcendentaux » (c'est-à-dire, les conditions de possibilité) des objets techniques.

domaine. Par exemple, les récits sur l'IA permettent de cristalliser et de produire des imaginaires technologiques spécifiques.

La littérature sur la communication et la vulgarisation scientifique a longtemps insisté sur la manière dont la communication contribue aux processus de production technologique et de découverte scientifique. Comme l'a souligné Bucchi (1996), la communication scientifique au niveau populaire influence la pratique scientifique de diverses manières, bien au-delà du simple soutien ou renforcement. Cette idée est développée ultérieurement par Goulden (2013), qui compare la relation entre la science et la vulgarisation scientifique à la rencontre entre l'eau douce et l'eau salée dans le delta d'une rivière. Plus récemment, Bucchi et Trench (2021) ont défini la communication scientifique comme « la conversation sociale autour des sciences et des techniques ».

*Deux utilisations connexes du terme « conversation » sont en jeu ici : un mode de communication interactive qui se distingue de la diffusion ou d'autres modes hiérarchiques, et un concept qui englobe tout ce qui est dit sur un sujet donné dans la société. Notre définition inclusive de la communication scientifique valide non seulement des activités telles que les cafés scientifiques et la comédie scientifique, qui sont orientées vers le plaisir, mais reconnaît également, comme faisant partie de la pratique plus large de la communication scientifique, l'utilisation « spontanée » dans la culture populaire d'images et d'idées issues de la science ou qui y sont liées (Bucchi et Trench 2021 : 6. Nous traduisons).*

On pourrait donc dire que la communication scientifique, en tant que conversation sociale, ouvre la science au débat entre experts et non-experts, en intégrant ces derniers dans les processus d'innovation technoscientifique. Il en va de même pour les produits de la communication scientifique, qui possèdent, en quelque sorte, leur propre ontologie, spécifique et indépendante des faits scientifiques et des productions scientifiques (comme les publications académiques) qui en découlent. Selon Jeanneret (2008 : 104), un article de revue de vulgarisation scientifique n'est pas seulement une leçon simplifiée sur la science, mais un objet complexe qui mobilise des conceptions quotidiennes, des jugements moraux, des définitions de caractère et des interactions sociales.

Au vu de ce qui a été discuté jusqu'à présent, il n'est pas surprenant de constater un décalage entre l'éthique de l'IA et le manque d'éthique dans la communication scientifique autour de l'IA. À titre d'exemple, prenons la manière dont la Communauté Européenne a choisi de communiquer visuellement sur une de ses initiatives majeures dans le domaine de l'éthique de l'IA, les « Lignes directrices en matière d'éthique pour une IA digne de confiance », en utilisant une image achetée sur Shutterstock<sup>3</sup>.

Comme nous l'avons souligné ailleurs (Romele et Severo, 2023), le problème avec de telles images n'est pas tant leur éloignement de l'IA « réelle », mais plutôt leur manque de « pensivité », un concept emprunté au philosophe Jacques Rancière. Les couleurs utilisées, notamment le bleu, et la patine évoquant les images pieuses (Dondero, 2007), ainsi que les sujets représentés (un robot qui « pense », la rencontre entre un doigt robotique et un doigt humain, etc.) produisent un effet anesthésiant. Au lieu d'inciter les non-experts à participer au débat sur l'IA et son éthique, ces images les en dissuadent. Dans un esprit similaire, Robert (2017 : 221) soutient que le Macro-Techno-Discours « ne fournit pas de cadres d'intelligibilité des TIC ; il remplit plutôt une fonction différente : nous submerger dans un flot discursif [...] qui occupe l'espace et empêche ainsi l'émergence de questions pertinentes, qu'elles soient politiques ou éthiques ».

Dans ce contexte, notre intérêt se porte non sur l'éthique de l'IA en tant que telle, mais sur la politique qui sous-tend l'utilisation de l'expression « éthique de l'IA ». Comme il est bien connu, le marxisme orthodoxe prend ses racines dans un renversement de la perspective hégélienne, passant de la matière à l'Esprit. Inspiré par les thèses de Feuerbach, Marx conclut que tout ce qui n'est pas matériel, que ce soit la religion, l'art ou le droit, n'est pas réel. C'est précisément pour cette raison que la lutte pour l'hégémonie dans la société se joue au niveau matériel. En effet, pour Marx et le marxisme orthodoxe, la révolution communiste implique une réappropriation des moyens de production détenus par la classe bourgeoise. Par ailleurs, il existe une branche du marxisme que l'on pourrait qualifier de symbolique, qui, sans nier l'importance de la

---

<sup>3</sup> <https://ec.europa.eu/futurium/en/ai-alliance-consultation.1.html>. Dernier accès le 29 février 2024. L'image est une des dizaines de variantes de La Création d'Adam de Michel-Ange en version humain-robot.

matière, reconnaît aussi une domination symbolique non moins significative – voir à ce propos Ricœur (2016).

Parmi les figures de ce marxisme symbolique, Antonio Gramsci occupe une place notable. L'intérêt du marxisme gramscien réside non seulement dans la « découverte » de la dimension symbolique de la domination, mais également dans sa reconnaissance que les dominés, bien que matériellement inférieurs, disposent toujours de marges de manœuvre sur le plan symbolique. C'est pour cette raison que le marxisme gramscien a rencontré un succès considérable dans le domaine des *cultural studies*, qui ont par ailleurs été déjà massivement mobilisés dans le domaine des études médiatiques (Maigret et Martin, 2020).

La théorie de Laclau et Mouffe a été justement qualifiée de « néo-gramscienne » (Martin, 2002). Trois éléments clés définissent cette perspective : (1) l'idée de la constitution discursive de la société, (2) l'analyse de discours spécifiques, et (3) le concept d'hégémonie. Concernant le premier point, Laclau et Mouffe voient le discours non pas comme un aspect parmi d'autres de la société, mais comme son fondement ontologique. Tout ce qui n'est pas articulé dans le discours d'une société est, en quelque sorte, inexistant au sein de celle-ci. Ce phénomène représente un premier acte d'inclusion et d'exclusion qui définit également le « politique », envisagé comme le moment fondateur d'une société. Pour le deuxième point, les termes et concepts utilisés dans le discours n'ont pas de valeur intrinsèque, mais uniquement une valeur différentielle, rendant les significations socialement hégémoniques vulnérables et instables. Enfin, concernant le troisième point, c'est ici que l'influence de Gramsci est la plus manifeste. Toutefois, alors que pour Gramsci, l'hégémonie est une pratique de classe, Laclau et Mouffe l'envisagent comme le principe d'articulation discursive de la société dans son ensemble. En somme, pour Laclau et Mouffe, la réalité sociale est le résultat d'une lutte pour l'hégémonie symbolique et, simultanément, de la dissimulation des traces de cette lutte par ceux qui en émergent victorieux.

C'est dans ce cadre théorique que Laclau a développé son concept de signifiant flottant. Le signifiant flottant ne représente pas simplement un cas de polysémie d'un terme ou d'un concept. Il désigne plutôt un terme

ou concept qui est « soumis à la pression structurelle de projets hégémoniques concurrents » (Laclau, 2008 : 155), se trouvant ainsi au cœur de contestations par différents projets politiques. Selon Laclau, les signifiants flottants émergent principalement en période de crise, lorsque le système symbolique nécessite une transformation radicale. Laclau lui-même explique :

*Nous savons maintenant que construire un « peuple » implique aussi de construire la frontière que le « peuple » présupposé. Les frontières sont instables et subissent un processus de déplacement constant. C'est pourquoi j'ai parlé de « signifiants flottants ». Cela entraîne un nouveau jeu hégémonique : tout nouveau « peuple » exigerait la reconstitution d'un espace à travers la construction d'une nouvelle frontière. La même chose arrive avec les « éléments extérieurs » au système : toute transformation politique implique non seulement une reconfiguration des demandes déjà existantes, mais aussi l'incorporation de demandes nouvelles (c'est-à-dire de nouveaux acteurs historiques) à la scène politique, ou son contraire, c'est-à-dire l'exclusion d'autres demandes (et d'autres acteurs) qui étaient présents dans la situation antérieure (Laclau, 2008 : 181).*

Dans une certaine mesure, le flottement des signifiants est au cœur même du politique. Pour Laclau, l'absence de flottement, le passage à une logique d'équivalences et de signifiants vides, signifie la mort du politique, au profit de ce que Rancière nomme la « police » (Laclau 2008 : 183).

La notion de signifiant flottant a également trouvé sa place dans le domaine des études médiatiques, notamment dans l'analyse du phénomène des *fake news* par Johan Farkas et Jannick Schou (2019). Ces auteurs critiquent l'approche du « solutionnisme basé sur la vérité » face aux *fake news*. Reconnaisant la propagation des *fake news* comme une menace pour les démocraties, ils mettent en garde contre la croyance en la toute-puissance de la raison et de la « vérité factuelle » pour résoudre ce problème. Ils argumentent que la démocratie ne repose pas uniquement sur la raison et la vérité, mais également sur l'inclusion et la représentation de divers projets et groupes politiques. Elle embrasse différentes visions de l'organisation sociale. Cela n'implique pas d'abandonner la quête de vérité, mais de la contextualiser et de



comprendre que la recherche de vérité et la promotion de l'agonisme dans nos démocraties ne sont pas mutuellement exclusives. Comme ils le soulignent ailleurs, l'expression « *fake news* » peut être vue comme un signifiant flottant :

*Au lieu d'entrer sur le terrain de ce qui définit la « vérité » ou la « fausseté », un champ de bataille dans lequel une multiplicité d'agents lutte pour définir ce qui compte comme valide ou trompeur, nous cherchons à comprendre « fake news » comme un signifiant discursif qui fait partie des luttes politiques. Nous prenons du recul et examinons comment les différentes conceptions de « fake news » servent à produire et à articuler des champs de bataille politiques sur la réalité sociale (Farkas et Shou, 2018 : 300. Nous traduisons).*

Particulièrement, à partir d'une analyse de la presse, ils identifient trois utilisations de l'expression « *fake news* » : (1) comme un terme utilisé pour critiquer le capitalisme numérique, c'est-à-dire l'idée que les plateformes génèrent des revenus sur la base du nombre de lecteurs ou de visiteurs, ce qui les incite à produire ou du moins à ne pas entraver la production de *fake news* ; (2) comme une critique de la politique et des médias de droite, souvent perçus comme de grands producteurs de *fake news* ; (3) comme une critique des médias libéraux et mainstream, un discours fréquemment utilisé par des figures comme Donald Trump, qui, accusé de diffuser des *fake news*, renvoie l'accusation vers les médias qui le critiquent. Ces différentes utilisations du terme révèlent une lutte pour l'hégémonie symbolique et culturelle entre divers projets politiques – par exemple, entre le pouvoir politique et le pouvoir journalistique.

Dans la section suivante, nous développons l'idée de mobiliser cette même intuition, ainsi qu'une méthodologie similaire, pour étudier les utilisations de l'expression « éthique de l'IA ».

## 2. « Éthique de l'IA », en pratique

Dans leur recherche empirique, Farkas et Shou (2018) ont identifié des discours issus de différentes sources où le terme « *fake news* » est utilisé comme un signifiant soutenant des agendas politiques spécifiques. Le matériel analysé comprenait des contenus de médias sociaux de Donald Trump ainsi que des articles journalistiques et des commentaires

académiques provenant de diverses sources telles que *The Guardian*, *The New York Times*, *The Conversation*, et *CNN*, sur une période de cinq mois. Dans notre cas, nous avons décidé de nous concentrer uniquement sur les articles journalistiques, suivant l'idée que la presse peut être vue comme une arène commune (ce qui ne signifie pas neutre) où se jouent des prétentions à l'hégémonie symbolique rivales<sup>4</sup>. Nous avons également décidé d'aller au-delà de la seule langue anglaise pour englober quatre différentes réalités européennes. Ce choix n'est pas destiné à montrer de manière comparative les différences (qui existent certainement, bien que non analysées ici), mais plutôt à démontrer la transversalité de cette lutte pour l'hégémonie<sup>5</sup>. En ce qui concerne la méthode, à l'instar de Farkas et Shou, nous nous référons principalement à l'analyse critique du discours (ACD).

L'ACD est une approche interdisciplinaire qui examine les liens entre le discours, le pouvoir et l'idéologie. Selon Fairclough (2013), l'ACD s'intéresse à la manière dont les pratiques discursives façonnent et sont façonnées par les structures sociales. Fairclough propose un modèle en trois dimensions pour l'analyse du discours, qui examine le texte, la pratique discursive, et la pratique sociale, permettant ainsi de comprendre les effets du langage dans les relations de pouvoir. Van Dijk (2015) met l'accent sur la manière dont le discours contribue à la reproduction de la dominance sociale par les élites, en se focalisant particulièrement sur les stratégies discursives qui influencent les représentations sociales et les attitudes. Pour van Dijk, l'analyse du discours doit aussi prendre en compte les contextes sociocognitifs qui influencent la production et la réception des textes.

---

4 Nous sommes conscients que les mêmes organes de presse ont leurs propres intérêts et orientations, que nous n'avons toutefois ni l'espace ni le temps de prendre en compte dans ce contexte. D'autre part, on pourrait argumenter que les journalistes représentent une quatrième sphère d'acteurs, distincte des trois que nous avons isolées, et qui avancent leurs propres prétentions hégémoniques dans l'utilisation de l'expression « éthique de l'IA ».

5 Il serait en effet pertinent de prêter davantage attention aux différences dans le niveau d'attention que divers organes de presse, dans différents pays, accordent au thème de l'éthique de l'IA. Il serait également intéressant d'étendre une éventuelle étude comparative au-delà des frontières européennes. Cela dépasse toutefois le cadre de cet article.

Notre travail se concentre exclusivement sur l'analyse des acteurs sociaux et leur manière de parler d'éthique de l'IA. En analysant comment différents acteurs utilisent le terme « éthique de l'IA », nous touchons à l'aspect de la pratique sociale de l'ACD, en révélant comment les discours sont affectés par et affectent les contextes socio-politiques du développement technoscientifique. Notre recherche ne se concentre pas sur tous les aspects de l'ACD, tels que l'analyse détaillée des structures linguistiques ou des styles discursifs. De plus, bien que l'aspect cognitif de la réception des discours soit central chez van Dijk, notre étude ne s'étend pas à la réception des discours par le public.

Pour notre analyse, nous avons consulté la base de données Europresse. Nous avons restreint notre recherche aux articles publiés dans la presse quotidienne sur une période de trois mois, du 9 mai au 9 août 2023. Nous avons recherché des articles contenant, soit dans le titre soit dans le texte, les termes « éthique » et « intelligence artificielle », ou « éthique » et « IA ». Cette recherche a été menée pour quatre pays européens : la France, l'Italie, le Royaume-Uni et l'Allemagne. Pour chaque pays, nous avons sélectionné deux journaux selon leur popularité et réputation : *Le Monde* et *Libération* pour la France, *Corriere della Sera* et *Repubblica* pour l'Italie, *The Guardian* et *The Independent* pour le Royaume-Uni, ainsi que *Frankfurter Allgemeine Zeitung* et *Süddeutsche Zeitung* pour l'Allemagne. Une recherche distincte a été réalisée pour chaque pays, les mots-clés étant traduits dans la langue respective. Le tableau ci-dessous résume les résultats de notre recherche :

Pays	Mots clés	Journaux	Résultats pertinents
France	« éthique » « intelligence artificielle (IA) »	<i>Le Monde</i> (journal + site)	22
		<i>Libération</i> (journal + site)	17
		<b>Total France</b>	<b>39</b>
Italie	« etica » « intelligenza artificiale (IA) »	<i>Corriere della Sera</i> (site)	24
		<i>Repubblica</i> (site)	13

		<b>Total Italie</b>	<b>37</b>
Royaume- Uni	« ethics » « artificial intelligence (AI) »	<i>The Guardian</i> (site)	31
		<i>The Independent</i> (site)	1
		<b>Total Royaume Uni</b>	<b>32</b>
Allemagne	« Ethik » « künstliche Intelligenz (KI) »	<i>FAZ</i> (site)	4
		<i>SZ</i> (site)	5
		<b>Total Allemagne</b>	<b>9</b>
		<b>Total</b>	<b>117</b>

Tableau 1 : articles contenant, soit dans le titre soit dans le texte, soit « éthique » et « intelligence artificielle », soit « éthique » et « IA » dans huit quotidiens européens dans la période 9 mai – 9 août 2023.

Dans notre analyse, nous avons identifié trois principaux groupes d'acteurs mobilisant l'expression « éthique de l'IA » : (1) les institutions, (2) les chercheurs et chercheuses, et (3) les entreprises. Bien que tous parlent d'éthique de l'IA, la signification qu'ils attribuent à « éthique de l'IA » varie considérablement. Notre hypothèse est que la manière dont ces acteurs abordent l'éthique de l'IA (ou, pour être plus précis, la manière dont leurs points de vue sont représentés dans les articles de presse) reflète un désir hégémonique dans ce domaine et, plus généralement, dans le champ de l'innovation en IA.

### **2.1 La normativité des institutions**

Dans notre étude, nous avons constaté que les institutions européennes et leurs représentants sont fréquemment sollicités. Une attention particulière est accordée à l'IA Act européen, c'est-à-dire le projet de loi destiné à réguler le développement et l'utilisation des systèmes d'IA en Europe dans les prochaines années. Par exemple, un article du *Frankfurter Allgemeine Zeitung* du 12 juin 2023, intitulé « Le Parlement européen s'accorde sur le cadre de la loi », évoque le vote sur cette législation qui a eu lieu deux jours plus tard, recueillant 499 voix pour, 28 contre et 93 abstentions. L'article soulève les questions des longues négociations préalables au vote, durées 18 mois, mettant en

lumière le clivage entre les députés percevant les systèmes d'IA puissants comme une menace et ceux inquiets d'une réglementation excessive, ainsi que les critiques à l'égard du caractère trop généraliste du cadre proposé par la Commission, ne distinguant pas suffisamment les secteurs et domaines d'application des systèmes d'IA à haut risque.

Outre les instances européennes, les institutions nationales sont également consultées. *Le Monde*, le 5 juillet 2023, publie un article intitulé « IA : le comité d'éthique propose de réguler davantage les services tels que ChatGPT que les modèles 'open source' », présentant la proposition du Comité national pilote d'éthique du numérique (CNPEN) de différencier les systèmes d'IA générative accessibles au grand public, tels que ChatGPT, des systèmes en open source moins accessibles. Cet article aborde aussi brièvement les tensions entre le Parlement européen et des acteurs comme Emmanuel Macron, qui, en accord avec des startups françaises telles que Mistral AI ou LightOn, craint que certaines mesures ne freinent le développement de champions européens dans ce domaine.

En effet, il est évident que la lutte pour l'hégémonie dans le domaine de l'IA oppose principalement deux groupes d'acteurs : les institutions et les entreprises, avec des chercheurs et chercheuses aligné.e.s de part et d'autre. Toutefois, il est également essentiel de reconnaître que l'harmonie n'est pas toujours présente au sein même des institutions. Par exemple, un article du *Monde* du 10 juin 2023, intitulé « Compétition mondiale pour la régulation de l'IA », rapporte les efforts du Royaume-Uni, sous l'égide de son Premier ministre Rishi Sunak, pour se positionner comme un hub mondial de la régulation de l'IA, bien que l'Europe semble avoir une avance dans ce domaine.

Les acteurs institutionnels attribuent à l'expression « éthique de l'IA » un sens normatif. Par normativité, nous entendons que, selon les institutions, l'éthique de l'IA devrait être mise en œuvre par l'imposition et la vérification du respect des normes par les innovateurs en IA, en particulier les entreprises. Un article du 16 mai 2023 publié par le *Corriere della Sera*, intitulé « ChatGPT, Bard, Bert : l'OMS appelle à une intelligence artificielle plus éthique et plus sûre dans les applications de santé », mentionne les six principes fondamentaux énoncés par l'OMS dans leur guide sur l'éthique et la gouvernance de l'IA dans le domaine de

la santé. Ces principes échoient à ceux formulés par d'autres institutions, tels que la Commission européenne dans son document de 2019 sur les lignes directrices pour une IA « digne de confiance ». Les institutions optent donc pour des approches normatives en matière d'éthique de l'IA, car c'est précisément par le contrôle des normes qu'elles exercent traditionnellement leur pouvoir. La normativité, en tant que caractéristique du discours institutionnel sur l'éthique de l'IA, est reconnue tant par les institutions elles-mêmes que par leurs détracteurs, qui ne remettent pas en question l'efficacité des normes mais les opposent aux impératifs de l'innovation technologique, souvent dans un contexte de concurrence mondiale, selon une rhétorique TINA (*there is no alternative*) (Defilippi, 2022).

## **2.2 La critique des chercheurs et des chercheuses**

Concernant les chercheurs et chercheuses, il est notable que dans la presse quotidienne, il n'y a généralement pas de distinction entre les représentants des sciences humaines et sociales (SHS) et ceux des sciences dures. Parmi les articles que nous avons examinés, nous trouvons autant d'interventions sollicitant l'opinion de philosophes et sociologues que celle d'ingénieurs et d'informaticiens. À titre d'exemple, nous pouvons citer un long entretien avec le philosophe Peter Singer publié dans *The Guardian* le 21 mai 2023, ainsi qu'un autre article paru le lendemain dans le même journal avec l'informaticienne Timnit Gebru, qui a été licenciée par Google en 2020 alors qu'elle codirigeait une équipe de recherche sur l'éthique de l'IA au sein de l'entreprise.

Cela ne signifie toutefois pas l'absence de luttes hégémoniques au sein de ce groupe d'acteurs. On pourrait même avancer l'hypothèse que l'émergence de l'éthique de l'IA est liée à la volonté des sciences humaines et sociales de participer à l'innovation et aux financements dans le domaine de l'IA. Pourquoi les sciences dures s'intéressent de plus en plus à l'éthique de l'IA ? Premièrement, il est à noter qu'au cours des dernières années, l'intégration d'évaluations éthiques dans les projets technoscientifiques est devenue un critère fondamental, voire obligatoire, pour l'obtention de financements publics. Deuxièmement, les chercheurs et chercheuses en sciences dures affiliés à des institutions publiques, telles que les universités, ont vu leur influence diminuer

significativement (en termes d'investissements et de capacités d'innovation) par rapport à leurs homologues du secteur privé.

Un article paru dans *Le Monde* le 31 mai 2023, intitulé « Le match perdu de l'université face au privé », se réfère à une étude publiée dans *Science* près de trois mois plus tôt (Ahmed et al., 2023), et souligne que de nos jours, 70 % des doctorants en IA sont recrutés par des entreprises privées et 40 % des communications dans les conférences proviennent de laboratoires privés, soit le double de ce qui était observé en 2012. Même des institutions prestigieuses comme le MIT à Boston font face à un manque de financements comparé aux laboratoires d'entreprise. Nous pouvons donc avancer l'hypothèse selon laquelle l'éthique de l'IA est en train de devenir un refuge (apparemment) moins coûteux pour les chercheurs et les chercheuses en sciences dures. L'article du *Monde* note aussi que même dans le domaine de l'éthique de l'IA, les entreprises dominent de plus en plus : « par exemple, elles ont soumis trois fois plus de travaux en 2022 qu'en 2021, représentant un tiers de la production académique ». La lutte pour l'hégémonie dans le domaine de l'IA est donc totale.

La conception de l'éthique de l'IA par les chercheurs et les chercheuses s'oriente principalement vers la critique. Nous définissons ici le terme « critique » comme une approche qui implique une vigilance constante vis-à-vis des pratiques existantes et des intentions des acteurs impliqués, notamment les entreprises. Si certains chercheurs et chercheuses plaident pour une normativité renforcée, d'autres semblent adhérer aux solutions technologiques avancées par les entreprises. Le 26 mai 2023, un article publié dans *Le Monde* sous le titre « Le défi de détecter les contenus issus des IA » discute de l'idée d'implémenter un filigrane cryptographique invisible sur les images et textes générés par l'IA. Cette proposition, soutenue par des entreprises telles que Microsoft, Google et Amazon, ainsi que par le gouvernement des États-Unis, rencontre également l'approbation d'Alexei Grinbaum, le chercheur cité dans l'article. Toutefois, adoptant une perspective justement critique, Grinbaum souligne que « les systèmes de détection [du filigrane] ne sont pas infaillibles ».

Dans les colonnes de *Libération*, le 20 juin 2023, un article intitulé « Après l'éco-anxiété, l'IA-anxiété ? » donne la parole à des penseurs critiques de l'IA tels qu'Antonio Casilli et Thierry Ménéssier. Cet article est issu d'un numéro spécial entièrement consacré à l'IA par le journal. Le même numéro propose également un entretien avec le théoricien des images Antonio Somaini. La critique nous rappelle l'importance des chercheurs et des chercheuses et de leur posture réflexive, soulignant que l'éthique de l'IA ne peut se résumer à une série de solutions technologiques ou de prescriptions normatives.

### **2.3 Le techno-solutionnisme des entreprises**

Concernant les entreprises, une approche distincte est requise. Notre analyse révèle que, contrairement aux chercheurs et chercheuses et aux institutions (qui sont presque également consulté.e.s), les entreprises sont rarement présentées comme les principaux sujets de ces discours. Elles sont généralement perçues comme des objets ou des cibles des discours critiques et normatifs des deux autres groupes. En examinant la presse quotidienne, il apparaît clairement que les entreprises ne sont souvent pas considérées comme de véritables parties prenantes dans ces discussions, mais plutôt comme des entités distantes, voire absentes. De manière intéressante, les médias donnent fréquemment la parole à celles et ceux qui se sont éloigné.e.s des entreprises pour lesquelles iels ont longtemps travaillé. En effet, il est bien connu qu'il n'y a pas de critique plus puissante d'une cause que celle émanant de personnes qui se sont converties ou qui se sont repenties. Généralement, cela montre, comme nous l'avons souligné dans une note précédente, que la presse n'est pas exempte d'intérêts. Et il est certain qu'en se tournant vers un autre type de presse (comme la presse économique et financière), les résultats seraient différents.

Il existe cependant des exceptions, comme l'article du 23 mai 2023 dans *La Repubblica*, consacré à une intervention de Bill Gates lors d'un événement organisé par Goldman Sachs et SV Angel. Cet article a été publié dans la section économie et finance du quotidien italien. Le 4 août 2023, *The Guardian* a publié un article rédigé par Dorothy Chou, responsable des affaires publiques chez Google DeepMind, sous le titre « Les entreprises d'IA n'ont pas peur de la réglementation – nous voulons



qu'elle soit internationale et inclusive ». Dans un extrait significatif de cet article, Chou déclare : « Les lois ne sont efficaces que si elles anticipent l'avenir. C'est pourquoi il est crucial pour les régulateurs d'envisager non seulement la manière de réglementer les chatbots aujourd'hui, mais aussi de promouvoir un écosystème où l'innovation et le progrès scientifique peuvent bénéficier à la société ».

Pour les entreprises et leurs représentants, l'éthique de l'IA est perçue comme une forme de techno-solutionnisme. Ce dernier valorise les entreprises dans le domaine de l'éthique de l'IA, car elles possèdent les compétences nécessaires pour élaborer les meilleures solutions technologiques. L'idée sous-jacente est que les entreprises n'ont besoin ni des institutions ni des chercheurs et chercheuses, étant donné qu'elles sont capables de se réguler de manière éthique par elles-mêmes.

À ce sujet, en mettant momentanément de côté les discours de presse pour examiner une portion du contenu en ligne sur l'éthique de l'IA proposé par un acteur majeur comme Google, on trouve une ouverture de page significative (nous traduisons) : « Bien que nous soyons optimistes quant au potentiel de l'IA, nous reconnaissons que les technologies avancées peuvent poser d'importants défis qui doivent être abordés de manière claire, réfléchie et proactive »<sup>6</sup>. Il y a donc une pleine reconnaissance des enjeux éthiques que l'IA peut soulever. Cependant, cette préoccupation n'éclipse pas un certain optimisme ambiant. La solution n'est pas la peur, mais plutôt une action réfléchie et efficace. Sur la même page, Google énonce six principes qui, à certains égards, rappellent ceux des institutions. Le dernier principe, celui de « l'excellence scientifique », est particulièrement révélateur. La section finale, intitulée « Mise en œuvre de nos principes », illustre comment ces principes sont appliqués matériellement par les technologies de Google (par exemple, l'évaluation de la couleur de la peau par apprentissage automatique), démontrant ainsi que le techno-solutionnisme est jugé comme le meilleur moyen de répondre de manière affirmative aux problématiques éthiques de l'IA.

---

<sup>6</sup> <https://ai.google/responsibility/principles>. Dernier accès le 29 février 2024.

## Conclusion

Dans cet article, nous avons exploré les usages de l'expression « éthique de l'IA ». Notre objectif n'était donc pas de définir précisément ce qu'est l'éthique de l'IA, ni de juger le caractère éthique ou non de certains systèmes d'IA. Nous avons plutôt cherché à analyser les diverses manières, souvent conflictuelles, d'utiliser et de comprendre cette expression. Aussi paradoxal que cela puisse paraître, l'éthique de l'IA, avant d'être un concept éthique, est une notion profondément politique. Il est fascinant de voir comment différents projets technico-politiques se confrontent autour de la définition de ce qu'il convient d'appeler « éthique de l'IA ».

Il n'est pas question pour nous de nier l'importance des enjeux éthiques liés à l'IA. Notre but est plutôt de souligner que la façon dont nous répondons et répondrons aux questions d'éthique de l'IA sera toujours conditionnée par le contexte, notamment politique et social. Avant même de déterminer ce qui est éthique ou non dans un système d'IA, nous devons nous pencher sur ce que nous pourrions appeler la « politique des éthiques de l'IA ».

Dans notre analyse empirique, nous avons identifié trois groupes d'acteurs engagés dans cette lutte pour l'hégémonie : les institutions, les chercheurs et chercheuses, et les entreprises. Chacun de ces groupes porte une vision spécifique de la manière dont l'éthique devrait être appliquée dans les systèmes d'IA. Nous avons évoqué la normativité des institutions, la critique portée par les chercheurs et les chercheuses, et le techno-solutionnisme des entreprises, illustrant trois approches très différentes de la mise en œuvre de l'éthique dans l'IA.

En conclusion, nous pouvons affirmer que l'issue de cette lutte pour l'hégémonie ne sera pas nécessairement une IA plus éthique, mais plutôt la prédominance d'un de ces groupes qui définira ce qui est éthiquement acceptable et donc la légitimité d'un système d'IA.

## Bibliographie

- Ahmed Nur, Wahed Muntasir, Thompson Neil C. (2023). The growing influence of industry in AI research. *Science*, vol. 379, no 6635, pp. 884–886.
- Brey Philip (2010). Philosophy of Technology after the Empirical Turn. *Techné: Research in Philosophy and Technology*, vol. 14, no 1, pp. 36–48.
- Bucchi Massimiano (1996). When scientists turn to the public: Alternative routes in science communication. *Public Understanding of Science*, vol. 5, no 4, pp. 375–394.
- Bucchi Massimiano, Trench Brian (2021). Rethinking science communication as the social conversation around science. *Journal of Science Communication*, vol. 20, no 3, pp. 1-11.
- Coeckelbergh Mark (2021). Time Machines: Artificial Intelligence, Process, and Narrative. *Philosophy & Technology*, vol. 34, no 4, pp. 1623-1638.
- Defilippi Fabrizio (2022). Il n'y a pas d'alternative : Les imaginaires de l'innovation dans les discours d'Emmanuel Macron. *Interfaces numériques*, vol. 11, no 1. <https://doi.org/10.25965/interfaces-numeriques.4755>
- Dondero Maria Giulia (2007). Reproductibilité, faux parfaits et contrefaçons : Entre fétichisme artistique et goût esthétique. *Nouveaux Actes Sémiotiques*. <https://doi.org/10.25965/as.3261>
- Farkas Johan, Schou Jannick (2018). Fake News as a Floating Signifier: Hegemony, Antagonism and the Politics of Falsehood. *Javnost - The Public*, vol. 25, no 3, pp. 298–314.
- Farkas Johan, Schou Jannick (2019). *Post-Truth, Fake News and Democracy: Mapping the Politics of Falsehood*. Londres et New York, Routledge.
- Fairclough Norman (2013). *Critical Discourse Analysis*. Londres et New York, Routledge.
- Floridi Luciano (2016). Tolerant Paternalism: Pro-ethical Design as a Resolution of the Dilemma of Toleration. *SSRN Electronic Journal*.
- Goulden Murray (2013). Hobbits, hunters and hydrology: Images of a « missing link », and its scientific communication. *Public Understanding of Science*, vol. 22, no 5, pp. 575–589.
- Jeanneret Yves (2008). *Penser la trivialité*. Paris, Hermès Lavoisier.
- Laclau Ernesto (2008). *La raison populiste*. Paris, Seuil.
- Maigret Eric, Martin Laurent (dir.) (2020). *Les cultural studies : Au-delà des politiques des identités*. Lormont, Le Bord de l'eau.
- Martin James (2002). The political logic of discourse: A neo-Gramscian view. *History of European Ideas*, vol. 28, no 1–2, pp. 21–31.

- Priest Susanne, Goodwin Jean, Dahlstrom Michael F. (dir.) (2018). *Ethics and practice in science communication*. Chicago, The University of Chicago Press.
- Ricœur Paul (2016). *L'idéologie et l'utopie*. Paris, Seuil.
- Robert Pascal (2017). L'imaginaire des TIC en question. *Interfaces numériques*, vol. 4, no 2.
- Romele, A. (2024). *Digital Habitus: A Critique of the Imaginaries of Artificial Intelligence*. Londres et New York, Routledge.
- Romele Alberto, Severo Marta (2023). Que veulent les images de l'IA ? Une exploration de la communication scientifique visuelle de l'intelligence artificielle. *Sociétés & Représentations*, no 1, pp. 179–201.
- Smith Dominic (2018). *Exceptional technologies : A continental philosophy of technology*. Londres, Bloomsbury.
- Van Dijk Teun A. (2001). Critical discourse analysis. Dans Tannen Deborah, Schiffrin Deborah, Hamilton Heidi E. (dir.), *Handbook of discourse analysis*. Oxford: Blackwell, pp. 352-371.