

# Enjeux numériques



## Pour une IA responsable et éthique

N°29 - MARS 2025



*Publiées avec le soutien  
de l'Institut Mines-Télécom*



## ENJEUX NUMÉRIQUES

ISSN 2781-1263 (en ligne)

ISSN 2607-9984 (imprimé)

Série trimestrielle - N°29 - Mars 2025

### Rédaction

Conseil général de l'Économie  
Ministère de l'Économie,  
des Finances  
et de la Souveraineté  
industrielle et numérique  
120, rue de Bercy  
Télédoc 797  
75572 Paris Cedex 12  
Tél. : 01 53 18 52 68  
<http://www.annales-des-mines.org>

**Grégoire Postel-Vinay**  
Directeur de la publication  
et Rédacteur en chef

**Alexia Kappelman**  
Secrétaire générale

**Daniel Boula**  
Secrétaire général adjoint

**Magali Gimon**  
Assistante de rédaction  
et Maquettiste

**Nuria Gorris**  
Webmestre et Maquettiste

### Publication

**Photos de couverture**  
Le Tricheur à l'as de carreau,  
Photo © RMN-Grand Palais  
(musée du Louvre) /  
Adrien Didierjean

**Iconographie**  
Daniel Boula

**Mise en page**  
Magali Gimon

**Impression**  
Duplprint Mayenne

### Membres du Comité de rédaction

**Pierre Bonis**  
Co-président  
**Anne-Lise Thouroude**  
Co-présidente  
**Edmond Baranes**  
**Godefroy Beauvallet**  
**Côme Berbain**  
**Hélène Brisset**  
**Serge Catoire**  
**Nicolas Chagny**  
**Jean-Pierre Dardayrol**  
**Éric Freyssinet**  
**Frédéric Garcia**  
**Francis Jutand**  
**Arnaud de La Fortelle**  
**Caroline Leboucher**  
**Bertrand Pailhès**  
**Grégoire Postel-Vinay**  
**Maurice Ronai**  
**Laurent Toutain**  
**Benjamin Vignard**

La mention au regard de certaines illustrations du sigle « D. R. » correspond à des documents ou photographies pour lesquels nos recherches d'ayants droit ou d'héritiers se sont avérées infructueuses.

Le contenu des articles n'engage que la seule responsabilité de leurs auteurs.

# Pour une IA responsable et éthique

## INTRODUCTION

- 05 **Pour une intelligence humaine**  
Nicolas CHAGNY
- 08 **Enjeux et perspectives pour une IA éthique et durable**  
Guillaume BOURGEOIS,  
Luciana GONDIM DE ALMEIDA GUIMARÃES  
et Vincent COURBOULAY

## L'IA POUR TOUS, TOUS POUR L'IA

- 15 **Démocratiser notre rapport à la technologie,  
un Café IA à la fois**  
Jean CATTAN
- 19 **Vers une société « IAApprenante »**  
Frédéric BARDEAU
- 22 **Le formateur augmenté : entre  
intelligence artificielle et intelligence émotionnelle**  
Brice GAILLARD
- 26 **Pour une éducation à la pluralité des altérités  
et des attachements numériques à l'heure  
de l'intelligence artificielle générative**  
Jean-François LUCAS
- 33 **La création artistique à l'épreuve  
de l'intelligence artificielle**  
Alain ASSOULINE

## ENJEUX ENVIRONNEMENTAUX

- 37 **Les impacts de l'IA sur l'environnement**  
Frédéric GARCIA et Sophie SCHBATH
- 43 **L'IA durable n'existe pas**  
Frédéric MARCHAND
- 50 **Pour un développement de l'IA  
au service du bien commun**  
Laure de LA RAUDIERE

- 58 **Intégrer l'IA dans un service éco-conçu :  
oxymore ou réalité ?**  
Christophe CLOUZEAU, Vincent COURBOULAY,  
Mathieu DELEMME, Jean-Luc MARINI, Emmanuel NURIT,  
Romuald RIBAULT et Claire VERDIER

## ENJEUX SOCIAUX

- 65 **L'intelligence artificielle et les droits humains :  
les insuffisances du cadre européen**  
Thomas DUMORTIER
- 70 **IA générative et mésinformation**  
Nicolas CURIEN
- 77 **AI-xiety : entre mythes et réalité,  
la superintelligence artificielle est-elle déjà là ?**  
Stéphanie POTTECHER et Aurélie GIARD-JACQUET
- 85 **Santé mentale au travail et intelligence artificielle :  
entre soutien psychologique et risque de dépendance**  
Christian MAKAYA et George KASSAR
- 91 **Législations et dialogue social européens  
autour de l'intelligence artificielle**  
Franck GAMBELLI
- 100 **IA et Communs : conjuguer  
puissance technologique et habitabilité terrestre**  
Emmanuelle ROUX
- 106 **Vers une intelligence artificielle "gender by design" ?**  
Peggy VICOMTE et Camille SALINESI

## ENJEUX POUR LA PROSPÉRITÉ

- 111 **IA et futur de la civilisation :  
dystopie transhumaniste ou métamorphose créatrice ?**  
Boris SIRBEY et Hervé BÉRAUD
- 117 **IA et transformations des métiers :  
création ou destruction ?**  
Guy MAMOUMANI et Axel MAMOUMANI
- 122 **L'organisation IA-compatible ou l'art  
de savoir recruter la technologie**  
Romain RABIER
- 129 **Construisons un cadre ambitieux et apaisé  
pour mettre les IA au service  
de l'éducation et de la formation**  
Oriane LEDROIT

135 **Intelligence artificielle et territoires**  
Fabien BAZIN

140 **Profils d'appropriation de l'intelligence  
artificielle générative dans l'éducation**  
Loubna MOURTAJJI et Nathalie CHISS

**ENJEUX POUR LA VIE PRIVÉE  
ET LA VIE PUBLIQUE**

150 **IA et libertés : un défi pour la régulation**  
Marie-Laure DENIS

156 **Le Droit à l'heure de "l'intelligence artificielle"**  
Didier GUÉVEL

158 **Les défis éthiques de la convergence de l'IA,  
des neurosciences, de l'informatique et de l'ingénierie**  
Dr Laure TABOUY

---

167 **Traductions des résumés**

175 **Biographies des auteurs**

*Ce numéro a été coordonné par  
Nicolas CHAGNY*

# Pour une intelligence humaine

Par Nicolas CHAGNY

Fondateur du cabinet de conseil NS Pulse,  
président de l'Internet Society (ISOC) France,  
membre de la Commission nationale consultative des Droits de l'Homme (CNCDH)

Ce numéro d'*Enjeux numériques* se concentre sur l'importance des individus et de la planète dans le contexte de l'intelligence artificielle (IA). Les contributeurs, choisis pour leur diversité d'expériences et de points de vue, incitent à une réflexion approfondie sur nos interactions humaines, tant personnelles que professionnelles.

Le numéro souligne que, bien que la technologie de l'IA existe depuis les années 1950, ce qui a véritablement évolué récemment est l'expérience utilisateur (« UX »). Cette évolution a permis une utilisation plus large et plus démocratisée de l'IA, similaire aux révolutions technologiques antérieures comme le *web 2.0*, l'iPhone, les assistants vocaux et la 5G.

Le numéro explore également l'IA Responsable et son impact sur l'accès démocratique à la technologie, les implications environnementales, sociales, économiques, ainsi que les questions de vie privée et publique.

Les femmes, les hommes et la planète avant tout. C'est ce que, j'espère, vous retiendrez de la lecture de toutes les autrices et auteurs de ce numéro. Bien entendu, je les ai sollicités pour leur expérience et leurs points de vue divers et variés. Et je me doutais bien qu'ils partageraient avec moi ces valeurs d'humanité.

Mais ce que vous allez lire dans les pages qui suivent va bien au-delà. Les autrices et auteurs de ce numéro vont vous inciter à repenser nos interactions humaines, personnelles et professionnelles, que l'on soit novice ou spécialiste, à l'heure de l'intelligence artificielle.

## L'EXPÉRIENCE UTILISATEUR

J'entends souvent les vieux informaticiens, c'est-à-dire ceux qui ont mon âge ou un peu plus, dire que l'intelligence artificielle n'est pas nouvelle. Ils ont factuellement raison, car les prémices de l'IA se trouvent dans les années 1950<sup>1</sup>. Mais ils ont tort sur l'usage. C'est malheureusement souvent le cas des informaticiens.

Moi aussi, quand je n'étais même pas encore adolescent, je lisais les pages sur les réseaux neuronaux du magazine *Pascalissime*<sup>2</sup>. Je n'y comprenais rien à l'époque et sans doute encore maintenant. Et, si on calcule bien, c'était il y a plus de 30 ans.

Mais ce qui a changé en 2022, ce n'est pas la technologie, c'est l'expérience utilisateur. Basée sur une technologie bien sûr, mais c'est l'expérience utilisateur qui provoque le changement et qui permet un usage large, massif, par toutes et tous. Nous avons vécu la

<sup>1</sup> Qu'est-ce que l'IA et Historique de l'IA par le Conseil national du numérique et l'Inria, <https://cafeia.org/qu-est-ce-que-l-ia/>

<sup>2</sup> *Pascalissime* était un magazine consacré aux langages Pascal et Turbo Pascal dont le premier numéro parut en 1984 et le dernier en novembre 1997. Anciens numéros disponibles auprès de l'Institut Pascal.

même chose en 2003 avec le *web* 2.0, en 2007 avec l'iPhone, en 2010 avec les assistants vocaux et en 2020 avec la 5G.

C'est l'expérience utilisateur (« UX ») qui vient régulièrement tout révolutionner. La révolution de l'intelligence artificielle que nous vivons depuis 2022 est, avant tout, une révolution de l'UX et, donc, une histoire de femmes et d'hommes.

Après une définition de ce qu'est l'IA Responsable, les autrices et auteurs de ce numéro nous font part de leurs points de vue sur la démocratisation de l'accès à l'IA, sur les enjeux pour l'environnement, pour notre société, pour la prospérité et pour la vie privée et la vie publique.

## L'IA POUR TOUS, TOUS POUR L'IA

De l'art à l'éducation, de la santé à la finance, des grandes entreprises du CAC 40 à la TPE, c'est toute la société qui doit s'emparer de l'intelligence artificielle, comme elle s'est déjà emparée d'internet. S'emparer d'un sujet, c'est expérimenter, comprendre les fondements et échanger sur les peurs, besoins et envies de toutes et tous, afin de ne pas se laisser dépasser par les décisions des autres. C'est ce type de démarches que mettent en œuvre, notamment, Frédéric Bardeau avec Simplon et Jean Cattan avec les Cafés IA au sein du Conseil national du Numérique. En mars 2024, la commission de l'intelligence artificielle<sup>3</sup> remettait au président de la République un rapport intitulé « IA, notre ambition pour la France »<sup>4</sup>. La première recommandation incitait à « créer les conditions d'une appropriation collective de l'IA et de ses enjeux afin de définir collectivement les conditions dans lesquelles elle s'insère dans notre société et nos vies quotidiennes ». L'enjeu est aussi de permettre à tout le système éducatif d'accéder à la puissance des outils d'IA.

## ENJEUX ENVIRONNEMENTAUX

Parmi les enjeux posés par l'intelligence artificielle, notre planète figure au premier plan des préoccupations. Le stockage des données, la puissance de calcul, la multiplication des *data centers* sont autant de défis qui posent question et sur lesquels les mesures doivent encore s'améliorer. Lors d'un récent Café IA au Campus Fonderie de l'Image<sup>5</sup>, Louis Magnes, responsable du plaidoyer au Conseil national du Numérique, a lancé le débat en demandant à la salle si « on assumerait la dépense énergétique nécessaire pour demander à une intelligence artificielle de résoudre les enjeux environnementaux posés par l'IA ». Un bel exemple de récursivité pour les cours d'informatique.

## ENJEUX SOCIAUX

Au premier plan, l'intelligence artificielle pose la question du respect des Droits humains, un sujet que nous abordons régulièrement au sein de la Commission nationale consultative des Droits de l'Homme<sup>6</sup>. La mésinformation, comme nous avons pu le voir dans les

---

<sup>3</sup> Lettre de mission du Premier ministre Gabriel Attal du 28 février 2024, [https://www.vie-publique.fr/files/rapport/lettre%20de%20mission/lettre%20de%20mission\\_6.pdf](https://www.vie-publique.fr/files/rapport/lettre%20de%20mission/lettre%20de%20mission_6.pdf)

<sup>4</sup> Rapport « IA : une ambition pour la France » (mars 2024), <https://www.info.gouv.fr/upload/media/content/0001/09/4d3cc456dd2f5b9d79ee75feea63b47f10d75158.pdf>

<sup>5</sup> CFA Campus Fonderie de l'image, Association, Bagnolet, <https://www.campusfonderiedelimage.org>

<sup>6</sup> Fondée en 1947, à l'initiative de René Cassin, Prix Nobel de la Paix, la Commission nationale consultative des Droits de l'Homme (CNCDH) est l'institution nationale de protection et promotion des Droits de l'Homme en France, accréditée auprès des Nations unies, <https://www.cncdh.fr/>

récentes élections, est accélérée par l'intelligence artificielle et pose donc de nouveaux défis pour les démocraties. Nous élargirons le champ de nos réflexions à ce sujet avec la santé mentale, les communs et les biais de genre.

## ENJEUX POUR LA PROSPÉRITÉ

Le travail et son évolution à l'heure de l'intelligence artificielle sont des sujets que nous avons toutes et tous en tête dès qu'on aborde les peurs que provoquent l'IA. « Création ou destruction ? », ce sont Axel et Guy Mamou-Mani qui tenteront de répondre à cette question. Preuve qu'on est bien dans une révolution, c'est l'éducation et la formation qui se sont emparés immédiatement de ce sujet majeur, en s'en méfiant, en l'interdisant, puis en l'appropriant. C'est aussi un enjeu pour nos territoires.

## ENJEUX POUR LA VIE PRIVÉE ET LA VIE PUBLIQUE

Enfin, les autrices et auteurs de ce numéro d'*Enjeux numériques* consacré à une IA responsable et éthique aborderont les défis que posent l'intelligence artificielle à nos libertés individuelles et à nos données personnelles, des sujets majeurs pour le droit et pour la régulation. Autant de défis posés notamment aux autorités indépendantes de notre pays, et plus particulièrement l'Arcep et la Cnil.

## ENJEUX POUR L'EUROPE ET LA FRANCE AU CŒUR DU DÉBAT

Début février 2025, la France, sous l'égide du président de la République, a été à l'initiative du Sommet pour l'action sur l'Intelligence Artificielle, dans un contexte géopolitique en mouvement avec de nombreuses annonces croisées d'investissement et de prises de position sur les sujets liés à l'IA. Retenons de ce sommet la « Déclaration sur une intelligence artificielle durable et inclusive pour la population et la planète »<sup>7</sup> signée par 60 pays et fixant 6 priorités, dont « Promouvoir l'accessibilité de l'IA pour réduire la fracture numérique » et « Rendre l'IA durable pour les populations et la planète », avec le refus remarqué d'engagement des États-Unis et du Royaume Uni.

Et, parmi les nombreuses annonces, la création d'une « Coalition pour une IA durable »<sup>8</sup> regroupant 91 partenaires, dont 37 entreprises technologiques, dans 10 pays. Ainsi que le lancement d'une initiative impliquant acteurs publics et privés baptisée « Current AI »<sup>9</sup> pour une intelligence artificielle d'intérêt général qui définit sa mission ainsi : « Ensemble, tirons profit de l'IA pour accroître les opportunités, renforcer la sécurité et soutenir un monde prospère pour tous les êtres humains ». Cette initiative est notamment soutenue par Reid Hoffman (cofondateur, LinkedIn ; associé, Greylock), Fidji Simo (présidente, Instacart ; membre du conseil d'administration, OpenAI) et Arthur Mensch (cofondateur et président, Mistral AI).

Faisons le vœu que les propositions du présent numéro inspirent les nombreux signataires et participants de ces engagements et nouvelles alliances.

Bonne lecture !

---

<sup>7</sup> Déclaration sur une intelligence artificielle durable et inclusive pour la population et la planète, publiée le 11 février 2025 : <https://www.elysee.fr/emmanuel-macron/2025/02/11/declaration-sur-une-intelligence-artificielle-durable-et-inclusive-pour-la-population-et-la-planete>

<sup>8</sup> Coalition for Sustainable Artificial Intelligence : <https://www.sustainableaicoalition.org/>

<sup>9</sup> Current AI : <https://www.currentai.org/fr>



# Enjeux et perspectives pour une IA éthique et durable

Par Guillaume BOURGEOIS

École de technologie supérieure, ÉTS, Canada et La Rochelle Université, France

Luciana GONDIM DE ALMEIDA GUIMARÃES

École de technologie supérieure, ÉTS, Canada et Universidade Potiguar, Brésil

Et Vincent COURBOULAY

La Rochelle Université, France

L'intelligence artificielle (IA) s'est imposée comme un moteur d'innovation dans divers secteurs, mais soulève des défis éthiques, sociaux et environnementaux. Cet article explore le concept d'IA responsable, une approche fondée sur des principes d'éthique, de transparence, de sécurité et d'inclusion sociale. Il analyse également l'impact environnemental de l'IA, en particulier la consommation énergétique des centres de données, et les risques de reproduction d'inégalités sociales. En proposant des solutions techniques et politiques, telles que l'optimisation énergétique et les régulations internationales, l'article vise à encourager un développement durable de l'IA, maximisant ses bénéfices tout en réduisant ses effets négatifs.

## INTRODUCTION

### Contexte et importance du sujet : définition de l'IA responsable

L'intelligence artificielle (IA) est rapidement devenue une force de transformation dans divers secteurs, tels que la santé, l'industrie, l'éducation et les transports (Nah *et al.*, 2023 ; Osasona *et al.*, 2024 ; Lu *et al.*, 2024). Cependant, cette adoption rapide suscite des questions éthiques, sociales et environnementales cruciales. Le concept d'IA responsable, ou durable, cherche à réduire les impacts négatifs de ces technologies tout en maximisant leurs bénéfices pour la société et l'environnement. Ce terme désigne ici une approche fondée sur des principes éthiques comme la transparence, la sécurité, l'inclusion sociale et la durabilité (Tigard, 2021 ; Trocin *et al.*, 2023). Ainsi, l'IA responsable dépasse les simples préoccupations de performance technologique pour embrasser une logique de régulation et de responsabilité envers la collectivité.

### Problématique : enjeux actuels de l'IA responsable et éthique

Malgré les progrès de l'IA, cette technologie pose des défis majeurs (Baeza-Yates, 2022 ; Osasona *et al.*, 2024). Comment s'assurer que ces systèmes respectent des normes éthiques ? Comment éviter que l'IA ne renforce les inégalités ? Quelles sont les répercussions environnementales des infrastructures nécessaires ? Ces questions sont cruciales, car les décisions actuelles influenceront l'intégration de l'IA dans la société.

Les préoccupations éthiques couvrent plusieurs domaines (Baeza-Yates, 2022) : les biais dans les algorithmes (Lu *et al.*, 2024), la transparence (Borenstein et Howard, 2021), le plagiat dans l'enseignement (Sozon *et al.*, 2024), l'usage des *deepfakes*, la protection des données (Nah *et al.*, 2023), l'impact sur le travail, et la justice sociale (Elendu *et al.*, 2023). Parallèlement, les préoccupations environnementales, comme la consommation énergétique des centres de données, augmentent (UNEP, 2024).

## Objectifs de l'article

Cet article explore le concept émergent d'IA responsable, en examinant ses dimensions éthiques, sociales et environnementales. Il analyse également les cadres réglementaires existants et les initiatives visant à promouvoir une IA équitable et durable. Enfin, il propose des pistes de réflexion à destination des chercheurs, des décideurs et des entreprises pour orienter le développement d'une IA véritablement bénéfique pour la société.

## ÉTAT DE L'ART

Bien que l'IA soit souvent perçue comme une innovation récente, ses racines remontent aux années 1950, avec des pionniers tels qu'Alan Turing et John McCarthy qui ont posé les bases des théories et algorithmes utilisés aujourd'hui. Cependant, ce n'est que ces dernières décennies que l'IA a véritablement gagné en ubiquité, grâce aux avancées dans la puissance de calcul, le stockage de données et le développement des algorithmes d'apprentissage profond, comme le montre la Figure 1 (Lanzetta, 2018).

Aujourd'hui, l'IA transforme profondément de nombreux aspects de la vie moderne, qu'il s'agisse de la personnalisation des recommandations en ligne ou de systèmes médicaux autonomes capables de diagnostics. Son rôle va au-delà de l'automatisation : elle participe désormais à la prise de décision, remplaçant parfois l'intervention humaine.

## Définir une IA responsable

Voici la définition que l'on peut donner à l'IA responsable : elle regroupe l'ensemble des processus visant à améliorer l'empreinte économique, sociale et environnementale des systèmes capables de simuler des fonctions cognitives humaines pour accomplir des tâches de manière autonome ou assistée.

De nombreuses initiatives, telles que le référentiel Act Label et Labelia Labs, cherchent à évaluer et à encadrer l'impact environnemental, social et éthique des systèmes d'IA. Ces initiatives fournissent des méthodologies rigoureuses pour accompagner entreprises et institutions dans l'évaluation de leurs systèmes (ACT, 2024 ; Labelia, 2024).

## L'impact environnemental de l'IA

Les grands projets d'IA menés par Google, Amazon et Microsoft génèrent une forte pression environnementale en raison de leur empreinte écologique. Ces technologies reposent sur des centres de données de grande capacité, essentiels pour traiter les volumes massifs d'informations nécessaires aux modèles d'IA. En 2022, les infrastructures de Google ont consommé autant d'énergie que la ville d'Atlanta (Les Numériques, 2024a).

L'impact environnemental dépasse la consommation électrique : le refroidissement des serveurs exige aussi des quantités d'eau significatives, problématique dans les régions à pénurie. Microsoft a utilisé 6,4 milliards de litres d'eau en 2022, soit 34 % de plus que l'année précédente, et Microsoft, Google, et Amazon ont consommé ensemble plus de 25 milliards de litres, soulevant des inquiétudes sur la soutenabilité (RTL, 2024). Cette pression appelle des infrastructures plus durables (Thomson, 2023 ; Li *et al.*, 2023).

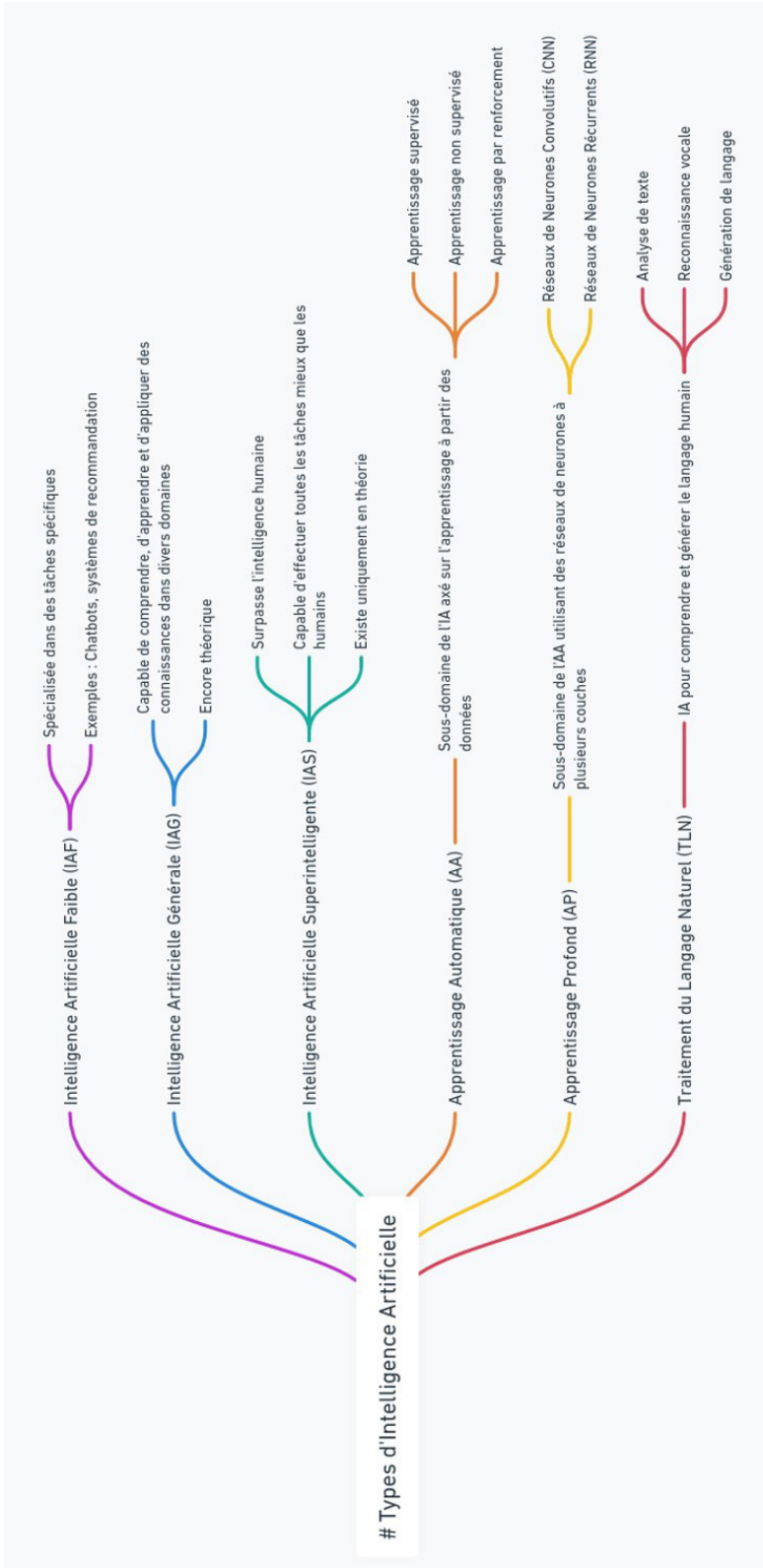


Figure 1 : Types d'intelligence artificielle (Source : auteurs).

En outre, l'IA demande d'énormes ressources énergétiques. Les algorithmes d'apprentissage profond sont très énergivores, et la formation de modèles avancés entraîne des émissions de gaz à effet de serre comparables à celles d'industries lourdes (UNEP, 2024). Par exemple, la formation de GPT-4 a requis des centaines de milliers d'heures de calcul, générant une empreinte carbone notable.

## IA et inclusion sociale et éthique

L'IA générative présente un risque d'accentuer les inégalités si son développement n'est pas inclusif. Un exemple est l'exploitation de travailleurs kenyans pour entraîner des modèles comme ChatGPT, avec des tâches répétitives et faiblement rémunérées, souvent dans des conditions précaires (Perrigo, 2024).

Cette situation souligne les déséquilibres géographiques et économiques, où les bénéficiaires de l'IA profitent aux pays développés, tandis que les tâches sous-payées sont transférées aux pays en développement. Au Kenya, certains travailleurs perçoivent seulement 1,32 dollar de l'heure pour modérer des contenus (Perrigo, 2024).

L'IA peut aussi reproduire des biais, restreignant l'accès de certaines populations à ses avantages, et toucher des secteurs comme la santé ou l'éducation (Trocin *et al.*, 2023 ; Borenstein et Howard, 2021). L'éthique devient alors cruciale : des initiatives, comme les Principes de l'OCDE et la directive IA de l'UE, visent à encadrer l'IA pour respecter les droits humains et favoriser une croissance inclusive (OCDE, 2024 ; Commission européenne, 2024 ; CNIL, 2024).

Ces efforts traduisent une volonté de prévenir les abus et de protéger les citoyens, renforçant la transparence et l'équité dans les systèmes intelligents (Karpus *et al.*, 2021 ; Xu *et al.*, 2019).

## L'économie de l'IA

L'impact de l'IA sur le marché du travail entraîne des mutations majeures dans divers secteurs. L'automatisation accrue provoque des suppressions d'emplois mais soulève aussi de nouvelles problématiques, comme l'"AI-anxiety", qui traduit une anxiété face à l'incertitude engendrée par l'automatisation (Johnson et Verdicchio, 2017). Cette inquiétude découle principalement de la crainte que les postes soient remplacés par des technologies de plus en plus autonomes.

Les professions vulnérables à cette automatisation incluent les assistants juridiques, journalistes, analystes de marché, enseignants et graphistes, confrontés à des bouleversements rapides et parfois inattendus (Wang, 2023). Par exemple, les systèmes d'IA capables de générer des articles ou d'effectuer des analyses financières menacent ces emplois d'obsolescence.

Toutefois, l'IA crée aussi des opportunités économiques : la maintenance de systèmes d'IA, la gestion des données et le développement d'applications illustrent ces nouvelles dynamiques. Ainsi, l'IA contribue à un environnement économique en mutation rapide et offre des possibilités inédites.

## PROPOSITIONS POUR UNE IA RESPONSABLE

Différentes initiatives peuvent garantir une utilisation éthique de l'IA. La Figure 2 montre qu'un avenir durable pour l'IA nécessite des efforts stratégiques. D'abord, une gouvernance ouverte doit inclure toutes les parties prenantes, notamment les populations marginalisées, et garantir des audits transparents. Ensuite, la reconversion professionnelle et la formation continue sont essentielles pour limiter l'impact de l'automatisation sur l'emploi. Il est également nécessaire de réduire la consommation énergétique en

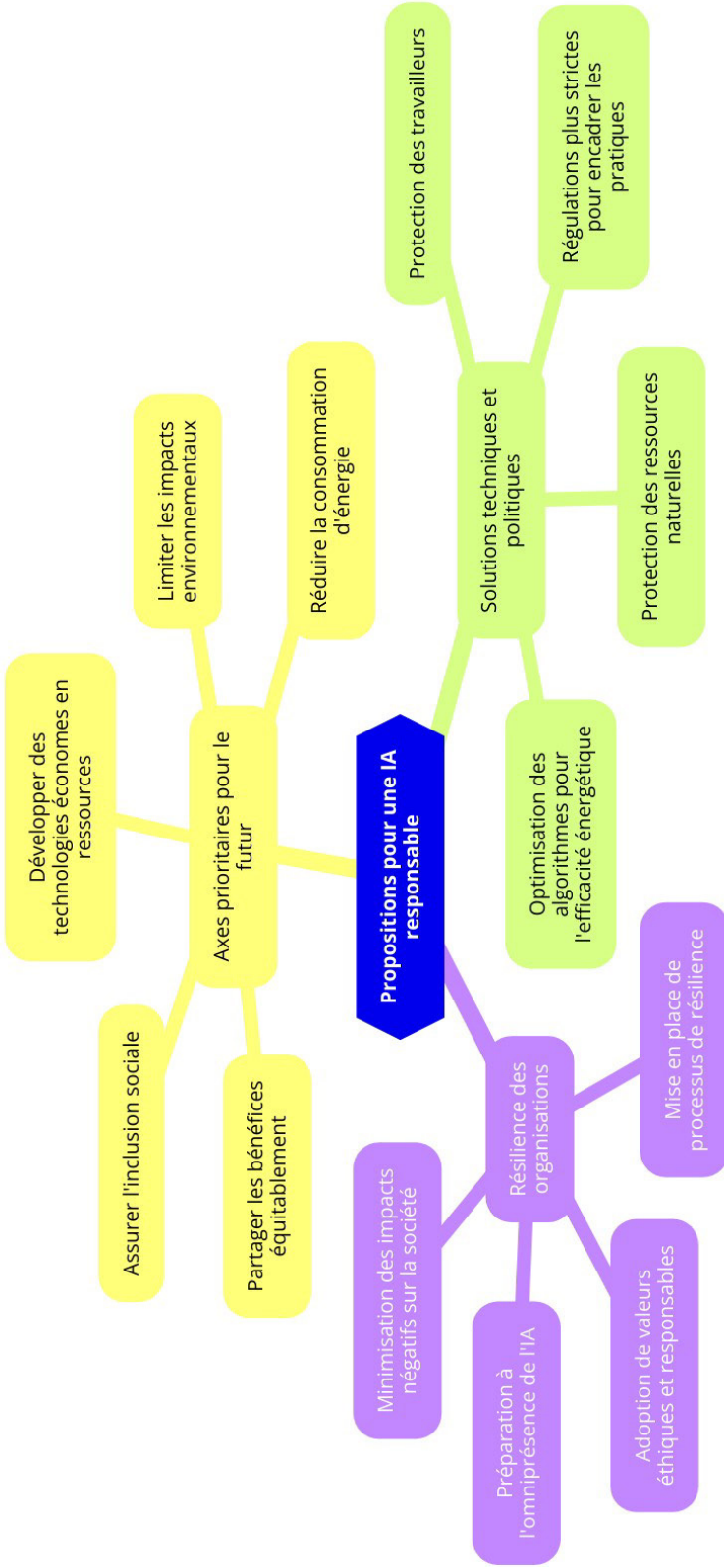


Figure 2 : Récapitulatif des propositions (Source : auteurs).

adoptant des technologies moins gourmandes en ressources et de promouvoir une répartition équitable des bénéfices technologiques pour ne laisser personne de côté.

L'optimisation des algorithmes pour réduire la consommation d'énergie devient cruciale. Les réseaux de neurones liquides sont une innovation prometteuse pour rendre l'IA plus durable et adaptable. Contrairement aux modèles classiques, ces réseaux s'ajustent en temps réel aux environnements changeants, réduisant ainsi les recalculs et la consommation énergétique. Cette capacité d'apprentissage continu permet de ne pas reformer le modèle à chaque nouvelle donnée, tout en assurant des performances stables avec des données limitées, ce qui en fait une solution efficace (Les Numériques, 2024b ; Le Big Data, 2024).

Sur le plan politique, il est essentiel de renforcer les réglementations pour mieux encadrer les pratiques des entreprises en matière de ressources et de droits des travailleurs. Les organisations doivent anticiper une société où l'IA est centrale, en adoptant des principes éthiques et responsables.

## CONCLUSION

Cet article met en lumière les défis que pose l'IA responsable, en analysant ses impacts environnementaux, sociaux et économiques, tout en suggérant des solutions concrètes pour promouvoir une IA durable et éthique. Il insiste sur la nécessité d'intégrer des principes moraux dans le développement de l'IA pour en maximiser les bénéfices et réduire ses effets négatifs. En outre, nous présentons une proposition d'IA responsable reposant sur trois piliers, la résilience des organisations, axes prioritaires pour le futur, et solutions techniques et politiques.

L'avenir de l'IA repose sur la capacité des gouvernements, entreprises et sociétés civiles à réguler cette technologie en accord avec des valeurs éthiques fondamentales. Un engagement continu est indispensable pour que l'IA contribue positivement à la société, tout en minimisant ses risques. La poursuite des recherches et la coopération entre les acteurs sont essentielles pour concevoir des réglementations et technologies favorisant une IA éthique et durable.

## RÉFÉRENCES

ACT (2024), "ACT Home", consulté le 10 octobre 2024, disponible sur : <https://act.mygreenlab.org/>

BAEZA-YATES R. (2022), "Ethical challenges in AI", ACM WSDM.

BORENSTEIN J. & HOWARD A. (2021), "Emerging challenges in AI and the need for AI ethics education", *AI Ethics*, Vol. 1, pp. 61-65.

CNIL (2024), « Entrée en vigueur du règlement européen sur l'IA : les premières questions-réponses de la CNIL », consulté le 10 octobre 2024, disponible sur : <https://www.cnil.fr/>

COMMISSION EUROPÉENNE (2024), « Législation sur l'IA », consulté le 10 octobre 2024, disponible sur : <https://digital-strategy.ec.europa.eu/>

ELENDU C. *et al.* (2023), "Ethical implications of AI and robotics in healthcare", *Medicine*, 102(50):e36671.

JOHNSON D. G. & VERDICCHIO M. (2017), "AI anxiety", *Journal of the Association for Information Science and Technology*, 68(9), pp. 2267-2270.

KARPUS J. *et al.* (2021), "Algorithm exploitation: Humans are keen to exploit benevolent AI", *iScience*, 24(6),102679.

- LABELIA (2024), “Responsible and trustworthy data science”, consulté le 10 octobre 2024, disponible sur : <https://www.labelia.org/>
- LANZETTA M. (2018), “Machine learning, deep learning, and artificial intelligence”, *In Artificial Intelligence for Autonomous Networks*.
- LE BIG DATA (2024), « Réseau de neurone liquide : l’invention du MIT va changer le monde de l’IA », consulté le 25 octobre 2024, disponible sur : <https://www.lebigdata.fr/>
- LES NUMÉRIQUES (2024a), « Le coût environnemental de l’IA fait de nouveau débat », consulté le 10 octobre 2024, disponible sur : <https://www.lesnumeriques.com/>
- LES NUMÉRIQUES (2024b), « Les réseaux de neurones liquides qui vont doper les possibilités de l’IA », consulté le 25 octobre 2024, disponible sur : <https://www.lesnumeriques.com/>
- LI P. *et al.* (2023), “Making AI less ‘thirsty’: Uncovering and addressing the water footprint of AI models”, arXiv.
- LU Q. *et al.* (2024), “Responsible AI pattern catalogue”, ACM Computing Surveys.
- NAH F. F. H. *et al.* (2023), “Generative AI and ChatGPT: Applications, challenges, and AI-human collaboration”, *Journal of Information Technology Case and Application Research*, 25(3), pp. 277-304.
- OCDE (2024), « Principes de l’IA », consulté le 10 octobre 2024, disponible sur : <https://www.oecd.org/>
- OSASONA F. *et al.* (2024), “Reviewing the ethical implications of AI in decision making”, *International Journal of Management & Entrepreneurship Research*, 6(22), pp. 322-335.
- PERRIGO B. (2024), “The \$2 per hour workers who made ChatGPT safer”, *TIME*, consulté le 10 octobre 2024, disponible sur : <https://time.com/>
- RTL (2024), « Pourquoi l’intelligence artificielle aggrave-t-elle la sécheresse ? », consulté le 10 octobre 2024, disponible sur : <https://www.rtl.fr/>
- SOZON M. *et al.* (2024), “Cheating and plagiarism in higher education”, F1000Research.
- THOMSON J. (2024), “2022 Environmental Sustainability Report”, Microsoft, consulté le 10 octobre 2024, disponible sur : <https://blogs.microsoft.com/>
- TIGARD D. W. (2021), “Responsible AI and moral responsibility”, *AI Ethics*, Vol. 1, pp. 113-117.
- TROCIN C. *et al.* (2023), “Responsible AI for digital health”, *Information Systems Frontiers*, Vol. 25, pp. 2139-2157.
- UNEP (2024), “AI has an environmental problem”, disponible sur : <https://www.unep.org/>
- WANG J. (2023), “Navigating the AI revolution”, *Advances in Economics, Management and Political Sciences*, Vol. 46, pp. 10-15.
- XU F. *et al.* (2019), “Explainable AI: A brief survey”, *In NLPCC 2019 Proceedings*, Springer.

# Démocratiser notre rapport à la technologie, un Café IA à la fois

Par Jean CATTAN

Secrétaire général du Conseil national du numérique

L'intelligence artificielle s'imisce dans tous les recoins de nos vies, souvent sans que nous en soyons pleinement conscients ou outillés pour bien définir nos usages. Entre fascination et inquiétudes, les enjeux liés à son usage nous concernent tous, qu'il s'agisse de comprendre ses applications dans nos métiers, dans l'éducation de nos enfants ou dans les services publics.

C'est pour répondre à ces défis que le Conseil national du numérique a initié Café IA : pour soutenir et encourager la constitution d'espaces d'échange, d'apprentissage et de dialogue ouverts à tous, où chacun peut s'exprimer, poser des questions, et imaginer ses propres conditions d'usage ou de non-usage de l'IA. À l'image du café, lieu informel et propice au partage de savoirs, ces rencontres visent à démocratiser la compréhension des enjeux technologiques et à renforcer une démocratie technique de proximité.

***« Je pense que le café est la partie  
la plus importante de la vie humaine »  
(J. Seinfeld).***

Mettons que vous êtes employé d'une structure professionnelle quelconque. Comme il est d'usage, certains services ne vous sont pas accessibles. Mais les outils d'IA génératives le sont eux. Est-ce que cela signifie tacitement que vous pouvez les utiliser pour exercer vos tâches ? Si oui, dans quelles conditions ? Très probablement, votre structure n'a pas répondu à ces questions et ne vous a pas entendu à ce sujet. Le résultat en est que tout aussi probablement, vous faites partie de la majorité des utilisateurs qui n'en parlent pas à leurs supérieurs. Et pourtant, tout le monde, à commencer par vos supérieurs, a un intérêt à engager cette conversation avec vous.

Imaginons maintenant que vous êtes un parent d'élève de CP et que vous retrouvez un soir un mot dans le carnet de correspondance de votre fille indiquant de flasher un QR code pour accéder à une application d'apprentissage différencié à la lecture assisté par IA. Sans nulle autre explication. Qu'est-ce que cela induit ? Que l'IA va apprendre à lire à votre fille ? Quelle grille d'appréciation de l'enseignant sur la base de quelles données transmises ? Quelle consigne d'usage ? Combien de temps par semaine ? Dans quelles conditions ? Toutes ces réponses ne vous seront probablement pas apportées d'entrée de jeu. Et pourtant vous en aurez besoin pour accompagner votre enfant.

Supposons enfin que vous êtes agent de la fonction publique. Un nouvel outil vous est indiqué comme devant être utilisé pour effectuer tel ou tel contrôle sur les citoyens. Vous ne comprenez ni les consignes qui l'instruisent, ni les résultats qui en ressortent ? Il est probable que personne n'ait pris la peine de tout vous expliquer. Et pourtant les lois du service public et le sens que vous mettez au travail en dépendent.



Nous pourrions multiplier ces situations à l'infini car elles nous concernent toutes et tous. Et encore nous n'avons pas évoqué la très grande majorité des cas où des personnes essaient un outil pour se rendre compte que, potentiellement, leur métier peut évoluer du tout au tout ou que les tâches qu'ils exécutent jour après jour peuvent être exécutées en un rien de temps. Ces situations bien familières nous montrent plusieurs choses. Tout d'abord que l'IA irrigue nos vies dans de nombreux recoins de nos existences, sans même que nous soyons des usages directs. Et que souvent, nous absorbons la situation comme si nous n'avions pas notre mot à dire ou que, de toute façon, à quoi bon ? Nous n'y comprendrions rien. Charge à nous alors de nous y faire. Pourtant, il n'en va pas nécessairement ainsi.

## POURQUOI CAFÉ IA ?

Si nous nous retrouvons dans cette situation empreinte de fatalité, c'est avant tout parce que nous n'avons pas le cadre qui nous permet d'ouvrir la boîte des possibles, des questions, des adaptations. Nous n'avons pas d'espace-temps qui nous permette d'avoir des échanges sur le pourquoi du comment de l'IA aujourd'hui. C'est précisément la raison pour laquelle le Conseil national du numérique a promu Café IA : pour offrir à l'entière de la population des espaces d'échanges, de compréhension et de mise en action sur l'IA. Pour que chacun se sente en capacité de tracer son chemin.

De 2021 à 2023, le Conseil national du numérique a mené le programme expérimental Itinéraires numériques pour mettre en lumière et en partage tous ces lieux, initiatives, collectifs et personnes qui font vivre le débat sur le numérique partout en France. Depuis la fin des confinements, toutes les semaines, une partie de l'équipe est dans au moins une localité de France pour participer, soutenir ou donner à voir une initiative de mise en partage d'éléments de débat ou de compréhension sur notre relation au numérique. Alors qu'une partie de l'équipe était dans un *Fablab*, la personne s'occupant du lieu leur a dit ce que nous soupçonnions sans pouvoir le formuler mais que probablement beaucoup savent déjà : « Vous savez ici, la machine la plus importante c'est la machine à café ».

Le point de rencontre informel incarné par le café est souvent celui qui assure la transmission des savoirs de pair à pair, l'endroit où s'exprime un savoir sur la structure ignoré par ses lieux de décisions. Ce constat rejoint d'ailleurs celui du Baromètre du numérique : c'est entre proches que pour la très grande majorité nous apprenons à nous servir des outils numériques. La plupart d'entre nous n'a pas fait d'école dédiée. Et pourtant nous arrivons à naviguer dans un environnement professionnel numérisé, plus ou moins bien il est vrai. Un des paris de Café IA est de capitaliser sur cette façon que nous avons de partager des savoirs au quotidien de manière informelle pour répondre aux besoins identifiés.

L'enjeu n'est pas de dire à untel ou une telle qu'il faut adopter l'IA, s'en servir, le faire comme-ci ou comme-ça. Non, c'est aux participants des cafés IA de le dire. Qu'ils décident de s'engager dans des parcours de formation plus approfondis ou non ou qu'ils déterminent leurs conditions d'usage ou de non-usage ne nous regarde que peu. Tout ce que nous croyons est que pour se mobiliser, comprendre, faire des choix, se positionner, dans quelque sens que ce soit, que ce soit en tant que pays ou en tant qu'individu, nous avons besoin d'échanger et de nous parler. Et que, ce faisant, nous pourrions nourrir un tissu de démocratie de proximité et d'apprentissage de pair à pair à l'échelle du pays qui soient assez robuste pour répondre aux risques sociaux qu'induisent potentiellement les technologies numériques.

## FAIRE ÉMERGER UNE RELATION DÉMOCRATIQUE À LA TECHNOLOGIE

Nous pensons au-delà que si nous pouvons faire émerger une voie populaire sur nos choix technologiques, alors nous pourrions renverser la très juste formule d'Irénée Régnauld et Yaël Benayoun « Technologie partout, démocratie nulle part » en « Démocratie partout et, quant à la technologie, c'est comme vous voulez ». Ce qui nous ramène à cette exigence d'avoir des espaces permettant d'assurer la compréhension des enjeux et des techniques mais aussi de nos contraintes et de ce qui nous anime. Des espaces qui soient aussi des espaces d'expression sur lesquels nous pouvons capitaliser pour dessiner une voie populaire du déploiement technologique. C'est peut-être d'ailleurs ce que d'aucuns pourraient appeler le progrès : définir les conditions de notre « à venir », pour reprendre l'expression de Derrida citée par Anne Alombert et Gaël Giraud, plutôt que de subir un futur imposé par quelques-uns. Et ce faisant nous pourrions alors voir émerger une culture commune de notre rapport à la technologie. Mais si nous ne faisons pas ce travail de rassemblement à la plus petite des échelles, à l'échelle où vraiment chacun a sa voix, alors nous manquerons d'avoir mis en place les moyens nécessaires pour dessiner notre avenir.

Car derrière Café IA, il y a bien une vision de la démocratie en tant que « conversation permanente », une démocratie vécue partout, tous les jours, à chaque instant, dans tout le pays et en tout lieu. Derrière Café IA, il y a cette idée que la démocratie demain ne peut fonctionner uniquement sur le fondement des institutions nationales représentatives. Qu'elle doit vivre en entreprise comme dans chacun des espaces de notre vie quotidienne.

## UN CAFÉ IA ET PUIS QUOI ?

C'est parce que l'ADN de Café IA est de nourrir la démocratie dans ses tissus les plus profonds que lorsque l'on nous pose la question des débouchés de Café IA, très clairement nous répondons qu'ils sont d'abord à la main des personnes qui font le Café IA. C'est à eux de décider de leur usage, non-usage, formation, choix d'outils, conditions d'utilisation, modes d'expression, concertations, etc. Le débouché immédiat d'un Café IA dans une école ne sera vraisemblablement pas le même que celui d'un Café IA en entreprise.

Au-delà des débouchés immédiats qui s'inscrivent dans le quotidien des personnes, les Cafés IA peuvent être le point de départ de différents exercices à l'échelle locale ou nationale. Suite à son ou ses Cafés IA, une personne peut vouloir partager son expérience, la synthèse ou certains aspects des échanges. Il pourra alors le faire dans un premier temps au travers d'un formulaire de remontées mis en ligne par le Conseil, venir partager son expérience lors d'un café en ligne qui se tient tous les jeudis de 13 h 30 à 15 h, ou encore lors d'un événement de rassemblement organisé à l'échelle locale ou nationale. Trois formes de partage donc, par écrit, en ligne ou lors d'un événement physique. L'enjeu est à chaque fois d'ouvrir toutes les portes possibles. Le fruit de ces remontées et moments d'échanges et de rencontres viendra alors alimenter la mécanique de Café IA mais surtout elles permettront de réunir et de transmettre aux dirigeants, publics comme privés, les perspectives de la population sur les développements et déploiements de l'IA.

## ET CONCRÈTEMENT, ÇA RESSEMBLE À QUOI UN CAFÉ IA ?

Le caractère protéiforme des Cafés IA vaut aussi pour son déroulé. De très nombreux formats existent allant de la simple discussion aux jeux de cartes, en passant par les ateliers d'écritures ou des séances de prise en main. Quel que soit le format choisi, le choix du format et le déroulé d'un Café IA doivent être adaptés à son public et à ses besoins.

C'est tout l'enjeu du travail d'animation. Si bien que la qualité première d'un animateur va avant tout être sa capacité à identifier les besoins de son public autant qu'à faire circuler la parole, à écouter, à relancer, à poser des questions et à mener la discussion vers un atterrissage, bien plus que de dicter un point de sortie et faire valoir sa science ou ses points de vue. Ce que nous avons appelé la recette inratable d'un Café IA tient à ça.

Selon la recette de base d'un Café IA, dans une première étape, l'animateur réunit par exemple entre 6 à 15 personnes pour 1 h ou 1 h 30 de discussions. Le matériel nécessaire est aussi simple qu'un espace calme, plutôt accueillant si possible, de quoi s'asseoir et noter quelques idées sur des bouts de papier ou des *post-it*. En début de réunion, après avoir accueilli tout le monde, l'animateur fait part aux convives de ce pourquoi ils sont réunis et des principes de la discussion. Il invite chacun à poser quelques idées, craintes, espoirs, usages, questions qu'il a en tête. Il est tout à fait possible de naviguer entre les trois registres de l'émotion, de la raison ou de la pratique. L'animateur regroupe alors les différentes idées posées sur des papiers ou *post-it* en nuages afin d'identifier un ou plusieurs thèmes prédominants sur lesquels l'échange portera en priorité. À lui de s'assurer alors que tout le monde a bien eu l'opportunité, le temps, l'espace de parler et d'être entendu. Mais aussi que les relances sont faites au bon moment, avec un degré d'ouverture ou de fermeture adapté au temps imparti ou à l'orientation des discussions. Les principes d'un Café IA – le désintéressement, la bienveillance, l'intégrité de l'information, etc. – sont toujours des guides utiles.

À l'approche de la fin de l'échange, il importe d'identifier les points d'atterrissage et d'ouverture qui s'ouvrent à eux. Assurer un échange avec telle personne responsable, suivre un processus d'apprentissage, essayer des outils, lancer un processus de concertation afin de décider d'un moratoire ou des conditions d'usage, etc. Encore une fois, cela vous appartient. Mais il est important que toutes ces perspectives soient verbalisées. Un mail récapitulatif avec des liens vers des ressources disponibles sur [cafeia.org](http://cafeia.org) sera optimal. Sachant que l'essentiel vient de se passer : des personnes se sont retrouvées pour échanger sur leurs usages, leurs espérances, leurs craintes, leurs questions pratiques. Ce qui est déjà un gain considérable en soi.

Enfin, sachez que pour vous accompagner dans votre démarche, le Conseil national du numérique rassemble les éléments de formation et de connaissances existants ou à venir, publie des modules pédagogiques élaborés, tient un recueil de recettes et formats d'animation, mobilise, sert de point de relai et vise à promouvoir à toujours plus large échelle une démarche de pollinisation des savoirs et des savoir-faire. Pour participer ou vous renseigner mais surtout pour animer et organiser, n'hésitez pas, une adresse pour cela : [cafeia.org](http://cafeia.org).

# Vers une société « IAApprenante »

Par Frédéric BARDEAU  
Président cofondateur de Simplon.co

L'intelligence artificielle est au cœur des transformations de nos sociétés. Si elle promet d'immenses opportunités, elle soulève aussi des défis éthiques, économiques, écologiques et sociaux. Comment éviter que cette révolution technologique ne laisse de côté une partie de la population ? Former le grand public à l'IA n'est pas seulement une nécessité, mais une urgence.

L'IA fait partie de notre vie quotidienne depuis des dizaines d'années. Les moteurs de recherche, les réseaux sociaux et de nombreux logiciels utilisant de l'IA dite « prédictive » peuplent discrètement nos vies personnelles et professionnelles. Depuis le 30 novembre 2022 et l'arrivée de ChatGPT, l'interface *web* qui permet d'utiliser les modèles construits par OpenAI, l'IA dite « générative » est entrée dans les conversations et les préoccupations d'une large couche de la population mondiale. Certains insistent sur les dangers et les risques, d'autres sur les promesses et les opportunités, une majorité reste encore dubitative ou éloignée de ces technologies, mais tout le monde s'accorde au moins sur un point : il faut former le grand public à l'IA. Pour autant, cela ne résout pas les questions clés suivantes : pourquoi est-il crucial de former le grand public à l'IA ? Quels sont les objectifs et les bénéfices de cette formation ? Comment former efficacement et inclusivement ? Qui doit le faire ?

Quelques mois après le dévoilement grand public de ChatGPT, une liste impressionnante d'entrepreneurs et de chercheurs, dont plusieurs Prix Turing et d'autres personnalités devenues depuis Prix Nobel – à l'exception de quelques absents très remarquables dont Yann Le Cun – signaient une déclaration comparant les risques d'extinction de l'humanité liés à l'IA à ceux des pandémies et des guerres nucléaires, et demandaient en parallèle un moratoire de 6 mois pour étudier les impacts de l'IA. Depuis, les peurs irrationnelles ont laissé la place à une course de vitesse et à la concurrence internationale effrénée malgré trois sommets sur la sécurité de l'IA et un quatrième qui va se tenir en France mais qui s'est renommé « sommet pour l'action ». Mais les apports indéniables de l'IA sur les découvertes scientifiques et techniques (des chercheurs en IA ou sur l'IA ont été récompensés de trois Prix Nobel en 2024, celui de physique, de chimie et d'économie), dans le domaine de l'éducation et de la santé, l'inclusion et l'émancipation, dans la contribution aux Objectifs de Développement durable des Nations unies sont réels. Que l'on soit enthousiaste face aux progrès de l'IA ou inquiet de ses dérives, une chose est claire : la compréhension et l'appropriation de ces technologies sont essentielles pour tous. Si comme on l'entend souvent « AI is here to stay » (« L'IA est là pour rester ») alors il faut former le plus grand nombre, et autant le faire vite, et le faire bien.

Mais que devrait-on mettre concrètement derrière cette intention ? S'agit-il d'acculturer à l'IA et de donner des repères, des connaissances et des éléments de compréhension ? Ou bien de faire monter en compétences les personnes sur l'IA, de leur apprendre certains usages, des bonnes pratiques transférables dans des contextes personnels ou professionnels et le cas échéant d'en certifier l'acquisition, ou non ? Si on en revient aux objectifs, et donc au cœur du problème, il faut donner au grand public des clés pour comprendre et

utiliser l'IA de manière efficace et responsable, développer une culture et des pratiques afin de ne pas subir cette évolution des technologies et des usages, et les impacts qui les accompagnent. Tout cela converge parfaitement avec les objectifs et les efforts qui sont en cours dans le domaine de la lutte contre l'illectronisme numérique (appelé également « illectronisme », ce phénomène touche plusieurs dizaines de millions de personnes en France et pas seulement au sein des personnes âgées et des couches les plus fragiles) au niveau des politiques publiques, des actions des entreprises et de la société civile en matière d'inclusion numérique et de promotion de la littératie et de la numératie numériques. Cette étape initiale devrait être prioritaire en termes de passage à l'échelle pour développer de la confiance chez nos concitoyens et des usages responsables, et pourrait ensuite ouvrir d'autres possibilités relevant plus d'un acte de formation (professionnelle) et pourrait donner lieu à des certifications au travers de badges numériques ou de certifications officielles enregistrées à France Compétences (Compétences Informatiques Fondamentales, Cléa Numérique, etc.).

Autre point important : pour que cette acculturation et le cas échéant cette formation soient efficaces, il faut qu'elles soient basées sur des principes forts qui seront de nature à en améliorer l'efficacité, l'impact et la popularité. La multiplicité des formats (en ligne, en présentiel, en hybride, en micro-*learning*) et des types de personnes en situation de mener ces actions (auto-formation, parents, éducateurs, formateurs, pairs, médiateurs, etc.) peut être un levier puissant pour toucher le plus grand nombre de par leur flexibilité et leur diversité, sous réserve qu'un certain nombre de conditions soient réunies. Idéalement, il faudrait que l'ensemble des actions allant dans ce sens soit référencé et hiérarchisé en fonction des objectifs qui sont visés et des catégories d'émetteurs (public, privé, non-marchand, professionnels de l'éducation). Les actions les plus intéressantes seront celles qui visent à démystifier pour comprendre, à apprendre en faisant (pédagogie active, *gamification*) pour s'approprier réellement les usages, à responsabiliser (pas d'injonction à l'usage, explications des limites, débats sur les impacts) sans polariser ni culpabiliser (« pour » ou « contre »). Les acculturations et les formations devront en outre être évaluées *a priori* et *a posteriori* par les apprenants et par la filière des spécialistes de la formation au numérique au regard du fait qu'elles sont agnostiques en termes de technologies et de services, ou que sont présentés l'ensemble des outils dans toute leur diversité, et qu'elles développent un esprit critique chez les utilisateurs en évitant les promesses fallacieuses ou les prédictions hasardeuses sur les impacts prévisibles de l'IA (qu'elles soient positives ou négatives). La tenue de débats sur les questions liées à l'IA (Cafés IA) est un format nécessaire et intéressant mais non suffisant et peut être contre-productif s'il n'est pas précédé d'une acculturation permettant de nourrir le débat avec une culture basique et des premières pratiques de ces outils.

Si Michael (2009) définit le grand public souvent de manière générique ou abstraite comme une masse homogène (« public en général »), la notion est divisée par Okamura (2016) en trois sous-groupes : le public attentif, caractérisé par un niveau élevé de connaissance et d'engagement désireux de participer aux processus de décision ; un public intéressé qui est curieux et suit les débats mais préfère laisser les experts prendre les décisions ; et le public résiduel souvent désengagé et manquant de connaissances ou de motivation pour participer à des discussions ou des processus de décision. C'est peut-être du côté de ce public résiduel composé de citoyens, de consommateurs, d'employés, d'élèves, de retraités qu'il faudrait faire porter les efforts de formation. Ce d'autant que ces publics sont souvent « hors radars » des dispositifs de formation et d'information, qu'ils seraient *a priori* les plus nombreux et que leur désintérêt pourrait également être dû à des contraintes socio-économiques, à un manque d'éducation scientifique ou à une absence de lien perçu avec les impacts des technologies qui pourraient être plus importants sur leurs propres vies personnelles et professionnelles. L'inclusion des publics périphériques, éloignés des circuits de la formation, en situation de fracture numérique, sociale ou professionnelle doit être une priorité. Les seniors, les zones rurales, les personnes fragilisées dans leur

insertion sociale et professionnelle (NEET<sup>1</sup>, personnes en situation de handicap, migrants) doivent être approchés de manière urgente, spécifique et non stigmatisante.

Quant au contenu de ces dispositifs d'acculturation et de formation, il faut véritablement donner au grand public toutes les possibilités de bien comprendre, de pratiquer et de s'interroger pour que l'appropriation responsable soit optimale. L'origine et l'histoire de l'IA, la place de l'IA dans le monde plus global des algorithmes et de l'informatique, le fonctionnement des différentes familles et services d'IA, un panorama des principaux outils et cas d'usages des IA au niveau personnel et professionnel et bien entendu des éléments scientifiques et documentés sur les impacts de l'IA (positifs et négatifs). Et au sein de ce vaste ensemble très flou qu'on appelle l'IA, il est déterminant que l'acculturation et la montée en compétences pratiques de base soient focalisées sur la dernière génération très médiatique des technologies d'IA qui se sont répandues comme une traînée de poudre : les IA génératives et notamment les modèles de langage de type ChatGPT et leurs équivalents commerciaux et *opensource*. C'est d'autant plus important que la littérature scientifique récente attribue à ces nouvelles interfaces humains-machines en langage naturel, multilingues et multimodales (capables de transformer des contenus indifféremment en texte, son, image, vidéo) des qualités d'inclusion et d'empouvoirement très importantes, équivalentes à celles qu'on prête aux *smartphones* et aux tablettes, contrastant ainsi avec le traditionnel « biais technologique » qui accompagne le plus souvent d'une exclusion forte de certains publics les déploiements de nouveaux outils.

Enfin, il est clair que former le grand public à l'IA est une responsabilité collective et non une mission qui doit être assignée aux professionnels de la médiation, de l'éducation et de l'enseignement, ou encore aux professionnels du numérique, de l'IA et de la formation au numérique qui ne doivent pas en constituer les experts désignés. L'État, les agences publiques et les collectivités locales – et pas seulement les régions qui sont en charge de l'orientation et de la formation professionnelles – seront des acteurs clés de ces dispositifs au sein de tous les territoires. Les entreprises et les corps intermédiaires (syndicats, branches professionnelles, monde associatif de l'éducation populaire, accès aux droits, secteur sanitaire et social) permettront également de toucher le grand public de manière spécifique et convergente. Favoriser les apprentissages entre pairs, dans les familles, au sein des groupes d'amis et entre les générations assureront des appropriations puissantes et rapides au sein de la population, quelles que soient les catégories socio-professionnelles et les publics spécifiques. Le rôle des médias et des médias sociaux sera clé pour diffuser cette culture et ces pratiques. L'interdisciplinarité (l'IA est une affaire qui a maille à partir avec l'informatique, la cognition, la psychologie, la linguistique mais qui intéresse aussi la philosophie, la sociologie, l'économie, la culture...) et la diversité des acteurs en situation de transmission seront garantes d'une acculturation équilibrée et profonde dans nos sociétés.

Seule une société véritablement « IAApprenante », où chacun a été en mesure de comprendre et de s'approprier l'IA, peut contribuer à des débats sereins sur l'IA et assurer une intégration citoyenne et démocratique de ces technologies dont les impacts seront durables et profonds. Cela nécessite une mobilisation massive et coordonnée de tous les acteurs et des financements spécifiquement fléchés sur ces actions (subventions, politique publique de formation, actions volontaristes des entreprises, initiatives philanthropiques) pour créer une dynamique d'apprentissage collectif et faire de l'IA ce dont nous avons tant besoin : une technologie au service de l'humanité, accessible et bénéfique pour toutes et tous.

---

<sup>1</sup> NEET, qui signifie *not in education, employment or training*, est une classification sociale d'une certaine catégorie de personnes sans emploi ne poursuivant pas d'études et ne suivant pas de formation.

# Le formateur augmenté : entre intelligence artificielle et intelligence émotionnelle

Par Brice GAILLARD  
Directeur général d'Apolearn

L'intelligence artificielle (IA) bouleverse le monde de la formation, transformant en profondeur les pratiques pédagogiques traditionnelles. Alors que 78 % des étudiants universitaires utilisent déjà l'IA dans leur cursus, les formateurs sont confrontés à un défi majeur : intégrer ces nouveaux outils sans perdre leur essence humaine. Loin de les remplacer, l'IA se présente comme un formidable levier d'augmentation, offrant des possibilités inédites de personnalisation et d'accessibilité de l'apprentissage.

Entre adaptation technologique et préservation de l'intelligence émotionnelle, le formateur de demain devient un architecte de l'apprentissage, capable de naviguer entre données numériques et accompagnement humain. Cette révolution soulève des questions majeures sur l'éthique, l'accessibilité et le rôle central de l'humain dans la transmission du savoir.

Le secteur de l'éducation et de la formation connaît une mutation radicale. L'intelligence artificielle (IA), notamment ses modèles génératifs, s'impose désormais comme un outil stratégique incontournable. Cette adoption massive traduit une transition profonde des pratiques pédagogiques, tant pour les apprenants que pour les formateurs. Une enquête mondiale du Digital Education Council en 2024 [1] révèle que 78 % des étudiants universitaires utilisent déjà l'IA dans leur cursus, symptomatique d'une appropriation technologique fulgurante. Les formateurs, quant à eux, font face à un impératif : intégrer ces outils pour ne pas être distancés par cette révolution numérique.

## UNE DIGITALISATION INACHEVÉE

La transformation numérique dans la formation, accélérée par la crise sanitaire du Covid-19, demeure inégale et parcellaire. La cohérence de cette transition reste problématique. Bien que la majorité des universités disposent aujourd'hui de plateformes numériques, force est de constater que seule une fraction minoritaire de ces outils est exploitée selon une logique pédagogique véritablement innovante, dépassant le simple partage documentaire.

Les défis persistent, exigeant une combinaison de compétences plurielles : pédagogiques, technologiques, graphiques et juridiques. Nombreux sont encore les établissements où la formation repose sur des cours magistraux rigides et des évaluations standardisées, un modèle hérité de l'ère industrielle, désormais en décalage avec les exigences contemporaines d'un apprentissage flexible et ubiquitaire.

Un paradoxe se dessine : les plateformes d'apprentissage en ligne, porteuses de potentielles révolutions pédagogiques, se transforment souvent en « cimetières à PDF », faute

d'accompagnement des formateurs. Ce constat révèle une numérisation mal maîtrisée qui, plutôt que de revitaliser les dispositifs de formation, risque de les fragiliser.

Plus préoccupant encore, les formateurs demeurent insuffisamment formés à ces nouveaux outils. À quoi servent des systèmes innovants si leurs utilisateurs n'en maîtrisent pas les fondamentaux ? Cette fracture dépasse la dimension technique pour devenir un enjeu politique et social, où l'investissement dans les ressources humaines reste largement en deçà des besoins effectifs.

## L'IA : UN LEVIER DE TRANSFORMATION NUMÉRIQUE

L'intelligence artificielle recèle un potentiel considérable pour repenser la conception, la diffusion et l'évaluation des formations. Telle un ingénieur pédagogique virtuel, l'IA peut guider les formateurs dans la scénarisation de leurs cours et l'intégration de principes comme la Conception Universelle des Apprentissages (CUA).

Par exemple, une IA pédagogique peut assister les formateurs dans l'application du modèle ADDIE, suggérant des activités alignées avec la taxonomie de Bloom pour garantir une progression cognitive optimale. Elle peut également analyser le contenu et proposer des améliorations inclusives, en appliquant les principes du Facile à Lire et à Comprendre (FALC) ou en suggérant des adaptations graphiques pour les apprenants daltoniens ou dyslexiques.

Les plateformes de Learning Management System (LMS) intègrent progressivement des fonctionnalités d'IA facilitant l'appropriation numérique. Apolearn, par exemple, intègre une IA qui aide les formateurs à scénariser leurs formations, à concevoir des modules interactifs et même à corriger les productions des apprenants.

L'efficacité de ces dispositifs est documentée : une étude de Harvard [2] révèle que les étudiants utilisant des tuteurs IA apprennent deux fois plus rapidement et manifestent une motivation accrue. Une étude de l'ISTF en 2024 [3] indique que 72 % des entreprises utilisent ou prévoient d'utiliser l'IA pour construire leurs formations.

Les outils d'intelligence artificielle offrent des possibilités prometteuses pour soutenir les apprenants en situation de handicap. Grâce à des interfaces adaptatives et des technologies innovantes comme la reconnaissance vocale, ces solutions peuvent significativement améliorer l'expérience d'apprentissage. Par exemple, les systèmes de reconnaissance vocale peuvent aider les apprenants dyslexiques en leur permettant d'interagir avec le contenu *e-learning* de manière plus fluide et personnalisée. Ces technologies ont le potentiel de rendre la formation plus accessible et inclusive.

## UN NOUVEAU PARADIGME PÉDAGOGIQUE

L'avènement de l'intelligence artificielle dans le domaine de la formation bouleverse profondément les méthodes d'enseignement traditionnelles. Cette révolution technologique ne se limite pas à la simple automatisation de tâches répétitives ; elle redéfinit les rôles des formateurs et transforme la nature même de l'apprentissage.

Le triangle pédagogique de Houssaye, longtemps considéré comme le fondement de la relation éducative, se voit aujourd'hui enrichi par l'intégration de l'IA. Cette dernière s'immisce dans les interactions entre l'apprenant, le formateur et le savoir, créant ainsi de nouvelles dynamiques d'apprentissage. L'IA intervient désormais dans les trois dimensions de ce triangle : la didactique, la pédagogie et la mathétique.



La didactique, qui concerne la structuration et la transmission du savoir, bénéficie grandement des capacités de l'IA à générer des contenus pédagogiques personnalisés. Des recherches récentes montrent que l'intégration des solutions d'IA dans l'enseignement peut significativement améliorer l'engagement des apprenants en permettant une personnalisation plus fine des approches pédagogiques.

Sur le plan pédagogique, l'IA offre de nouvelles possibilités pour gérer la progression et la motivation des apprenants. Selon une étude du Pôle Léonard de Vinci, 79 % des apprenants estiment que l'IA enrichit leur capacité à résoudre des problèmes complexes et 65 % d'entre eux affirment qu'elle augmente leur productivité et leur performance [4].

Cependant, la mathématique, qui concerne les stratégies cognitives et métacognitives de l'apprenant, reste un domaine où l'humain conserve un avantage certain. Une méta-revue publiée en 2021, analysant les données de 40 000 apprenants et 1 500 formateurs à travers le monde, souligne l'importance de la posture du formateur dans l'engagement et l'autonomie des apprenants [5].

Cette évolution du paysage éducatif soulève des questions éthiques et pratiques. La protection des données personnelles des apprenants, la transparence des algorithmes utilisés et l'équité d'accès à ces technologies sont autant de défis à relever. L'IA Act européen, dont certaines dispositions sont entrées en vigueur en juin 2024, impose de nouvelles réglementations qui façonneront l'utilisation de l'IA dans la formation à partir de 2027.

Le rôle du formateur, loin d'être obsolète, se transforme. Il devient un facilitateur, un guide qui doit maîtriser non seulement son domaine d'expertise, mais aussi les outils d'IA à sa disposition. Cette transition nécessite une formation continue et une réflexion approfondie sur les pratiques pédagogiques. Une enquête récente auprès d'enseignants européens montre que la majorité d'entre eux considère que leur rôle s'est complexifié avec l'arrivée de l'IA, tout en reconnaissant que cette évolution a enrichi leur pratique professionnelle.

L'intégration de l'IA dans la formation n'est pas sans soulever des inquiétudes. En France, les préoccupations sont particulièrement marquées, avec seulement 49 % des enseignants qui voient l'IA comme une compétence nécessaire pour l'avenir, contre 74 % aux États-Unis et 70 % au Royaume-Uni [6]. Il est donc impératif pour les institutions éducatives de communiquer clairement sur la place irremplaçable de l'humain dans le processus d'apprentissage.

Ce nouveau paradigme pédagogique, alliant expertise humaine et puissance de l'IA, promet de révolutionner l'apprentissage. Il offre la perspective d'une formation plus personnalisée, plus efficace et plus engageante.

## **RÉINVENTER LA TRANSMISSION À L'ÈRE DE L'INTELLIGENCE ARTIFICIELLE**

La révolution de l'IA dans l'éducation n'est pas simplement technologique – elle est avant tout humaine. Elle nous invite à repenser fondamentalement notre conception de l'apprentissage et du rôle du formateur.

Le formateur n'est plus un simple transmetteur de connaissances, mais un architecte de l'apprentissage, orchestrant une symphonie complexe où l'IA joue le rôle d'un instrument puissant mais non autonome.

Cette mutation profonde exige une remise en question de nos certitudes pédagogiques. L'IA n'est pas une menace pour le formateur, mais une opportunité de transcender ses limites actuelles. Elle libère le formateur des tâches répétitives pour lui permettre de se

concentrer sur ce qui fait l'essence même de son métier : l'accompagnement humain, la transmission de valeurs, et le développement de l'esprit critique.

Le formateur de demain sera celui qui saura naviguer entre deux dimensions : celle de l'intelligence artificielle et celle de l'intelligence émotionnelle. Son expertise ne résidera plus uniquement dans sa maîtrise du savoir, mais dans sa capacité à créer des ponts entre ces deux dimensions, à insuffler du sens dans l'apprentissage, et à cultiver chez ses apprenants les compétences essentiellement humaines que sont la créativité, l'empathie et la pensée complexe.

L'avenir n'appartient ni aux technophiles aveugles ni aux technophobes réfractaires, mais aux équilibristes visionnaires qui sauront harmoniser le meilleur des deux mondes. La formation de demain sera augmentée ou ne sera pas, mais cette augmentation doit être pensée, réfléchi et maîtrisée pour servir un objectif fondamental : l'épanouissement et la réussite de chaque apprenant.

## BIBLIOGRAPHIE

[1] RONG H. & CHUN C. (2024), "Digital Education Council Global AI Student Survey 2024", Digital Education Council.

[2] MILLER K. & KESTIN G. (2024), "AI Tutoring Outperforms Active Learning", Research Square.

[3] ISTF (2024), « Baromètre : les chiffres 2024 du *digital learning* », ISTF.

[4] MASSIAS J. & BERTRAND L. (2024), « L'impact des IA génératives sur les étudiants », Pôle Léonard de Vinci, RM conseil et Talan.

[5] REEVE J. & CHEON S.H. (2021), "Autonomy-supportive teaching: Its malleability, benefits, and potential to improve educational practice", *Educational Psychologist*, Vol. 56, pp. 54-77.

[6] CAPGEMINI RESEARCH INSTITUTE (2024), "Harnessing the value of generative AI, 2<sup>nd</sup> edition 2024: Top use cases across sectors".

# Pour une éducation à la pluralité des altérités et des attachements numériques à l'heure de l'intelligence artificielle générative

Par Jean-François LUCAS

Délégué général de Renaissance Numérique

Cet article soutient la nécessité d'une éducation à l'intelligence artificielle qui permette aux enfants de comprendre et d'utiliser ces dispositifs sociotechniques de manière éclairée, soit de faire des choix responsables. Cela doit avoir lieu à l'école, et en dehors de l'école. Toutefois, à l'école, et en dépit des bonnes intentions, le développement des connaissances et compétences nécessaires à appréhender et saisir les opportunités de l'intelligence artificielle se heurte au manque de moyens et de ressources. Le développement de projets et d'initiatives « hors l'école » est à ce titre intéressant.

L'article conclut en soulignant l'importance de développer les compétences émotionnelles des enfants, et des citoyens de manière plus large, afin qu'ils puissent naviguer et saisir les opportunités des altérités et des attachements numériques auxquels ils vont être confrontés, et qu'ils vont développer, notamment au regard du développement des systèmes d'intelligence artificielle générative.

## LES ENFANTS ET L'IA GÉNÉRATIVE

En janvier 2023, soit un peu plus d'un mois après sa sortie officielle (le 30 novembre 2022), ChatGPT-3 comptait plus de 100 millions d'utilisateurs actifs, faisant de cette application grand public celle ayant connu la croissance la plus rapide de l'histoire<sup>1</sup>. Dans ce contexte, les enfants ont rapidement développé des usages diversifiés de l'intelligence artificielle générative (IAgen) : synthèse et traduction d'informations, génération de contenus, aide à la réalisation aux devoirs, discussion avec des *chatbots* personnalisés, etc. Ainsi, 90 % des élèves de seconde scolarisés en région Nouvelle-Aquitaine ont déclaré l'avoir utilisée pour faire leurs devoirs<sup>2</sup>. Aux États-Unis, ce sont 47 % de 1 017 adolescents américains âgés de 13 à 18 ans qui ont déclaré utiliser des outils d'IA tels que ChatGPT, plusieurs fois par semaine, voire plus<sup>3</sup>.

---

<sup>1</sup> HU K. (2023), "ChatGPT sets record for fastest-growing user base", 2 février 2023, <https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01>

<sup>2</sup> BRUYEN C. & FIALAIRE B. (2024), « L'IA et l'avenir du service public », Rapport thématique 3 – IA et éducation. Délégation à la prospective, Sénat.

<sup>3</sup> HAUSENLOY J. & GULATI S. (2024), "The 2024 generation AI survey", <https://survey.youth-ai.org>

Au-delà des solutions dédiées, l'IA générative est déjà intégrée dans des outils éducatifs, des plateformes sociales ou encore des jeux interactifs. Si chaque dispositif sociotechnique fait l'épreuve d'usages et de représentations ambivalents<sup>4</sup>, l'IA générative doit spécifiquement nous interroger quant aux risques relatifs à la désinformation et aux hypertrucages (*deep fakes*)<sup>5</sup>, notamment à caractère pornographique. Les enfants n'y sont d'ailleurs pas étrangers : les résultats de l'étude menée par YouthAI et YouGov indique que 59 % des répondants sont préoccupés par la désinformation générée par l'IA et 58 % par les *deep fakes* (*ibid.*).

Les premiers outils et services issus des systèmes d'IA modifient déjà en profondeur nos rapports à l'information et à la connaissance, à la démocratie, mais aussi les relations et les attachements que nous tissons et entretenons avec. Si les appels à une éducation au sens critique sont nombreux, il s'agit également de favoriser une éducation au pluralisme des modes d'engagements<sup>6</sup> et des modes relationnels aux dispositifs numériques, notamment d'IA, qui peuplent désormais notre quotidien.

## ÉDUCUER AU RAPPORT À L'INTELLIGENCE ARTIFICIELLE, À L'ÉCOLE

### En théorie

Dès 2018, « l'Éducation nationale s'est emparée du sujet de l'IA dans le cadre de la stratégie du numérique pour l'éducation et de la stratégie nationale pour l'intelligence artificielle (SNIA), rattachée au volet "maîtrise de technologies numériques souveraines et sûres" du plan France 2030 »<sup>7</sup>. Pourtant, l'introduction de l'IA dans les programmes scolaires reste limitée et différenciée selon les niveaux d'enseignement. Si l'éducation aux médias et à l'information (EMI), qui est enseignée de l'école primaire au lycée, a pour objectif « d'apprendre aux élèves à lire, à décrypter l'information et l'image, à aiguïser leur esprit critique, à se forger une opinion, compétences essentielles pour exercer une citoyenneté éclairée et responsable en démocratie »<sup>8</sup>, l'enseignement explicite de l'IA n'apparaît véritablement qu'au collège et au lycée.

Au collège, l'IA est ainsi introduite *via* des modules intégrés à l'EMI<sup>9</sup> (éducation aux médias et à l'information) ou à des matières comme la technologie. Par exemple, des exercices sur Pix, la plateforme nationale d'évaluation des compétences numériques, permettent aux élèves de comprendre les biais algorithmiques. Au lycée, l'enseignement de l'IA est davantage approfondi, souvent dans des filières technologiques ou scientifiques.

<sup>4</sup> MUSSO P., COIFFIER S. & LUCAS J.-F. (2014), *Innover avec et par les imaginaires*, Manucius, 150 pages.

<sup>5</sup> RENAISSANCE NUMÉRIQUE (2019), « Les *deep fakes* : un défi pour nos démocraties », <https://www.renaissancenumerique.org/publications/deepfakes-defis-democratie>

<sup>6</sup> THEVENOT L. (2006), *L'action au pluriel. Sociologie des régimes d'engagement*, La Découverte, Paris, 312 pages.

<sup>7</sup> SABATOU A., CHAIZE P. & NARASSIGUIN C. (2024), « ChatGPT, et après ? Bilan et perspectives de l'intelligence artificielle », Office parlementaire d'évaluation des choix scientifiques et technologiques, p. 23.

<sup>8</sup> VIE PUBLIQUE (2024), « L'éducation aux médias et à l'information (EMI) face aux défis du numérique », <https://www.vie-publique.fr/eclairage/274092-leducation-aux-medias-emi-face-aux-defis-du-numerique>

<sup>9</sup> <https://www.clemi.fr/>

Des modules spécifiques, liés à l'analyse de données ou à la programmation, permettent aux élèves de manipuler des concepts d'IA. Cependant, cet enseignement reste optionnel et dépend fortement des spécialités choisies.

La promulgation de la loi visant à sécuriser et réguler l'espace numérique (SREN) pourrait marquer un tournant : en ajoutant une mention explicite à l'IA dans l'article L.312-9 du code de l'éducation, la loi introduit une obligation de sensibilisation des élèves « au bon usage des outils numériques et de l'intelligence artificielle, de tous types de contenus générés par ceux-ci et des réseaux sociaux ainsi qu'aux dérives et aux risques liés à ces outils et aux contenus générés par l'intelligence artificielle ainsi qu'à la lutte contre la désinformation »<sup>10</sup>. Cette avancée s'inscrit dans le cadre plus large des évolutions européennes à l'heure où l'IA est en cours d'intégration dans le Digital Competence Framework for Educators (DigCompEdu).

### En pratique

L'enseignement de l'IA à l'école reste inégalement réparti selon les établissements et les ressources disponibles. En France, l'EMI est un enseignement transversal qui a vocation à être assuré par tous les professeurs et les professeurs documentalistes, qui sont les « maître(s) d'œuvre de l'acquisition par les élèves d'une culture de l'information et des médias »<sup>11</sup>. Malheureusement, il souffre d'un manque de moyens, notamment pour former les enseignants et professeurs documentalistes, et d'une « absence de démarche structurée et de cadre d'usage partagé ». Conséquemment, « l'intégration des outils d'IA dans la pédagogie reste largement fonction de l'appétence de chacun »<sup>12</sup>. Enfin, les inégalités territoriales exacerbent le problème. Certains territoires bénéficient de ressources technologiques avancées, tandis que d'autres peinent à offrir des équipements de base. Ces disparités s'inscrivent dans une problématique plus large dites de « fractures numériques », liées aux inégalités sociales et économiques qui sont renseignées depuis des années<sup>13</sup>.

## ÉDQUER AU RAPPORT À L'INTELLIGENCE ARTIFICIELLE, EN DEHORS DE L'ÉCOLE

### Initiatives extrascolaires

Chaque année, en France, en dépit « des articles 28 et 29 de la Convention internationale des droits de l'enfant qui garantissent à chacun le droit à l'éducation, également rendue obligatoire par la loi Jules Ferry de 1882 »<sup>14</sup> des milliers d'enfants sont privés d'éducation : situation de handicap, personnes isolées, situation de précarité, etc. Pour pallier ce

---

<sup>10</sup> CODE DE L'ÉDUCATION (2024), « Article L.312-9 », [https://www.legifrance.gouv.fr/codes/article\\_lc/LEGIARTI000049571494](https://www.legifrance.gouv.fr/codes/article_lc/LEGIARTI000049571494)

<sup>11</sup> BULLETIN OFFICIEL DE L'ÉDUCATION NATIONALE, DE LA JEUNESSE ET DES SPORTS (2017), « Les missions des professeurs documentalistes », <https://www.education.gouv.fr/bo/17/Hebdo13/MENE1708402C.htm>

<sup>12</sup> BRUYEN C. & FIALAIRE B., *ibid.*

<sup>13</sup> ANCT, CREDOC, Université Rennes 2, CREAD-M@rsouin (2023), « La société numérique française : définir et mesurer l'éloignement numérique », <https://agence-cohesion-territoires.gouv.fr/la-societe-numerique-francaise-definir-et-mesurer-leloignement-numerique-1128>

<sup>14</sup> UNICEF(2023), « Enfants non-scolarisés : l'angle mort de l'école de la République », publié le 4 septembre 2023, <https://www.unicef.fr/article/enfants-non-scolarises-langle-mort-de-lecole-de-la-republique>

manque, de nombreux projets éducatifs ont lieu au-delà des murs de l'école. Orange, par exemple, en partenariat avec la Fédération française de Rugby, Tralalère et l'association Internet Sans Crainte ont mis en place des ateliers d'éducation au numérique pour les moins de 12 ans et moins de 14 ans<sup>15</sup>. Par extension, des sujets relatifs à l'intelligence artificielle pourrait être envisagés. Aux États-Unis par exemple, il existe des hackathons pour enfants, tels que ceux organisés par Hackathon4Kids et des événements similaires tels que FutureHacks, qui visent à initier les jeunes à des concepts technologiques et à la programmation, y compris l'intelligence artificielle.

## La parentalité numérique

« Parce qu'ils sont les premiers éducateurs confrontés aux questions pratiques posées par l'omniprésence du numérique dans notre quotidien, les parents sont des partenaires essentiels de l'éducation au numérique »<sup>16</sup>. La parentalité numérique se définit comme l'ensemble des pratiques, des compétences, et des attitudes que les parents adoptent pour accompagner leurs enfants dans l'usage des technologies numériques (interactions des enfants avec les écrans, protection du bien-être numérique, développement d'un usage responsable et créatif des technologies...) <sup>17</sup>.

Depuis plusieurs années, de nombreuses initiatives visent à sensibiliser et accompagner les familles dans cette mission. Les guides pédagogiques, tels que ceux publiés par le Centre pour l'éducation aux médias et à l'information (CLEMI), les parcours thématiques comme ceux de Pix dédiés aux parents, ainsi que des ateliers, formations ou événements comme Super Demain, à Lyon, constituent des ressources précieuses pour les parents. Ces dispositifs offrent des cadres structurés pour aborder des enjeux comme le temps d'écran, la désinformation, ou encore la sécurité en ligne. Par ailleurs, en 2024, 62 % des structures de médiation numérique proposent des activités pour accompagner le développement de la parentalité et l'éducation au/par le numérique<sup>18</sup>.

Ces exemples témoignent d'une reconnaissance croissante du rôle clé des parents dans l'éducation numérique. Toutefois, cela nécessite des efforts concertés pour inclure toutes les catégories de foyers et pallier les disparités structurelles qui persistent. Car, malgré cette offre abondante, une étude de 2022 révèle qu'environ un parent sur deux ne se sentait pas, ou pas suffisamment, accompagné dans l'encadrement des pratiques numériques<sup>19</sup>. Or, si les causes de ce déficit d'accompagnement sont multiples (disparités socio-économiques et culturelles, manque de temps ou d'accès aux ressources, ainsi qu'une méconnaissance des enjeux numériques), ce sont encore les familles qui ont un capital social, économique et culturel le moins élevé, qui sont les moins disposées à saisir et se saisir des ressources et opportunités numériques, de manière proactive et efficace.

<sup>15</sup> MUZEAU S. (2022), « École de Rugby : Éducation au numérique avec Orange », 9 février 2022, <https://www.ffr.fr/actualites/partenerariat/ecole-de-rugby-education-au-numerique-avec-orange>

<sup>16</sup> Fréquence Écoles, Futur composé (2024), « Repenser l'éducation au numérique. Pour développer des méthodes d'intervention et de conception émancipatrices », octobre 2024.

<sup>17</sup> LIVINGSTONE S. & BLUM-ROSS A. (2020), *Parenting for a digital future: How hopes and fears about technology shape children's lives*, Oxford University Press.

<sup>18</sup> LA MEDNUM (2024), « Observatoire de l'inclusion numérique. Réalités et perspectives du secteur de la médiation numérique en 2024 ».

<sup>19</sup> Observatoire de la Parentalité et de l'Éducation Numérique (Open), Union Nationale des Associations Familiales (Unaf) (2002), Étude « Parents, enfants et numérique ».

## ÉDUIQUER À LA PLURALITÉ DES ALTÉRITÉS ET DES ATTACHEMENTS NUMÉRIQUES

### Former les citoyens de demain

« L'exigence de la transition numérique semble admise comme une évidence mais il est moins fréquent d'en décliner les dimensions et surtout de les traduire en politiques d'éducation au numérique. Le risque est grand de lister un certain nombre d'activités à offrir dans une classe sans pour autant énoncer les principes qui guident cette éducation, sans mettre en lumière la vision cohérente de ces enseignements (...) Car il ne s'agit pas d'empiler un savoir supplémentaire mais bien de former des citoyennes et des citoyens et donc des actrices et acteurs à chaque instant de choix informés et responsables »<sup>20</sup>.

Former les citoyens de demain exige d'atteindre un certain niveau de compréhension de la technologie (comment cela fonctionne), des effets socioculturels (quels en sont les effets) et des manières d'utiliser les dispositifs (comment sont-ils utilisés)<sup>21</sup>. Or, si « les programmes à l'échelle internationale (...) ont ainsi bien intégré la nécessité d'ajouter une compréhension des enjeux sociaux à l'éducation au numérique », ils risquent « lorsqu'ils sont pensés en termes d'effets ou d'impacts, sans tenir compte de toutes les avancées des sciences sociales en matière d'analyse sociotechnique (Akrich, Callon, Latour, 2006) (...) de juxtaposer des savoirs sans donner les moyens de les combiner et de les associer à une expérience pratique de choix responsables »<sup>22</sup>.

Par ailleurs, d'aucuns considèrent que le développement d'une pensée critique et algorithmique préparerait les jeunes à résoudre des problèmes complexes dans divers domaines, au-delà de la programmation<sup>23</sup>, visant ainsi le développement de compétences afin de s'adapter à un avenir complexe, et incertain. Cette exploration d'un inconnu sans crainte passe aussi par un système éducatif véritablement inclusif qui valorise la créativité et l'expression personnelle des élèves, leur donnant confiance dans les choix qu'ils font, ce que François Taddei rappelle quand il évoque par exemple l'importance de concevoir une « société apprenante » grâce à une éducation qui favorise l'expérimentation et la projection dans des scénarios nouveaux<sup>24</sup>.

Enfin, former les citoyens de demain, c'est aussi les projeter dans un scénario où ils peuvent se « passer du numérique ». D'abord parce que toute technologie est faillible. Ensuite, dans le cas de l'IAgen, on observe qu'elle peut servir de « béquille mentale », avec le risque d'appauvrir les élèves et étudiants en techniques mentales<sup>25</sup>. En ce sens, Daniel Andler propose ainsi d'adopter le « bilinguisme éducatif »<sup>26</sup> pour deux raisons principales.

---

<sup>20</sup> BOULLIER D., CHESSEL-LAZZAROTTO F., LIEGEOIS G. *et al.* (2023), « Un modèle pluraliste d'éducation numérique, l'expérience du canton de Vaud en Suisse », *Distances et médiations des savoirs*, 43, <http://journals.openedition.org/dms/9344>

<sup>21</sup> "Bildung in einer digitalen vernetzten Welt" : <https://dagstuhl.gi.de>

<sup>22</sup> BOULLIER *et al.*, *ibid.*

<sup>23</sup> TOURETSKY D. S., GARDNER-McCUNE C, MARTIN F. & SEEHORN D. (2019), "Envisioning AI for K-12: What should every child know about AI?", *Proceedings of the AAAI Conference on Artificial Intelligence*, 33(01), 9795-9799. <https://doi.org/10.1609/aaai.v33i01.33019795>

<sup>24</sup> TADDEI F. (2018), *Apprendre au XXI<sup>e</sup> siècle*, Calmann Lévy.

<sup>25</sup> RENAISSANCE NUMÉRIQUE (2024), « Intelligence artificielle et intelligence humaine, avec Daniel Andler », <https://www.renaissancenumerique.org/publications/daniel-andler-ia>

<sup>26</sup> ANDLER D. (2023), *Intelligence artificielle, intelligence humaine : la double énigme*, 432 pages, Gallimard.

La première pour être autant capable de travailler avec les technologies que de s'en passer. La seconde, pour que cela permette aux individus de dépasser « la peur de la page blanche » en leur redonnant confiance dans leur capacité d'expression et de créativité.

## Littératie en IA

Pour permettre aux citoyens de comprendre les impacts tangibles et possibles de l'IA, ses possibilités et limites, son utilité ou encore son efficacité pour le bien commun, il est nécessaire que chacun puisse atteindre « un certain niveau de compétence en matière d'IA, notamment en termes de connaissances, de compréhension, d'aptitudes et d'orientation des valeurs »<sup>27</sup>. Ce « package » de savoirs, savoir-faire ou encore savoir-être est souvent désigné par le terme de « littératie »<sup>28</sup>, bien que celui-ci recouvre des périmètres et approches diverses selon les contextes et les époques<sup>29</sup>.

Or, si la littératie numérique pouvait tendre, d'une certaine manière, à définir les bases nécessaires qu'un individu doit acquérir pour utiliser des équipements, logiciels et applications, et saisir leurs opportunités pour « agir avec » et en extraire des bénéfices (en termes d'information, de connaissance, de socialisation, etc.), la littératie algorithmique s'intéresse de manière plus spécifique aux algorithmes et données qui sous-tendent les systèmes automatisés, souvent pour en comprendre les logiques sous-jacentes. Quant à la littératie en matière d'IA, l'Unesco considère par exemple qu'elle comprend « à la fois la littératie relativement aux données, ou la capacité de comprendre comment l'IA collecte, nettoie, manipule et analyse les données, et la littératie concernant les algorithmes, ou la capacité de comprendre comment les algorithmes d'IA trouvent des modèles et des connexions dans les données, qui peuvent être utilisés pour les interactions homme-machine. Il s'agit d'une tentative d'encadrer la portée, la structure et les principales catégories du domaine émergent de l'alphabetisation en IA »<sup>30</sup>.

À l'image de cette définition, les différentes tentatives de circonscription de la notion de littératie en IA sont souvent très englobantes, reprenant régulièrement des aspects des littératies existantes. Toutefois, il est possible d'identifier dans l'émergence de ces définitions la préoccupation grandissante à développer des compétences émotionnelles.

## Les compétences émotionnelles à l'ère de l'intelligence artificielle

Les compétences émotionnelles, définies comme la capacité à identifier, comprendre, réguler et utiliser les émotions de manière constructive<sup>31</sup>, prennent une importance accrue dans le contexte du développement des systèmes d'IA générative. Car, au-delà de la volonté de leur conférer des capacités à reconnaître et à simuler des réponses

<sup>27</sup> UNESCO (2023), « Enseigner l'intelligence artificielle au primaire et au secondaire. Une cartographie des programmes validés par les gouvernements », p. 13, [https://unesdoc.unesco.org/ark:/48223/pf0000380602\\_fre](https://unesdoc.unesco.org/ark:/48223/pf0000380602_fre)

<sup>28</sup> Terme emprunté de l'anglais qui n'a pas d'équivalent officiel français.

<sup>29</sup> LE DEUFF O. (2012), « Littératies informationnelles, médiatiques et numériques : de la concurrence à la convergence ? », *Étude de communication*, 38 | 2012, pp. 131-147, <https://journals.openedition.org/edc/3411>

<sup>30</sup> UNESCO, *ibid.*, p. 11.

<sup>31</sup> MAYER J.D. & SALOVEY P. (1997), "What is emotional intelligence?", in P. Salovey & D. Sluyter (Eds.), *Emotional development and emotional intelligence: Educational implications* (pp. 3-31), New York, Basic Books.



émotionnelles afin de personnaliser les interactions avec les utilisateurs (en se basant par exemple sur l'analyse des expressions faciales, des tonalités vocales ou encore des signaux physiologiques – ce qui n'est pas sans poser des questions éthiques liés à la manipulation et à l'exploitation des émotions humaines<sup>32</sup>), il s'agit de favoriser l'apprentissage de compétences telles que l'empathie ou la gestion des conflits dans des contextes réels<sup>33</sup>, mais aussi de prévenir la possible standardisation des réponses émotionnelles<sup>34</sup>.

Plus généralement, il s'agit de préparer les individus à la multiplicité des modes relationnels qui se tissent, et qu'ils vont tisser avec les objets technologiques, dont chacun est caractérisé par un imaginaire relationnel et narratif, un régime psychologique ou encore des inducteurs relationnels spécifiques<sup>35</sup>. L'éducation au pluralisme des altérités technologiques et des attachements<sup>36, 37</sup> apparaît dès lors comme une priorité pour que les citoyens d'aujourd'hui et de demain puissent être acteurs.

---

<sup>32</sup> McSTAY A. (2018), *Emotional AI: The rise of empathic media*, SAGE Publications, 248 pages.

<sup>33</sup> LIVINGSTONE S. & SEFTON-GREEN J. (2016), *The class: Living and learning in the digital age*, New York, New York University Press.

<sup>34</sup> HOLMES W., BIALIK M. & FADEL C. (2019), *Artificial intelligence in education: Promises and implications for teaching and learning*, Boston, Center for Curriculum Redesign.

<sup>35</sup> MUSSO P., COIFFIER S. & LUCAS J.-F., *ibid.*

<sup>36</sup> HENNION A. (1993), *La passion musicale. Une sociologie de la médiation*, Paris, Métailié.

<sup>37</sup> BOULLIER D. (2013), « L'attention : un bien rare en quête de mesures », *Sciences de la société : Les cahiers du LERASS*, 87, pp. 128-145.

# La création artistique à l'épreuve de l'intelligence artificielle

Par Alain ASSOULINE

Ex-président du réseau des écoles WebForce3

Dans notre rapport présent et futur à l'IA, la question de la création artistique concentre probablement la plupart des problèmes qu'il faudrait se poser.

La danse, les arts plastiques, la littérature, la musique et tous les autres arts existent depuis que le genre humain est apparu. Ce sont les différentes expressions des émotions et des relations humaines. C'est la part la plus profondément humaine d'une culture.

Qu'est-ce que l'IA va faire de cela ? Que doit-on lui permettre de faire ? Avons-nous encore le choix ?

Une IA au service de l'humanité, des femmes et des hommes, peut augmenter et peut-être magnifier encore plus les arts. Une IA au service de logiques algorithmiques d'efficacité, de rentabilité ou de pouvoirs, réduirait l'art à un outil de communication sans doute très performant mais en annihilant la diversité, l'originalité, bref, le génie humain.

Il est urgent de mettre en place des outils de régulation de l'utilisation de l'IA dans tous les domaines artistiques.

Dans notre rapport présent et futur à l'IA, la question de la création artistique concentre probablement la plupart des problèmes qu'il faudrait se poser.

La danse, les arts plastiques, la littérature, la musique et tous les autres arts existent depuis que le genre humain est apparu. Ce sont les différentes expressions des émotions, de l'inconscient et des relations humaines. C'est la part la plus profondément humaine d'une culture.

Qu'est-ce que l'IA va faire de cela ? Que doit-on lui permettre de faire ? Avons-nous encore le choix ?

Beaucoup de choses ont déjà été écrites sur l'IA et la création artistique sous l'angle des outils, de l'aide qu'elle peut apporter, et des nouveaux enjeux pour les droits d'auteur. Dans ce domaine, comme dans tous les autres, la tendance est de parler de l'IA comme d'un instrument, certes très puissant, mais un instrument comme l'a été le numérique, qu'il suffit de maîtriser pour ne pas se laisser dépasser.

L'auteur de ces lignes a été un des pionniers de l'introduction du numérique dans la chaîne de création graphique. À l'époque (les années 1990), les plus méfiants étaient les directeurs artistiques qui affirmaient que le numérique ne permettait pas de produire des images et des typographies avec la même qualité que les outils existants (photocomposition, photogravure...) et qu'il enserrerait la créativité dans un carcan. La preuve a été faite depuis que c'est le contraire qui s'est produit. Si le numérique a fait disparaître nombre de métiers de la chaîne graphique, les gagnants ont été les directeurs artistiques

à qui le numérique a donné une maîtrise complète de leurs projets en leur ouvrant des perspectives créatives plus vastes.

Alors, l'IA n'est-elle finalement qu'un nouveau développement du numérique ?

De plus en plus nombreux sont les meilleurs experts de l'IA qui nous alertent : l'IA est bien plus qu'un nouvel outil, elle introduit un changement de paradigme fondamental qui va bouleverser l'activité humaine. Nous prenons de plus en plus conscience que ce n'est pas simplement un nouvel outil, une nouvelle technologie plus performante mais qui resterait sous contrôle humain. Et même quand certains semblent la maîtriser, combien sont-ils, et pour combien de temps ? Quid de tous les autres qui sont l'écrasante majorité ?

Si l'IA peut remplacer l'humain dans bien des domaines, que restera-t-il à l'humain ?

Peut-être l'art...

Mais en est-on bien sûr ? Cela dépendra-t-il seulement de ce que nous en ferons et des règles que nous fixerons ?

Avez-vous essayé de faire une image avec l'IA ? On rentre un *prompt* bien conçu (il y a déjà des outils pour ça) et on obtient un résultat bluffant. Plusieurs même, immédiatement. Mais tous se ressemblent et si on veut autre chose, on peut demander un style « à la manière de » tel ou tel artiste, ou avec tel ou tel outil, crayon, fusain, peinture ou appareil photo de telle marque avec telle focale. C'est incroyable, le résultat est toujours rapide et répond aux demandes, rien n'est laissé au hasard.

Au hasard... Le hasard qui surprend l'artiste, cette chose lumineuse qui sort de lui, cette inspiration soudaine où son inconscient rencontre sa conscience, c'est pourtant cela qui est unique !

Est-ce que l'IA mise au service du créateur peut créer ce moment, cette rencontre ?

L'IA a-t-elle un inconscient ? Non, bien sûr. L'IA est intelligente, elle le sera de plus en plus, et dépassera sans doute l'intelligence humaine. Mais elle n'a ni conscience, ni inconscient. Elle est d'une froide rationalité, et c'est bien là le problème quand on parle de création artistique.

Peut-elle aider l'humain à trouver l'inspiration, à exprimer ses émotions ? Peut-être, sûrement même. Mais attention, ce que propose l'artiste est issu d'une intuition singulière et subjective, alors que l'IA nous livre la résultante complexe d'un certain nombre de choix statistiques. Ce sont deux productions de natures radicalement différentes.

Bien sûr, l'artiste monte toujours sur les épaules de tous ceux qui l'ont précédé, il subit l'influence de son environnement et de ses rencontres (Picasso et l'art africain, les surréalistes et l'art précolombien), consciemment ou inconsciemment. Il travaille beaucoup pour atteindre son sentiment de perfection. Mais à un moment donné, fort de ce patrimoine, de ce savoir-faire, de cette richesse culturelle, il crée une rupture, expression à la fois de sa personnalité, de ses émotions et de son époque.

L'IA produit des images en s'appuyant sur une base de données de millions d'images et en répondant à une demande précise. Saura-t-elle créer cette rupture ou, au contraire, va-t-elle produire et reproduire toujours ce qui répond au dénominateur commun de tout ce qui lui est demandé et de tout ce qu'elle retrouve dans les bases de données et les réseaux sociaux ?

Si l'IA utilise les réseaux sociaux dont les algorithmes réduisent, comme le dit Harari dans *Nexus*, « l'éventail des émotions humaines – haine, amour, indignation, joie, confusion – à une seule catégorie fourre-tout : l'engagement », quel art produira-t-elle ?

Le cerveau humain a tendance à chercher la ligne de moindre résistance. L'IA alliée à un cerveau humain fainéant, qui cherche la pente la plus douce pour produire le moins

d'efforts possible, risque de dominer le paysage culturel. On verra alors tout notre environnement visuel (affiches, images, vidéos, publicités, mais aussi architectures, etc.) et sonore (musiques, bruits, alertes) se ressembler comme deux gouttes d'eau et les créatifs (*design* graphique, directeurs artistiques, créateurs de sons, etc.) soumis à rude concurrence. Ce sera aussi le cas des créations plastiques, littéraires ou musicales qui deviendront hyper efficaces pour plaire au plus grand nombre. Et on dira que c'est de l'art, l'avènement d'un art démocratique accessible à tous ! Mais que restera-t-il de l'artiste ? Du génie humain de la création ?

Des recherches ont permis de créer des IA spécifiques complexes qui sont capables de créations originales et inattendues. Ces IA sont inventées pour l'usage d'un artiste ou d'un groupe d'artistes, à grands frais et avec des moyens considérables comparés au prix d'une plume, d'un pinceau ou d'un instrument de musique. Tous les artistes y auront-ils accès ? En réalité, 99 % des artistes auront accès à Midjourney et à tous les avatars d'outils de création d'images générés par Chat GPT. Et c'est ce qui dominera le paysage culturel.

Il serait alors intéressant d'explorer la ressemblance de l'art produit par l'IA avec ce que furent en leur temps le réalisme socialiste ou l'art nazi : un art normé par une conception idéologique totalitaire, pour exprimer une réalité qui ne doit pas être remise en cause, politiquement correct, avec un culte de la beauté défini une fois pour toutes.

Dans sa publicité, un producteur de méthode pour créer des prompts efficaces dit : « Je fixais mon écran, bouche bée. L'image que je venais de générer avec l'IA était... parfaite ».

Mais qu'est-ce qu'une image « parfaite » ? L'artiste est un artisan qui fabrique de l'originalité. Il est souvent fabricant en même temps que créateur. L'IA peut jouer le rôle de partenaire de travail pour augmenter sa créativité, pousser son imagination et ses limites artistiques, lui éviter des tâches longues et fastidieuses. L'avenir de l'art pourrait être une fusion parfaite entre imagination humaine et efficacité de l'intelligence artificielle. Mais si l'intelligence artificielle réalisait la production parfaite de ce que l'humain imaginerait, que resterait-il à produire pour l'humain ? N'est-ce pas là que l'art se cache ? Au moment de la production, dans les imperfections et les doutes ?

Il est clair que l'IA est un accélérateur pour la création. Les artistes peuvent y avoir recours dans leur travail pour stimuler leur créativité, augmenter leur inspiration, découvrir d'autres techniques. Wayne McGregor le dit : « Dans la danse, l'IA nous aide à bousculer nos habitudes ». S'en tiendra-t-elle là ?

Alors à quoi faut-il être attentif pour sauver l'art, comme production humaine, expression de son âme, moment de rencontre de sa conscience avec son inconscient ?

Je lance ici quelques pistes qui peuvent s'avérer impraticables et même vaines, mais il faut réfléchir ensemble d'urgence à ces sujets car sans régulation, tout va très vite et ce qu'on observe déjà est inquiétant.

D'abord, il faudrait développer plusieurs IA pour chaque domaine artistique, chacune avec des approches différentes en prenant soin de définir les règles déontologiques et le but qu'on lui fixe de manière qu'elle soit alignée avec les objectifs ultimes de la préservation d'un art humain.

Ensuite, il faudrait réguler la façon dont elle se connecte aux réseaux sociaux, comment elle évite le piège des algorithmes basés sur l'engagement, ce qu'elle va y chercher et ce qu'elle en fait. On peut ici mettre en place une autorégulation par l'IA elle-même.

Et il faudra y ajouter une régulation par la communauté des artistes qui devrait garantir à la fois une utilisation raisonnée de l'IA et surveiller ses productions. Les démocraties se sont toujours efforcées de ne pas réglementer la création artistique. « Toute licence en art ». Ce sont les dictatures qui ont voulu imposer un art officiel. Mais devant le risque

d'un art IA qui ressemblerait tellement à ce type d'art officiel, ne faut-il pas réglementer l'utilisation de l'IA pour produire de l'art ? Attention à une sorte d'IA populiste qui deviendrait la seule représentante de l'art populaire. Ce sont les artistes, même minoritaires, qui doivent représenter l'art.

Et puis surtout, il faudra non pas moquer mais protéger et encourager les artistes qui n'utiliseront pas l'intelligence artificielle...

Une IA au service de l'humanité, des femmes et des hommes, peut augmenter et peut-être magnifier encore plus les arts. Une IA au service de logiques algorithmiques d'efficacité, de rentabilité, de pouvoirs ou d'idéologies, réduirait l'art à un outil de communication sans doute très performant mais en annihilant la diversité, l'originalité, bref, le génie humain.

Il est urgent de mettre en place des outils de régulation de l'utilisation de l'IA dans tous les domaines artistiques.

Dans ce domaine, tout reste à débattre, définir et concevoir...

# Les impacts de l'IA sur l'environnement

Par **Frédéric GARCIA**

Université de Toulouse, INRAE, Unité MIAT

Et **Sophie SCHBATH**

Université Paris-Saclay, INRAE, Unité MaIAGE

Avec le développement rapide de l'intelligence artificielle (IA), en particulier l'IA générative, le lien entre IA et environnement suscite de nombreuses questions.

D'un côté, l'IA apparaît comme une technologie clé au service de la transition écologique, décuplant nos capacités prédictives et décisionnelles pour une utilisation plus durable des ressources naturelles et une compréhension plus fine des systèmes sociotechniques et environnementaux.

De l'autre, l'IA mobilise de nombreux équipements informatiques très gourmands en énergie, eau et métaux rares, ce qui engendre des impacts dévastateurs sur l'environnement lui-même, tant en termes d'épuisement des ressources, de perte de biodiversité que d'émission de gaz à effet de serre.

Vu l'explosion actuelle de l'usage de l'IA par le grand public, nous sommes en droit de nous demander si les appels unanimes au développement d'IA « frugales » suffiront à ce que les bénéfices environnementaux de l'IA dépassent ses dommages environnementaux.

## INTRODUCTION

Avec le développement rapide de l'intelligence artificielle (IA) et en particulier la vague actuelle de l'IA générative et des grands modèles de langage (LLM), la question du lien de l'IA à l'environnement est l'objet de nombreux travaux académiques et rapports d'experts (Crawford, 2024 ; OECD, 2022 ; Tatot et Vermot Desroches, 2024). L'IA apparaît en effet comme au cœur de la nouvelle révolution technologique qui impacte ou impactera prochainement l'ensemble de nos sociétés. Elle amplifie les transformations engagées au XX<sup>e</sup> siècle avec l'avènement de l'ère du numérique, en rendant possible le traitement intelligent d'informations numérisées à des échelles de performance et pour un spectre de domaines encore inimaginables il y a peu (Suleyman et Bhaskar, 2023).

L'IA est ainsi régulièrement présentée comme une des technologies clés au service de la transition écologique. En décuplant les capacités des technologies du numérique à traiter efficacement des problèmes de plus en plus complexes, l'IA offre l'espoir d'une technologie salvatrice, qui saura guider nos choix vers de nouveaux modes de production, de nouveaux usages, de nouvelles gouvernances, pour un développement durable qui respecte les limites planétaires.

Néanmoins, cette vision techno-positiviste de l'IA est aujourd'hui contrebalancée par des analyses de l'impact environnemental, actuel ou potentiel, de cette technologie. On sait que les précédentes révolutions technologiques associées à l'invention de la machine à vapeur ou de l'électricité ont été la source directe ou indirecte d'immenses transforma-

tions sociales et environnementales à l'échelle de la Terre, dont nous prenons pleinement conscience aujourd'hui avec les crises écologiques et environnementales que nous traversons. Forts de notre meilleure compréhension des déterminants de ces crises, il est alors légitime de chercher à anticiper l'ampleur à venir des impacts négatifs de l'IA sur l'environnement afin de les contrer.

## LES PROMESSES DE L'IA

L'IA est à la base une discipline scientifique construite au milieu du XX<sup>e</sup> siècle sur le projet ambitieux de concevoir les méthodes et algorithmes associés permettant la mise en œuvre de machines informatiques « intelligentes », c'est-à-dire comme le proposait Minsky, « capables de faire des choses qui nécessiteraient de l'intelligence si elles étaient faites par des humains » (Konieczny et Prade, 2020). L'histoire de cette discipline est remplie de vagues successives traduisant enthousiasmes et parfois déceptions de la société quant aux potentialités des technologies qui en découleraient (les réseaux neuronaux, les systèmes experts, l'ingénierie des connaissances, l'apprentissage automatique...).

Les algorithmes issus de l'IA traditionnelle, principalement axée sur les représentations symboliques et le raisonnement à base d'heuristiques, se retrouvent au cœur de nombreux systèmes industriels, que ce soit en phase de conception ou de pilotage (optimisation combinatoire, ordonnancement et planification, diagnostic, plus court chemin dans les graphes, etc.). Toutefois, c'est aujourd'hui le développement et la diffusion extrêmement rapide au sein de la société des techniques d'IA à base d'apprentissage profond et d'IA générative, basées elles sur l'emploi des réseaux de neurones artificiels, qui impressionnent. Cette IA neuronale, qui repose sur l'exploitation massive de données et le calcul intensif, apparaît en effet comme très efficace pour traiter des tâches de reconnaissance et de simulation dans les domaines de l'image et de la vidéo, du son, de la parole et du langage, ainsi que pour des tâches de prédiction (Microsoft Aurora) et de décision (DeepMind AlphaGo). Cette technologie IA, qui pourrait encore profiter dans un futur proche de l'intégration de ses deux courants symbolique et neuronal (Cervoni *et al.*, 2024), laisse alors entrevoir à la société la possibilité de disposer de moyens de calcul et de raisonnement qui nous permettraient de combattre au mieux la crise environnementale actuelle, causée et entretenue par notre mode de développement.

C'est ainsi que des solutions à base d'IA sont actuellement développées et mises en œuvre par les entreprises privées et les pouvoirs publics de nombreux États pour favoriser des activités de production, de distribution ou de services plus respectueuses de l'environnement. Par exemple, grâce aux techniques IA de modélisation, d'optimisation et d'apprentissage, une meilleure connaissance du climat et de meilleures prédictions météorologiques permettent de mieux prédire en temps réel les besoins en énergie des foyers et des industries ainsi que les productions d'énergie renouvelable, informations indispensables pour la bonne gestion des réseaux. Elles permettent également d'optimiser la gestion de l'énergie dans les bâtiments, usines et entrepôts. Plus globalement, les capacités prédictives de l'IA nous permettent de mieux comprendre les tendances, les régularités, les structures des systèmes sociotechniques et environnementaux. Associées aux méthodes d'optimisation et de raisonnement dans l'incertain que développe l'IA, elles permettent de prendre des décisions rationnelles pour une utilisation plus efficace de l'énergie, des matières premières, de l'eau et de l'ensemble des ressources critiques. Ces mêmes capacités prédictives et décisionnelles permettent à l'IA de proposer des solutions pour mieux calculer l'empreinte carbone des produits et activités, concevoir des produits éco-responsables, réduire les pollutions et les émissions de CO<sub>2</sub>, en particulier dans le transport et la logistique ou encore en agriculture (Bellon Maurel *et al.*, 2022). Enfin, une dimension également importante est celle de l'utilisation de l'IA dans la surveillance des écosystèmes, pour l'anticipation et la minimisation pour les populations des risques

environnementaux liés aux événements extrêmes plus fréquents, mais également pour la prévention de la déforestation, la protection de la biodiversité, ou le suivi des impacts des changements climatiques.

## LES IMPACTS NÉGATIFS ATTENDUS

Néanmoins, les espoirs et solutions qu'apporte l'IA pour la transition écologique sont paradoxalement contrebalancés par des impacts négatifs sur l'environnement lui-même. Au niveau mondial, le numérique représente environ 4 % des émissions de gaz à effet de serre (GES), équivalent au secteur de l'aérien, et pour la France, l'Ademe évalue cette part à 4,4 % des GES émis en 2022 (Brilland *et al.*, 2025), soit près de 30 millions de tonnes CO<sub>2</sub>e (un peu moins que le secteur des poids lourds). Du fait entre autres de l'avènement de l'IA, l'Ademe prévoit une multiplication par trois de cette quantité de GES en 2050 si aucune mesure n'est prise pour limiter les impacts du numérique.

Les impacts de l'IA, qui contribuent aux impacts environnementaux plus généraux du numérique, se caractérisent par une augmentation forte et rapide, reflétant ainsi les besoins croissants en calcul et ressources informatiques pour collecter les données d'apprentissage, entraîner les modèles d'IA et exécuter les algorithmes de prédiction. Des travaux voient ainsi le jour pour estimer l'empreinte carbone de modèles d'apprentissage profond ; par exemple, plusieurs centaines de kg à plus de 100 tonnes CO<sub>2</sub>e pour apprendre les millions de paramètres de certains modèles de langage (Strubell *et al.*, 2020 ; Bouza *et al.*, 2023). Pour 2024, le collectif GreenIT a estimé les émissions de gaz à effet de serre (GES) dues à l'IA, à l'échelle de l'Europe, à plus de 3 millions de tonnes CO<sub>2</sub>e (ce qui équivaut à près de 20 milliards de km effectués en voiture thermique).

La complexité croissante des modèles d'apprentissage, la masse de données manipulées et la quantité impressionnante de calculs nécessaires pour déployer ces IA nécessitent de nombreux équipements électroniques toujours plus puissants et plus fréquemment renouvelés. La fabrication de ces équipements a un impact considérable sur l'épuisement des ressources naturelles, en particulier les métaux rares qui sont aussi largement convoités pour le photovoltaïsme, et sur l'émission de GES. Le nombre de centres de données ne cesse également d'augmenter, et les centres existants s'agrandissent, d'où la recommandation dans l'avis d'experts de l'Ademe (Ademe, 2024) d'inscrire tout nouveau centre de données dans l'objectif « Zéro Artificialisation Nette » de la loi Climat et Résilience de 2021. Le renouvellement trop rapide de ces équipements génère également de nombreux déchets électroniques. Le poids de ces derniers est en effet en forte croissance au niveau mondial avec un taux de recyclage encore très faible et des conséquences néfastes sur la santé et la biodiversité à proximité des sites de traitement ou de stockage dues à la pollution de ces sites (Ficher *et al.*, 2023 ; Baldé *et al.*, 2024).

Même si l'efficacité énergétique des micro-processeurs s'améliore d'années en années, la consommation électrique de tous ces équipements numériques augmente drastiquement (Trystram *et al.*, 2021 ; Crawford, 2024), ce qui ajoute une tension pour le partage de la ressource électrique avec les autres secteurs d'activité. Les besoins en eau sont également très importants et interviennent à de multiples niveaux (Bouveret *et al.*, 2024) : refroidissement des centres de données, fabrication des serveurs et équipements, production électrique. De premières études estiment par exemple que la demande mondiale en eau de l'IA en 2027 sera équivalente à la moitié de la demande en eau du Royaume-Uni (Li *et al.*, 2023). Ces forts besoins en eau peuvent ainsi s'avérer critiques en période de forte chaleur et de stress hydrique (Ademe, 2024).

On constate ces derniers mois, une multiplication de rapports, avis et recommandations sur la mise en garde des impacts de l'IA sur l'environnement et la nécessité d'investir rapidement dans le développement d'IA dites « frugales » (Afnor, 2024 ; Tatot et Vermot Desroches, 2024). L'une des principales raisons de l'urgence est la croissance vertigineuse



de l'usage de l'IA par le grand public et dans tous les pans de la société (objets connectés, loisirs, mobilité, etc.). L'époque où l'étape d'apprentissage des modèles, entre autres des grands modèles de langage, dominait largement l'empreinte environnementale de l'IA sur l'étape d'utilisation est révolue. Ainsi, Google a annoncé dans son rapport environnemental 2024 (Google, 2024) que ses émissions de GES ont augmenté de 48 % en 2023 par rapport à 2019 du fait de l'explosion des besoins en IA ce qui compromet leur objectif de réduction de leurs GES ; un problème que rencontrent également ses concurrents Microsoft et Amazon (*La Tribune*, 2024).

## DISCUSSION

Comme le mentionne le Conseil économique, social et environnemental dans son avis sur l'IA et l'environnement (Tatot et Vermot Desroches, 2024), « l'IA représente une vraie rupture d'impact et d'usage par rapport au domaine plus large du numérique ». Les recommandations du CESE, partagées par de nombreux experts (Green It, 2024), mettent en avant plusieurs axes d'amélioration :

- Une meilleure évaluation de l'empreinte environnementale de l'IA, de la conception à l'utilisation, et pour cela un meilleur accès aux données environnementales en lien avec l'IA, en particulier de la part des entreprises privées. Cela permettra en particulier aux utilisateurs d'orienter leurs choix vers les IA les plus sobres.
- Une orientation de la part de tous les acteurs technologiques de l'IA vers des systèmes plus frugaux, depuis la conception des centres de données jusqu'aux terminaux utilisateurs et aux grands modèles de langage utilisés. En particulier, une intégration systématique de la démarche d'écoconception des nouveaux équipements, en s'appuyant sur la norme (Afnor, 2024).
- Une priorisation de la recherche publique en IA vers le *design* de systèmes frugaux ou leur application à des questions environnementales, comme l'illustre par exemple le projet SHARP du PEPR IA (<https://project.inria.fr/sharp>).

Malgré ces recommandations importantes, nous partageons l'avis de nombreux experts qui expriment des doutes sur le fait que le développement d'une telle IA frugale serait suffisant pour en maîtriser l'ensemble des impacts environnementaux négatifs (Trystram et Ménessier, 2024). Tant qu'un usage large d'objets numériques à base d'IA (ordinateurs, *smartphones*, objets connectés, véhicules autonomes, etc.) sera encouragé dans la société, on peut prévoir que toute avancée à court terme vers une IA plus frugale se traduira à moyen terme, selon un classique « effet rebond », par un accroissement de l'usage de ces objets numériques à base d'IA et de leur impact négatif sur l'environnement, renforçant un mode de vie non soutenable à plus long terme.

Ainsi, au-delà d'une IA plus frugale, nous pensons nécessaire d'aller vers des recommandations fortes pour une plus grande sobriété dans l'usage de l'IA et de tout le numérique qui l'accompagne. Le rapport du CESE évoque cela, en recommandant une responsabilisation de l'ensemble des usagers, privés, publics, citoyens *via* des campagnes d'information sur l'empreinte environnementale de l'IA, pour que tous soient à même d'éviter les surutilisations inutiles.

## RÉFÉRENCES

ADEME (2024), « Avis d'experts. Les *data centers* ou centres de données. Au centre de la transition numérique », <https://librairie.ademe.fr/ged/9341/Avis-Data-Centers-Centre-de-Donnees-2024.pdf>

AFNOR (2024), « Référentiel général pour l'IA frugale - Mesurer et réduire l'impact environnemental de l'IA. AFNOR SPEC 2314 ».

BALDE C. P. *et al.* (2024), “The global e-waste monitor”, International Telecommunication Union and United Nations Institute for Training and Research, <https://ewastemonitor.info/the-global-e-waste-monitor-2024/>

BELLON MAUREL V., BROSSARD L., GARCIA F., MITTON N. & TERMIER A. (2022), « Agriculture et numérique. Tirer le meilleur du numérique pour contribuer à la transition vers des agricultures et des systèmes alimentaires durables », INRIA, Livre blanc, 195 pages.

BOUVERET S., BUGEAU A., ORGERIE A.-C. & QUINTON S. (2024), « De l'eau dans les nuages », *Annales des Mines - Enjeux numériques*, septembre 2024.

BOUZA L., BUGEAU A. & LANNELONGUE L. (2023), “How to estimate carbon footprint when training deep learning models? A guide and review”, *Environmental Research Communications*, 5, 115014, <https://doi.org/10.1088/2515-7620/acf81b>

BRILLAND T., FANGEAT E., MEYER J., WELLHOFF M., ADEME (2025), « Évaluation de l'impact environnemental du numérique en France. Mise à jour de l'étude Ademe-Arcep », 35 pages, <https://bibliothèque.ademe.fr/changement-climatique/7880-evaluation-de-l-impact-environnemental-du-numerique-en-france.html>

CERVONI L., CLERGERIE E. & ROUSSEAUX F. (2024), « Réconcilier les intelligences (artificielles) », ActuaIA, Dossier : Réconcilier les IA, août 2024.

CRAWFORD K. (2024), “Generative AI's environmental costs are soaring - and mostly secret”, *Nature*, 626, p. 693.

FICHER M., BAUER T. & LIGOZAT A.-L. (2023), « Les DEEE numériques en France », Rapport technique, Saclay, CNAM, 2023, <https://hal.science/hal-04098638>

GOOGLE(2024), Googleenvironmentalreport, <https://sustainability.google/reports/google-2024-environmental-report/>

GREEN IT (2024), « Quels sont les impacts environnementaux de l'IA ? », <https://www.greenit.fr/2024/08/15/quels-sont-les-impacts-environnementaux-de-lia/>

KONIECZNY S. & PRADE H. (2020), *L'intelligence artificielle. De quoi s'agit-il vraiment ?*, Cépaduès, 104 pages.

LA TRIBUNE (2024), « Google voit ses émissions carbone exploser à cause de l'IA », *La Tribune*, 3 juillet 2024, <https://www.latribune.fr/climat/energie-environnement/google-voit-ses-emissions-carbone-exploser-a-cause-de-l-ia-1001368.html>

LI P., Yang J., ISLAM M. A. & REN S. (2023), “Making AI Less ‘Thirsty’: Uncovering and Addressing the Secret Water Footprint of AI Models”, Published in arXiv.org, *Environmental Science, Computer Science*, 6 April 2023, <https://arxiv.org/abs/2304.03271>

OECD (2022), “Measuring the environmental impacts of artificial intelligence compute and applications: The AI footprint”, OECD Digital Economy Papers, 341, OECD Publishing, Paris.

STRUBELL E., GANESH A. & McCALLUM A. (2020), “Energy and Policy Considerations for Modern Deep Learning”, *Proceedings of the AAAI Conference on Artificial Intelligence*, <https://doi.org/10.1609/aaai.v34i09.7123>

SULEYMAN M. & BHASKAR M. (2023), *La Déferlante : Technologie, pouvoir et le dilemme majeur du XXI<sup>e</sup> siècle*, Paris, Fayard, 384 pages.

TATOT F. & VERMOT DESROCHES G. (2024), « Impacts de l'intelligence artificielle : risques et opportunités pour l'environnement », Journal officiel de la République française, avis du Conseil économique, social et environnemental (CESE), septembre, 66 pages.

TRYSTRAM D., COUILLET R. & MENISSIER T. (2021), « Apprentissage profond et consommation énergétique : la partie immergée de l'IA-ceberg », *The Conversation*, 8 décembre 2021, <https://theconversation.com/apprentissage-profond-et-consommation-energetique-la-partie-immeree-de-lia-ceberg-172341>

TRYSTRAM D. & MENISSIER T. (2024), « L'IA peut-elle vraiment être frugale ? », *The Conversation*, 13 mai 2024, <https://theconversation.com/lia-peut-elle-vraiment-etre-frugale-226274>

# L'IA durable n'existe pas

Par **Frédéric MARCHAND**

Cofondateur et CEO de Digital4Better

Une IA durable peut-elle exister ? Cet article aborde les implications environnementales de l'intelligence artificielle (IA), souvent ignorées face à ses bénéfices. L'IA, tout en promettant des avancées, génère des externalités négatives, notamment environnementales. Avec la démocratisation de l'IA générative, sa consommation énergétique augmente significativement, illustrée par l'augmentation exponentielle de la demande en calcul informatique. Une IA utile mesurée et non une IA futile démesurée, c'est le chemin préconisé par l'auteur.

Il est fort probable que l'IA va révolutionner notre monde mais l'IA pourrait-elle à l'inverse amener les pays à la révolution ?

L'IA est un levier de progrès dans de nombreux domaines, comme la santé par exemple. Ce n'est pas pour autant qu'il faut ignorer les externalités négatives de l'IA. Le faire serait comme conduire les yeux fermés sa voiture pied au plancher.

Toute action a des externalités négatives et positives. Les externalités positives de l'IA sont nombreuses et il est indéniable que l'IA puisse avoir un impact positif pour peu que nous en fassions bon usage. Mais afin de savoir si la balance est positive, il faut déjà appréhender les externalités négatives de l'IA.

Les biais sociaux de l'IA ne manquent pas. Dans cet article, nous ne traiterons que de la partie environnementale de l'IA. Et c'est déjà beaucoup pour un seul article.

Souvent présentée comme une solution à la transition écologique, l'IA est aussi l'une des sources du problème. Pourquoi une telle explosion à venir ? En effet, l'IA n'est pas nouvelle. La démocratisation de l'IA générative (comme ChatGPT) est la principale raison des trajectoires que nous observons. Une citation pour poser un peu le contexte : « Le besoin de calcul informatique pour l'IA a été multiplié par un million en 6 ans et il décuple chaque année » déclarait en mai 2024 Sundar Pichai, patron de Google. Le ton est donné.

## L'EMPREINTE ENVIRONNEMENTALE DE L'IA, DE QUOI PARLE-T-ON ?

Regardons ce qu'implique l'IA en termes d'empreinte environnementale, en partant du haut et en allant jusqu'à la fabrication des équipements dont elle a besoin.

### Les phases de l'IA

L'empreinte de l'IA de manière schématique se structure en deux phases : phase d'entraînement du modèle (apprentissage) et la phase d'usage (inférence). Auparavant, la phase d'apprentissage, nécessitant beaucoup de temps et de données, représentait la majeure partie de l'empreinte environnementale, mais dorénavant, avec le déploiement massif de l'IA générative, la phase d'inférence est beaucoup plus significative et peut représenter jusqu'aux deux tiers de l'empreinte.



Figure 1 : L’empreinte environnementale de l’IA  
(Source : Image générée par une IA à partir du chapô de l’article).

## Des terminaux, du réseau et des serveurs

L’IA c’est une part importante du côté des *data centers*. C’est surtout du calcul. Alors que pour les autres systèmes numériques, la part la plus importante est dans les terminaux (PC, *smartphone*...), qui représentent 75 % en moyenne de l’empreinte du numérique par rapport aux couches réseaux et serveurs.

### Les *data centers*

Les centres de données utilisent de l’électricité pour alimenter les serveurs mais également pour les refroidir. Et cette consommation n’est pas neutre. Pour donner un ordre de grandeur, cette demande électrique des *data centers* dans le monde, c’est l’équivalent de la consommation électrique de la France pour toute l’année 2023.

## La phase de fabrication

À cela s'ajoute la phase de fabrication des équipements. Le numérique est tout sauf virtuel et l'IA a une matérialité cachée très importante. Comme toute activité numérique l'IA nécessite de mettre en action de nombreux équipements numériques : des serveurs, des équipements réseaux, des terminaux pour y accéder... Tout ce processus de fabrication de l'ensemble de nos équipements numériques a un poids environnemental très lourd. Ce poids est estimé à ce jour à 4 % des émissions mondiales de GES et devrait tripler dans les années à venir selon différents scénarios, dont l'étude de l'Ademe sur le sujet.

Pour fabriquer tous ces équipements, il est nécessaire d'extraire de nombreuses terres rares, essentielles à l'élaboration des différents composants électroniques. L'extraction des métaux est à la base de la fabrication des équipements nécessaires à l'IA. L'exploitation des terres rares a un impact néfaste sur l'écosystème global associé à l'extraction et au traitement pour la purification de ces métaux bruts obtenus à partir de la croûte terrestre. L'extraction de terres rares implique souvent des pratiques dommageables pour l'environnement telles qu'une contamination de l'eau, une pollution de l'air... associées souvent à un impact social : conditions de travail des personnes dans les mines, impacts sur les populations environnantes qui se retrouvent avec des cours d'eau pollués...

Ce poids environnemental de l'IA impacte tous les indicateurs environnementaux : écotoxicité aquatique, radiations ionisantes, acidification des sols...

## L'IMPACT CONSTATÉ DE L'IA ET LES SCÉNARIOS

L'IA serait-elle devenue le nouveau prétexte pour les grands acteurs afin de ne pas respecter leurs engagements de décarbonation ? Microsoft devait atteindre la neutralité carbone pour 2030, pour autant dans son dernier rapport climat sorti en 2024, il a affiché une détérioration de 29 % en 3 ans de ses émissions ! Une trajectoire carbone à l'exact inverse de son engagement. L'excuse donnée : l'IA ! Même son de cloche côté Google, qui lui, a vu ses émissions bondir de 48 % en 4 ans atteignant 14,3 millions de tonnes de CO<sub>2</sub> en 2023. L'IA est à la fois une révolution pour les entreprises mais également ce qui va ruiner leurs engagements de décarbonation si nous faisons de l'IA à tout va, sans réfléchir.

L'IA devient donc le mot d'excuse parfait des mauvais élèves. Les mêmes qui indiquaient que l'IA allait résoudre tous les problèmes. Étonnant.

L'impact réel de l'IA reste encore difficile à estimer. En effet, les grands acteurs se gardent bien de partager les données en toute transparence. Nous manquons donc de données mais même si cela reste une estimation avec des inconnues, il est important d'essayer de définir l'empreinte environnementale de l'IA.

## Électricité

Les *data centers* sur la partie IA et Crypto représentent presque 2 % de la consommation électrique mondiale en 2022, soit une consommation électrique de 460 TWh, ce qui correspond à peu près à la consommation électrique de la France pour toute l'année 2023. La consommation électrique du numérique était globalement à 10 %, celle-ci va continuer à augmenter et surtout la part due à l'IA va fortement croître.

Selon les données de l'Electric Power Research Institute (EPRI), les *data centers* vont consommer entre 4,6 et 9,1 % de l'électricité américaine d'ici 2030. Le Boston Consulting Group (BCG) prévoit de son côté que la proportion pourrait atteindre 7,5 % d'ici 2030. L'ordre de grandeur est donc là et il n'est pas neutre.

Pour bien comprendre les choses, prenons un élément de comparaison : l'interface ChatGPT 3.5 c'est 2,9 wattheures (Wh) *versus* Google 0,3 wattheures (Wh), soit 10 fois plus.

Les versions d'après de ChatGPT (4o, puis o1) vont encore augmenter d'un ou deux ordres de grandeur cette consommation.

## Terres rares

Entre 1980 et 2008, la demande globale de métaux a augmenté de 87 % (France Stratégie, 2020). D'ici à 2050, la quantité de métaux requis pourrait représenter 3 à 10 fois les volumes de production actuels. Nvidia, *leader* de la fabrication de puces pour l'IA, a vu sa demande en terres rares exploser pour les produire.

## Eau

Une étude de la Cornell University établissait en octobre 2023 que l'IA pourrait consommer en 2027 entre 4,2 et 6,6 milliards de mètres cube d'eau.

Une conversation avec ChatGPT consomme environ 50 cl d'eau. Si vous multipliez cela par des milliards de conversations, le résultat devient effrayant.

## GES

Certaines études estiment que l'entraînement du modèle d'IA BLOOM émet 10 fois plus de GES qu'un Français en une année.

Les estimations portent à 10 113 tonnes CO<sub>2eq</sub> pour GPT-3 pour le seul mois de janvier 2023 (estimation opérée par l'association Data For Good dans son livre blanc de juillet 2023).

## VERS UNE GUERRE GÉOPOLITIQUE ?

Plusieurs visions s'opposent sur l'IA, d'un côté une IA sans limite portée par Elon Musk, de l'autre des pays comme le Canada qui portent la vision d'une IA de confiance, durable. Les scénarios qui vont en découler sont un excellent terrain de films de science-fiction pas si lointaine. L'empreinte de l'IA pourrait en effet générer une guerre géopolitique inattendue.

En septembre 2017, Vladimir Poutine déclarait sur l'IA « Celui qui deviendra *leader* en ce domaine sera le maître du monde ».

Chaque pays est convaincu que l'IA est une clef maîtresse pour influencer le monde. Dans un contexte où le terme de souveraineté numérique prend de plus en plus d'ampleur, où beaucoup de pays se renferment sur eux-mêmes avec la montée en puissance de certaines politiques peu favorables à partager leurs richesses et protéger la planète, la guerre de l'intelligence artificielle pourrait bien avoir lieu.

Nous connaissons la guerre des métaux rares, décrite dans l'excellent livre de Guillaume Pitron. Voici maintenant la guerre de l'intelligence artificielle, dans une ambiance électrique de terres rares et de stress hydrique. Où pourrait nous mener l'IA ?

## Guerre électrique

En octobre 2024, Google annonce vouloir acheter 6 à 7 mini réacteurs nucléaires (SMR) pour une production cible de 500 mégawatts.

Cette explosion de besoins en *data centers*, nécessaire notamment pour l'IA va accroître les besoins en électricité. Tous les pays auront-ils la capacité de fournir une telle demande

électrique ? Certains acteurs de la tech anticipent le problème en se lançant dans les mini centrales nucléaires pour leurs *data centers*.

Une étude publiée par l'EPRI en mai 2024 : "Powering Intelligence: Analyzing Artificial Intelligence and Data Center Energy Consumption" donne beaucoup d'informations et montre l'enjeu sur les différents scénarios. L'étude met en avant notamment que dans le scénario le plus pessimiste, pour certains États américains, la part de la consommation électrique des *data centers* pourrait atteindre 50 % de la consommation électrique totale de l'État. Cela relève bien tout l'enjeu du sujet. Quel État peut avoir une telle dépendance énergétique ?

Nous pourrions assister à une pénurie d'électricité, et notamment des refus de certains opérateurs de raccorder des centres de données. Ce cas s'est d'ailleurs déjà produit avec le refus d'Hydro Québec en Gaspésie.

En Île-de-France, l'Institut Paris Région estime qu'il faudrait jusqu'à sept réacteurs nucléaires d'ici 2030 pour répondre à cette demande de l'IA.

L'installation de *data centers* nécessaires au fonctionnement de l'IA va potentiellement devenir un enjeu géopolitique. Une région sera-t-elle d'accord pour accueillir un *datacenter* qui ne servira pas que ses intérêts quitte à frôler une crise énergétique ?

## Guerre de l'eau

En mai 2023, en pleine période de sécheresse, l'Espagne s'inquiète du projet de *datacenter* de Meta. Meta (Facebook) doit lancer la construction d'un *datacenter* d'une puissance électrique de 248 mégawatts sur 180 hectares à Talavera de la Reina, en pleine région aride. El País a estimé la consommation totale d'eau, c'est-à-dire d'eau potable et non potable, à 120 litres par seconde dans le centre de données et 33 litres par seconde dans le reste des installations ; ce qui ferait une consommation d'eau de 4,8 milliards de litres par an.

Alors même que l'eau est devenue une ressource rare dans certaines régions, l'IA va accentuer ce stress hydrique. Arrivera-t-on demain à un dilemme : veux-tu ton IA ou de l'eau pour te nourrir et ton agriculture ? La guerre de l'eau pourrait bien avoir lieu.

## QUEL CADRE POUR L'IA

L'un des tous premiers cadres pour l'IA date de 2018. Dans la déclaration de l'IA responsable de Montréal, il est évoqué dix principes :

- bien-être ;
- respect de l'autonomie ;
- protection de l'intimité et de la vie privée ;
- solidarité ;
- participation démocratique ;
- équité ;
- inclusion de la diversité ;
- prudence ;
- responsable ;
- développement soutenable.

Cela reste cependant des principes vertueux qui n'ont pas encore donné lieu à un cadre législatif.

À l'inverse, l'Union européenne occulte à ce jour la partie environnementale, avec une approche portée par les risques, mais qui a l'avantage de se traduire par des lois. L'Acte



sur l'IA de l'Union européenne pose des principes de risques afin d'encadrer l'IA. L'AI Act classe les applications de l'IA en quatre niveaux de risque.

Dans cette lutte de pouvoir sur l'IA, certains pays sont prêts à une IA à tout prix. Un cadre n'est cependant pas un frein, il permet au contraire de se développer sereinement. Pouvez-vous imaginer élever vos enfants sans le moindre cadre, sans la moindre règle. Vous les laissez choisir de dormir quand ils le souhaitent, manger tout ce qu'ils veulent, ne pas aller à l'école... À quel résultat arriverez-vous 20 ans plus tard ? Imaginez le même scénario avec l'IA. En IA d'ailleurs, il y a la phase clef dont le nom est la phase d'apprentissage...

## RETOUR AU BON USAGE

En numérique responsable, nous évoquons souvent la règle des 3U : Utile, Utilisable et Utilisé. Pour chaque fonctionnalité, il faut se poser à chaque fois la question des 3U.

Est-il pertinent de déployer un outil d'IA généraliste auprès de tous ses salariés qui vont l'utiliser pour tout et rien, comme en arrière-plan de leurs visioconférences pour faire les comptes rendus ou tout simplement pour lui demander la météo ? C'est pourtant la réalité que vivent certains acteurs, qui constatent les dégâts, tant sur leur empreinte environnementale que sur leur facture...

Mais que pouvons-nous faire afin de faire une IA plus durable ? Voici quatre leviers nécessaires pour y arriver.

## Le modèle qui sait tout faire c'est bien, mais non...

Plus le modèle d'IA est spécialisé, plus il est petit et moins il consomme. Les modèles généralistes comme ChatGPT 4 ne doivent donc être que très peu utilisés ; pour les tâches spécifiques il est préférable d'utiliser des modèles spécialisés beaucoup moins impactant.

## *Open source*

Il permet beaucoup de réutilisation et donc de limiter l'impact des phases d'entraînement. Il est essentiel de voir apparaître le maximum de modèles spécialisés en *open source*.

## La frugalité est source d'efficience

Comment faire plus avec moins ? Cela passe déjà par questionner nos usages. À chaque usage, la bonne approche. Il convient ensuite d'écoconcevoir nos systèmes d'IA. La recherche d'efficience doit se faire le plus en amont possible.

## Sensibiliser et former les usagers

Toutes les parties prenantes sont responsables mais si nous ne leur expliquons pas les externalités négatives et les biais de l'IA ils ne peuvent pas agir en conséquence. Avant tout déploiement massif d'IA, il est important de former les personnes, de donner un cadre, de partager des bonnes pratiques.

## Pas d'amélioration sans mesure

Nous ne pouvons pas améliorer ce que nous ne mesurons pas. C'est la base. La mesure doit s'opérer dès le début. Cette mesure, même si elle reste une évaluation, pourra être communiquée en toute transparence, permettant à chacun d'agir mieux en toute connaissance cause.

## CONCLUSION

Une IA utile mesurée et non une IA futile démesurée, tel est le chemin que nous devons prendre.

L’empreinte environnementale de l’IA est réelle et il serait inconscient de ne pas en prendre conscience. Néanmoins, les effets positifs de l’IA sont également réels, l’IA est une des solutions aux dérèglements climatiques pour rester sur les sujets environnementaux.

Cette réflexion sur la balance entre les externalités positives et négatives sera essentielle dans nos approches de l’IA.

Les ROI carbone, productivité et financier de l’IA ne sont pas encore prouvés (notamment le carbone), la mesure de ce ROI ainsi que la mesure des impacts négatifs définira la performance globale réelle de l’IA et permettra de prendre les bonnes décisions.

## POUR ALLER PLUS LOIN AVEC SOBRIÉTÉ SUR L’EMPREINTE ENVIRONNEMENTALE DE L’IA

### 3 Livres blancs

Association Hub France IA, ChatGPT : usages, impacts et recommandations, février 2024 : [https://www.hub-franceia.fr/wp-content/uploads/2024/03/Hub-France-IA-Propositions-IA-Generative\\_0803.pdf](https://www.hub-franceia.fr/wp-content/uploads/2024/03/Hub-France-IA-Propositions-IA-Generative_0803.pdf)

Association Data For Good, Livre blanc IA Générative, juillet 2023 : <https://dataforgood.fr/iagenerative/>

AFNOR - Référentiel général sur l’IA frugale : <https://www.boutique.afnor.org/fr-fr/norme/afnor-spec-2314/referentiel-general-pour-lia-frugale-mesurer-et-reduire-limpact-environneme/fa208976/421140#>

### 3 Études scientifiques

Estimating the environmental impact of Generative-AI services using an LCA-based methodology, Adrien Berthelot, Eddy Caron, Mathilde Jay, Laurent Lefèvre. Estimating the environmental impact of Generative-AI services using an LCA-based methodology. CIRP LCE 2024 - 31<sup>st</sup> Conference on Life Cycle Engineering, Jun 2024, Turin, Italy. pp. 1-10. (hal-04346102v2) : <https://inria.hal.science/hal-04346102>

Power Hungry Processing: Watts Driving the Cost of AI Deployment?, @article{Luccioni2023PowerHP, Sasha Luccioni and Yacine Jernite and Emma Strubell, 2023: <https://arxiv.org/abs/2311.16863>

Impacts environnementaux de l’IA : quels réels bénéfiques ?, Julien Lefevre, Aurélie Bugeau, Jacques Combaz, Laurent Lefèvre, Anne-Laure Ligozat *et al.*, Collection numérique de l’AMUE, Agence de mutualisation des universités et établissements d’enseignement supérieur, 2023, 30, fhal-04486682 : <https://hal.science/hal-04486682/document>

### 1 Podcast

Podcast «IA Café», Épisode 62 - Mesure de l’impact environnemental de ChatGPT - Entrevue avec Anne-Laure Ligozat : <https://www.buzzsprout.com/1431712/episodes/12921515-episode-62-mesure-de-l-impact-environnemental-de-chatgpt-entrevue-avec-anne-laure-ligozat>

# Pour un développement de l'IA au service du bien commun

Par **Laure de LA RAUDIERE**

Présidente de l'Autorité de régulation des communications électroniques, des postes et de la distribution de la presse (Arcep)

L'émergence et la diffusion rapide de l'intelligence artificielle générative au grand public depuis la fin de l'année 2022 sont autant vectrices d'opportunités que d'enjeux économiques et sociétaux. En tant qu'architecte et gardienne des réseaux d'échanges comme biens communs, l'Arcep alerte sur plusieurs impacts potentiels des IA génératives. En devenant les nouvelles portes d'entrée de notre accès à internet, les grands acteurs de l'IA générative pourraient remettre en cause le principe d'ouverture d'Internet. En outre, le développement de l'IA générative interroge quant à sa soutenabilité environnementale. Sur ces enjeux, plusieurs leviers doivent être mobilisés pour un développement de l'IA au service du bien commun.

*Un grand merci aux « Arcépiens »,  
Chiara Caccinelli, Sandrine Elmi Hersi, Tom Nico,  
Marion Panfili, Oriane Piquer-Louis et Nathan Salmon,  
qui ont contribué activement à la rédaction de cet article.*

L'intelligence artificielle (IA) générative connaît une diffusion rapide dans la société depuis fin 2022 et l'émergence de ChatGPT. Elle représente une avancée technologique importante par sa capacité à produire du contenu « nouveau » et de diverses natures (texte, images, vidéos, son...) à partir de simples instructions et de données d'entraînement. Ce type d'IA pourrait offrir des opportunités considérables, tant d'un point de vue économique que sociétal. Néanmoins, l'IA générative soulève également un certain nombre de défis à relever pour s'assurer de son développement au service du bien commun.

Si l'adoption de l'IA générative à grande échelle a été marquée par l'émergence de plusieurs acteurs, elle soulève néanmoins des préoccupations en matière de concurrence. L'IA repose sur des ressources clés – telles que les services *cloud*, les données, la puissance de calcul et les compétences techniques – qui peuvent être fournies ou contrôlées par un nombre limité de grandes entreprises. La course actuelle à l'innovation et à l'adoption des services constitue ainsi un *momentum* qui dessinera les dynamiques concurrentielles futures tant de l'IA générative que des services numériques en général. Le développement d'un marché concurrentiel et ouvert pour l'IA générative et les services qui l'intègrent apparaît donc nécessaire pour que cette technologie bénéficie au plus grand nombre.

En outre, face à l'essor d'un usage grand public de l'IA générative, il est fondamental d'évaluer son impact potentiel sur les droits et la liberté des utilisateurs en ligne. L'IA générative pourrait devenir une nouvelle porte d'accès aux contenus et services internet, et remplacer, en tout ou partie, les moteurs de recherche. Si cette tendance se confirme, les utilisateurs auront-ils toujours la possibilité d'accéder à tous les contenus en ligne ? Et les contenus qu'ils partageront en ligne seront-ils toujours accessibles par tous ? En d'autres termes, le développement des IA génératives pourrait avoir des implications sur un droit essentiel, celui de l'accès à un internet ouvert. L'Arcep souhaite porter dans le

débat public cet enjeu, afin de définir les conditions permettant de concilier IA générative et internet ouvert.

Enfin, au vu des importantes ressources que requiert l'intelligence artificielle générative et de la croissance attendue de son usage, l'Arcep souhaite mettre à disposition son expertise pour alimenter le débat public sur les impacts environnementaux potentiels de cette technologie.

## **L'INTELLIGENCE ARTIFICIELLE GÉNÉRATIVE COMME PORTE D'ENTRÉE À INTERNET, UN ENJEU MAJEUR**

À l'instar des moteurs de recherches, des systèmes d'exploitation ou des réseaux sociaux, l'IA générative pourrait devenir, pour de nombreux internautes, une nouvelle porte d'entrée pour accéder aux contenus et aux services sur internet. Une récente étude de Gartner prévoit ainsi une baisse de 25 % du volume de requêtes des moteurs de recherche d'ici 2026 en raison du développement des *chatbots* utilisant de l'IA générative (Gartner, 2024). Par ailleurs, les systèmes d'IA générative sont de plus en plus intégrés dans les terminaux utilisateurs et les services numériques traditionnels : moteurs de recherche, réseaux sociaux et systèmes d'exploitation en particulier.

La place de l'IA générative comme interface « intelligente » entre l'utilisateur et le contenu sur internet pourrait avoir des conséquences importantes sur le partage et la diffusion d'informations et de données en ligne et menacer l'ouverture d'internet, c'est-à-dire la capacité des utilisateurs à accéder et à partager le contenu de leur choix, principe garantissant la liberté de communication et d'innovation. Ce principe d'ouverture s'inscrit au cœur du modèle numérique de l'Union européenne, garanti par le règlement 2015/2020 et reconnu par la déclaration européenne sur les droits et principes numériques. Ces enjeux pourraient être d'autant plus prégnants qu'un nombre pour l'instant limité d'acteurs domine le marché de fourniture de services d'IA générative<sup>1</sup>.

### **L'IA générative, une nouvelle passerelle non neutre entre l'utilisateur et son contenu**

Les utilisateurs, en s'appuyant sur des IA génératives pour accéder aux contenus, délèguent leur choix à des systèmes souvent opaques quant aux sources et aux critères de sélection (Shah et Bender, 2024). C'est d'autant plus préoccupant que l'IA générative repose sur des systèmes algorithmiques très puissants mais qui présentent encore des difficultés à expliquer ce qui a guidé leurs recommandations, y compris les sources utilisées...

Les technologies d'IA générative ne se contentent plus de partager ou d'organiser du contenu produit par des tiers ; elles génèrent des contenus *ad hoc*, amplifiant le poids des choix et des paramétrages opérés par le fournisseur d'IA. L'IA générative pourrait aussi accentuer les risques d'enfermement et de « bulles algorithmiques » du fait d'une perte de contrôle de l'utilisateur sur ses choix en ligne, devant s'en remettre à un nouvel intermédiaire. Ce phénomène de « délégation » est commun à d'autres services numériques, mais prend une nouvelle dimension lorsqu'on se réfère à l'IA générative. Ainsi, les

---

<sup>1</sup> L'Arcep a par exemple partagé ces questionnements avec la Commission européenne dans sa réponse à la consultation publique sur les enjeux concurrentiels de l'IA générative ouverte de janvier à mars 2024 ainsi que dans son rapport sur l'état de l'internet 2024.

choix effectués par les fournisseurs d'IA générative en termes de corpus d'entraînement, paramètres, choix de modèles sont susceptibles de déterminer le contenu à disposition des utilisateurs. Il faut notamment prendre en compte les limites techniques des modèles de langage utilisés pour nos services d'IA générative : les risques de renforcement des biais sur internet ou encore « d'hallucinations » pouvant favoriser la diffusion de fausses informations en ligne sont déjà largement documentés<sup>2</sup>.

Ces questionnements sur les atouts et limites de l'IA générative comme nouvelle porte d'entrée pour accéder aux contenus en ligne sont d'autant plus cruciaux qu'aujourd'hui seule une poignée de grands acteurs du numérique dominant le marché de l'IA générative. Cette situation pourrait nous amener à qualifier ces fournisseurs de services comme possibles « générateurs de contenus structurants », puisque que leur positionnement leur permettrait de paramétrer les conditions selon lesquelles les utilisateurs accèdent aux contenus sur internet. Ce diagnostic appelle ainsi à une responsabilisation de ces acteurs afin de limiter leur potentielle capacité à restreindre l'information à laquelle les utilisateurs accèdent sur internet.

### **L'impact de l'IA générative sur la richesse, la diversité des contenus sur internet et le pluralisme**

L'IA générative comme nouvelle porte d'entrée vers l'internet implique également une profonde évolution des modes de partage et de diffusion du contenu. Il est certain que l'IA générative facilite la production de contenu : il n'a jamais été aussi facile de produire un texte, une image ou une ligne de code. Les potentialités d'aide à la création des outils d'IA générative peuvent ainsi stimuler l'innovation dans de nombreux secteurs.

Néanmoins, l'utilisation de l'IA générative pour produire du contenu soulève d'importants questionnements en matière de droit d'auteur. Par ailleurs, l'absence de « marquage »<sup>3</sup> des contenus générés par ou avec une IA ne permet pas aux utilisateurs de reconnaître les contenus originaux ou de leur donner une valeur différenciée. L'effort additionnel que peut présenter la création de contenu sans IA générative n'est pas reconnu. Si l'on ajoute à cela les risques d'appropriation de leur contenu par l'IA générative, les utilisateurs finals et professionnels pourraient être désincités à créer. Par exemple, les contributeurs aux communs numériques, comme Wikipédia, pourraient être découragés d'enrichir la base de connaissance si le contenu est progressivement remplacé pour certains usages par une IA générative, sans citation des multiples sources contributives. Dans le milieu artistique, les créateurs de contenus humains vont aussi être mis en concurrence avec des algorithmes qui, par ailleurs, peuvent les copier<sup>4</sup>.

Cela aura donc un double impact : sur la création comme évoqué, mais aussi sur la qualité des futurs modèles d'IA. En effet, du fait des limites de la récursivité algorithmique, lorsque des modèles d'IA sont réentraînés sur des contenus créés par d'autres systèmes

---

<sup>2</sup> Voir notamment le rapport « IA générative et mésinformation » de l'Académie des technologies, publié en 2024, <https://www.academie-technologies.fr/publications/ia-generative-mesinformation/>

<sup>3</sup> Le marquage désigne les techniques permettant d'informer l'utilisateur quant à l'origine du contenu qu'il consulte, c'est-à-dire s'il a été ou non produit avec l'appui ou par une IA générative. Il s'inspire des indications dont les contenus publicitaires font l'objet. Pour une reconnaissance *a posteriori* par des publics plus informés (journalistes, etc.), des pratiques telles que le « watermarking », (par exemple, SynthID de Google) pourraient être envisagées, sous réserve de leur efficacité. Le marquage présente un intérêt en matière de propriété intellectuelle, sans se limiter à cette application.

<sup>4</sup> Plusieurs artistes ont d'ores et déjà été victimes d'imitation de leurs créations (Shan *et al.*, 2023).

d'IA générative, la qualité et la diversité des données peuvent progressivement se détériorer (Shumailov *et al.*, 2024). La nature statistique des modèles d'IA générative pourrait aussi favoriser la répétition de contenus largement diffusés ou facilement accessibles, au détriment d'informations plus spécialisées. La richesse et la diversité des contenus accessibles en seraient ainsi affectées.

L'IA générative redéfinit en profondeur les relations technico-économiques sur le marché du contenu. Les fournisseurs de contenu traditionnels (les médias mais aussi les modèles de communs numériques type Wikipédia) auront-ils toujours la capacité à rendre disponible du contenu aux utilisateurs ? Quelle rémunération restera-t-il aux auteurs et aux fournisseurs si les contenus traditionnels deviennent en quelque sorte « invisibles » sur l'internet ? La remise en cause des modèles économiques des fournisseurs de contenus traditionnels par les IA génératives s'annonce majeure.

Pour monétiser l'utilisation de leurs contenus, ils devront sans doute conclure des accords avec les fournisseurs d'IA générative. Mais comme seul un nombre limité d'acteurs domine le marché, il n'est pas certain que la libre concurrence soit suffisante pour garantir les conditions d'une négociation équilibrée et la normalisation des relations entre fournisseurs d'IA générative et de contenus traditionnels. Les générateurs de contenu structurants pourraient choisir de conclure des accords exclusivement avec certains fournisseurs de contenu, créant ainsi un risque d'éviction des autres fournisseurs de contenu.

À titre d'exemple, des acteurs dominant le secteur de l'IA générative multiplient des contrats bilatéraux avec certains grands médias nationaux, plutôt que de conclure des accords à l'échelle de la filière. Les autres médias se retrouvent ainsi privés des utilisateurs accédant à l'information par les outils d'IA générative concernés. Cela pose aussi la question du pluralisme de l'information délivrée par les outils d'IA générative. Cette problématique rejoint celle déjà posée dans le secteur de la distribution de la presse au moment de l'émergence de kiosques numériques, mettant à disposition plusieurs titres de presse. Des normes ont été établies par la loi pour une République numérique et la loi Bichet dont la mise en œuvre permet de garantir la neutralité et pluralité de l'information que délivrent ces kiosques et l'absence de discrimination entre les titres de presse. Nous devons certainement établir, en France et dans l'Union européenne un cadre approprié pour garantir le pluralisme de l'information à l'ère des IA génératives. Il y a en effet peu de chances que le pluralisme s'établisse par défaut dans ce nouveau mode d'accès à l'information.

## Concilier internet ouvert et IA générative

Face à ces enjeux, l'Arcep pense essentiel de travailler collectivement – puissance publique, acteurs économiques, monde académique et société civile – à la mise en place des conditions permettant de concilier développement de l'IA et internet ouvert.

Parmi les pistes d'actions à envisager, la promotion de standards élevés d'efficacité, de transparence et d'auditabilité des modèles d'IA générative semble une étape essentielle pour garantir le pluralisme et préserver la richesse et la qualité des contenus sur internet. Cela permettra également aux utilisateurs d'être davantage informés sur le fonctionnement et les limites des IA génératives.

En complément, la mise en place de moyens concrets pour que les internautes puissent choisir les modalités de leurs usages numériques, y compris lorsqu'ils ont recours à une IA générative, pourrait être une réponse à ces enjeux, par exemple en instaurant un droit au paramétrage et au pluralisme algorithmique.

Enfin, une régulation technico-économique pourrait aussi se révéler bénéfique. En effet, comme pour d'autres services numériques contrôlés par une poignée de grands acteurs

– et régulés au titre du règlement européen “Digital Market Act” –, il est nécessaire de veiller à garantir les conditions d’un marché concurrentiel de l’IA générative en levant les possibles verrous à l’arrivée de nouveaux entrants sur le marché, déjà par la mobilisation des outils existants<sup>5</sup>. Mais cela ne sera sans doute pas suffisant. À la lumière de notre expérience en matière de neutralité du net, il sera nécessaire de veiller à ce que le pouvoir de marché des « générateurs de contenu structurants » n’entrave pas la richesse du contenu d’internet et laisse suffisamment de liberté aux utilisateurs et innovateurs de consulter et partager le contenu de leur choix.

## **POUR UN DÉVELOPPEMENT RESPONSABLE DE L’IA PAR LA MAÎTRISE DE SON EMPREINTE ENVIRONNEMENTALE**

L’Agence Internationale de l’Énergie (AIE) prévoit que la consommation mondiale d’électricité des centres de données pourrait doubler entre 2022 et 2026 (IEA, 2024), notamment à cause de la croissance de l’IA générative. Le développement fulgurant de l’IA soulève donc des interrogations sur sa soutenabilité. Promouvoir les usages de l’IA générative implique donc de se préoccuper de son empreinte environnementale.

### **Un besoin de mesurer les impacts environnementaux croissants de l’IA et d’approfondir les travaux en cours**

À l’instar de tout service numérique, l’IA repose sur des infrastructures et équipements physiques, nécessitant une quantité importante de ressources pour faire fonctionner les serveurs et les terminaux (Ligozat et al, 2022). Ces derniers ont des impacts tout au long de leur cycle de vie, de l’extraction des métaux nécessaires à la fabrication des équipements (notamment des puces), en passant par l’utilisation et enfin leur fin de vie.

Les serveurs, ainsi que leurs composants, ont besoin de ressources abiotiques métalliques et minérales pour être fabriqués. La fabrication des puces nécessite beaucoup d’énergie et d’eau. Les serveurs ont également besoin d’être alimentés en électricité, acheminée par le réseau électrique national. L’impact carbone des serveurs dépend donc du pays où ils sont hébergés. Par exemple, un modèle d’IA émettra 31 fois plus de gaz à effet de serre s’il s’exécute en Pologne plutôt qu’en Islande.

Par ailleurs, les centres de données comprennent d’autres équipements non-informatiques comme des onduleurs, des *switchs* ou des routeurs. Les serveurs doivent aussi être refroidis par des fluides frigorigènes, de l’eau ou de l’air. Il est ainsi nécessaire de considérer l’ensemble des infrastructures, *via* l’indicateur d’efficacité énergétique (PUE).

Enfin, l’IA peut avoir des impacts sur les réseaux, nécessaires aux transferts de données, et potentiellement sur le renouvellement des terminaux pour acquérir des terminaux adaptés à l’IA.

Les impacts environnementaux de l’IA sont aujourd’hui encore peu évalués (Ligozat *et al*, 2022) et n’ont pas été anticipés, comme le montrent les annonces des grands acteurs tels que Google ou Microsoft mi-2024 de renoncement à leur objectif de neutralité carbone à 2030, en particulier du fait du développement des IA génératives. Or, il s’agit d’un enjeu

---

<sup>5</sup> Dans le cadre du règlement européen sur les données (Data Act) et de la loi visant à sécuriser et réguler l’espace numérique (dite « loi SREN »), l’Arcep s’est notamment vue confier la mission d’élaborer un cadre tarifaire et technique pour lever un certain nombre d’obstacles au changement de fournisseur et au multi-*cloud*.

majeur compte tenu des prévisions de croissance des usages de l'IA et de ses conséquences sur les trajectoires énergétiques. Afin de mesurer l'impact environnemental de l'IA, il est nécessaire d'améliorer les connaissances en adoptant une approche multicritère, multi-composants et multi-étapes<sup>6</sup>.

À ce stade, des travaux récents sur l'IA générative permettent d'établir un premier socle de connaissances des impacts environnementaux malgré l'opacité des fournisseurs de services. Luccioni *et al.* (2023) montrent par exemple que la phase d'entraînement est la plus consommatrice en ressources (serveurs, GPUs) et en électricité, par rapport à la phase d'usage, qui est la phase où le modèle est mis à disposition des utilisateurs finals. La dynamique des impacts liés à la phase d'entraînement est préoccupante, car le nombre de modèles, leur taille ainsi que leur nombre de paramètres n'ont cessé de croître au cours des dernières années : 120 millions de paramètres pour le premier modèle de ChatGPT sorti en 2018 contre 1 000 milliards pour le modèle GPT-4 sorti en 2024... Au-delà de la consommation d'électricité, l'entraînement de GPT-3 a consommé 700 000 litres d'eau en 15 jours pour le refroidissement des centres de données, pour une empreinte en eau totale (impact lors de la production d'électricité<sup>7</sup>) de 3,5 millions de litres (Li *et al.*, 2023).

Concernant la phase d'usage, les données restent encore très limitées. Il s'agit *a priori* de l'étape la moins intensive en électricité, mais celle qui se déroulera le plus fréquemment, notamment pour les services les plus employés, comme ChatGPT. Il est ainsi possible que la part des émissions de gaz à effet de serre (GES) liées à l'usage puisse finalement dépasser celle liée à l'entraînement, en fonction du nombre d'utilisateurs et de la fréquence d'utilisation du service, notamment dans le cas de l'IA générative (Data For Good, 2023). Notons qu'il y a une interdépendance forte entre ces deux phases : les utilisateurs finals, de plus en plus nombreux, vont demander une précision accrue.

Ces travaux doivent être approfondis. Et pour cela, un meilleur accès aux données des fournisseurs de services d'IA est nécessaire, pour permettre une amélioration de la quantification des impacts environnementaux de l'IA et ainsi identifier des leviers d'action pour réduire son empreinte environnementale.

## La nécessaire réduction des impacts environnementaux de l'IA par l'écoconception

Sans attendre davantage de données, l'application de principes d'écoconception à l'intelligence artificielle pourrait dès à présent permettre de minimiser les impacts environnementaux associés à son développement rapide. Par ailleurs, l'efficacité et la sobriété des modèles pourraient offrir un avantage compétitif aux solutions ainsi développées, notamment en réduisant les capacités énergétiques qu'elles nécessitent.

Des bonnes pratiques et recommandations existent d'ores et déjà à partir de l'état de l'art actuel. À titre d'exemple, le référentiel général de l'écoconception des services numériques<sup>8</sup> (RGESN, Arcep et Arcom, 2024) intègre une série de questions et de pratiques que les concepteurs peuvent mettre en œuvre pour construire une démarche volontaire

---

<sup>6</sup> Comme précisé par la recommandation UIT-T L. 1410 sur l'analyse de cycle de vie des services numériques ou par le référentiel par catégorie de produit (RCP) de l'Ademe sur les services numériques.

<sup>7</sup> L'empreinte eau se base notamment sur l'indicateur EWIF (Energy Water Intensify Factor) qui quantifie la quantité d'eau nécessaire pour la production d'électricité. Selon le rapport annuel de Microsoft pour l'année 2022, l'empreinte eau était de 6,4 milliards de litres.

<sup>8</sup> Prévu par la loi relative à la réduction de l'empreinte environnementale du numérique en 2021, ce document de référence a été élaboré par l'Arcep et l'Arcom en lien avec l'Ademe et en collaboration avec la Dinum, la Cnil et l'Inria.



d'écoconception pour leurs services intégrant des systèmes d'IA. Ce dernier mentionne par exemple la nécessité de définir le besoin du service et des utilisateurs cibles afin d'adapter la conception du modèle d'IA envisagé, en évitant des développements accessoires dont l'utilité n'est pas démontrée. Après avoir défini la pertinence du modèle, il est aussi nécessaire de choisir les options les plus sobres, par exemple en utilisant des modèles pré-entraînés à chaque fois que possible. Il s'agit également de prévoir des modalités d'entraînement des algorithmes qui sont proportionnées aux besoins essentiels du service en termes de quantité, de fréquence, de compression ou de données collectées.

Ces recommandations du RGEN ont depuis été complétées par le référentiel spécialisé sur l'IA frugale (Afnor, 2024). Le document met en avant des définitions communes, des méthodologies d'évaluation environnementale, ainsi que des bonnes pratiques d'écoconception pour les services fondés sur l'IA.

Ainsi, ces travaux au niveau national constituent de premières briques pour œuvrer au développement d'une IA durable. Ces thèmes doivent être portés aux niveaux européen et international, afin de permettre le développement d'IA sobres, plus responsables et plus transparentes, au bénéfice de la compétitivité, de l'environnement et de grands enjeux de société. La mise en place d'un code de conduite sur l'écoconception de l'IA au niveau communautaire serait déjà une première étape pour tendre vers cet objectif. Plus largement, il serait intéressant que les enjeux environnementaux puissent être intégrés par défaut aux politiques d'innovation soutenant l'industrie de l'IA. Ces IA sobres européennes pourraient alors bénéficier d'un avantage concurrentiel face aux grands acteurs déjà établis.

Ces défis appellent à mobiliser toutes les expertises pour construire un avenir désirable d'internet, à l'heure de l'IA générative. L'Arcep collaborera avec l'ensemble des parties prenantes, experts et chercheurs intéressés pour étoffer ce diagnostic et les pistes de recommandations. Nous pouvons agir pour que l'IA générative se développe au service des utilisateurs et de l'innovation sur internet, tout en tenant compte des enjeux écologiques contemporains.

## BIBLIOGRAPHIE

AFNOR, MINISTÈRE DE LA TRANSITION ÉCOLOGIQUE ET DE LA COHÉSION DES TERRITOIRES (2024), « Référentiel général pour l'IA frugale ».

ARCEP (2024), "Arcep's contribution to the call for contributions on competition in generative AI".

ARCEP (2024), « L'état d'internet en France », rapport d'activité, chapitre 8.

ARCEP & ARCOM (2024), « Référentiel général de l'écoconception des services numériques ».

DATA FOR GOOD (2023), « Les grands défis de l'IA générative ».

GARTNER (2024), "Predicts 2024: How GenAI Will Reshape Tech Marketing".

IEA (2024), "Electricity 2024".

LIGOZAT A.-L., LEFÈVRE J., BUGEAU A. & COMBAZ J (2022), "Unraveling the Hidden Environmental Impacts of AI Solutions for Environment Life Cycle Assessment of AI Solutions", *Sustainability*, 14, pp. 51-72.

LI P., JIANYI Y., MOHAMMAD A. I. & SHAOLEI R. (2023), "Making AI less 'thirsty': Uncovering and addressing the secret water footprint of AI models".

LUCCIONI A.-S., VIGUIER S. & LIGOZAT A.-L. (2023), “Estimating the carbon footprint of BLOOM, a 176B Parameter Language Model”, *Journal of Machine Learning Research*, 24 (253), pp. 1-15.

SHAH C. & BENDER E. (2024), “Envisioning Information Access Systems: What Makes for Good Tools and a Healthy Web?”, *ACM Trans. Web*, 18, 3, Article 33, pp. 1-24.

SHAN S. CRYAN J., WENGER E, ZHENG H., HANOCJA R. & ZHAO B. (2023), “Glaze: Protecting artists from style mimicry by Text-to-Image Models”, Proceedings of USENIX Security Symposium.

SHUMAILOV I. *et al.* (2024), “AI models collapse when trained on recursively generated data”, *Nature*, 631, pp. 755-759.

# Intégrer l'IA dans un service éco-conçu : oxymore ou réalité ?

**Par Christophe CLOUZEAU**

Expert Green-UX et Écoconception numérique chez Temesis

**Vincent COURBOULAY**

Maître de conférences HDR en informatique à La Rochelle Université

**Mathieu DELEMME**

Président de Ctrl-a

**Jean-Luc MARINI**

Directeur de l'agence de Lyon et du Laboratoire IA d'OpenStudio

**Emmanuel NURIT**

PDG et co-fondateur d'OpenStudio

**Romuald RIBAUT**

Vice-président de l'AGIT

**Et Claire VERDIER**

*Datascientist* et doctorante en mathématiques appliquées chez OpenStudio

L'intégration de l'intelligence artificielle (IA) dans un service éco-conçu soulève des paradoxes majeurs. D'un côté, l'écoconception cherche à réduire l'impact environnemental des services numériques en prônant sobriété et optimisation. De l'autre, les IA, et notamment les IA génératives, nécessitent d'énormes quantités d'énergie, d'eau et de ressources informatiques qui utilisent des métaux critiques et des terres rares.

Face à une consommation électrique liée à l'IA qui pourrait bientôt rivaliser avec celle de certains pays, l'idée même d'associer IA et écoconception semble contradictoire. Pourtant, des pistes émergent pour concevoir des IA plus responsables. Mais l'enjeu principal reste de déterminer si et dans quelles conditions leur utilisation peut réellement s'intégrer à une démarche durable.

En 2023, les universités de Californie, du Colorado et du Texas ont calculé que poser 25 questions à ChatGPT revenait à verser un demi-litre d'eau douce par terre (Li, Yang, Islam et Ren, 2023). Cette équivalence très parlante nous fait comprendre à quel point les intelligences artificielles (IA) génératives font flamber la facture environnementale du secteur du numérique. D'après un article publié dans la revue *Joule* (10 octobre 2023), Alex de Vries, étudiant en thèse à l'université libre d'Amsterdam, s'est basé sur les prévisions de production de nouveaux serveurs pour estimer que d'ici 2027 la consommation mondiale d'électricité liée à l'IA passerait de 85 à 134 Wh par an, soit l'équivalent de la consommation annuelle d'électricité de la Suède. Cela signifie que lorsque l'IA sera développée à grande échelle, elle représenterait 0,5 % de la consommation d'électricité mondiale, toujours selon les calculs d'Alex de Vries.

Dans ce contexte où nous sommes encore au stade des estimations et où il est encore très difficile de mesurer finement l'impact global des IA sur l'environnement, l'association des deux termes « IA » et « écoconception » ressemble à un bel oxymore. Comment un service numérique pourrait-il rester éco-conçu en intégrant un modèle d'IA réputé énergivore ? Cette question soulève effectivement des paradoxes.

Il est donc indispensable tout d'abord de revenir sur ce que recouvre exactement l'écoconception, pour se demander ensuite si l'IA pourrait avoir sa place dans cette démarche. Attention également à ne pas résumer « les » IA aux IA génératives qui sont basées sur d'énormes modèles, entraînés avec de gigantesques quantités de données. Les IA sont elles-mêmes des solutions numériques dont la conception peut suivre de bonnes pratiques pour les rendre moins voraces en ressources naturelles. Encore faut-il se poser la question de la nécessité absolue d'utiliser l'IA pour donner vie à un service numérique, surtout si celui-ci revendique une volonté de diminuer son impact sur le vivant.

## ÉTAT DE L'ART SCIENTIFIQUE SUR L'ÉCOCONCEPTION

L'écoconception vise à intégrer des critères de durabilité environnementale dès les premières étapes de conception de produits et services. Née dans les années 1960 avec la montée de la conscience environnementale, l'écoconception a évolué pour devenir un domaine interdisciplinaire, touchant aujourd'hui des secteurs variés, de l'industrie manufacturière au numérique. En réponse aux enjeux climatiques et à l'épuisement des ressources naturelles, l'écoconception cherche à réduire l'impact écologique à toutes les étapes du cycle de vie des produits, de leur production à leur fin de vie, et à promouvoir une économie circulaire.

Dans le domaine des services numériques, l'écoconception *design* occupe une place croissante, car les infrastructures numériques – bien que souvent perçues comme immatérielles – sont énergivores et consomment des ressources naturelles importantes. Les centres de données, les réseaux de télécommunication et les dispositifs électroniques contribuent de manière significative aux émissions de gaz à effet de serre. Ainsi, l'écoconception de services numériques propose des stratégies pour réduire l'empreinte environnementale des services numériques, en optimisant l'utilisation des ressources, en améliorant l'efficacité énergétique et en concevant des solutions plus durables.

Dans l'Afnor Spec 2201 concernant l'écoconception des services numériques, on trouve la définition suivante : « Approche méthodique qui prend en considération les aspects environnementaux du processus de conception et développement dans le but de réduire les impacts environnementaux négatifs tout au long du cycle de vie d'un produit. Appliquée aux services numériques, l'écoconception a pour objectif de réduire ou limiter les impacts environnementaux de ces services, de l'expression des besoins jusqu'à leur fin de vie ».

Cette définition s'accompagne de nombreux défis dont, par exemple, la prise en compte des usages réels et des comportements des utilisateurs. En effet, comprendre l'utilisation réelle des services et modéliser les comportements d'utilisation permet d'adapter les solutions pour qu'elles soient non seulement performantes mais aussi écologiquement optimisées.

Ainsi, l'écoconception des services numériques s'impose comme une réponse essentielle aux enjeux contemporains de la durabilité numérique, ouvrant la voie vers des pratiques de conception responsables, transparentes et respectueuses des limites planétaires.

Si l'écoconception des services numériques s'attache à réduire l'impact environnemental des infrastructures numériques et à repenser les usages, elle doit également intégrer les avancées technologiques, comme l'intelligence artificielle, dont l'empreinte environne-

mentale suscite de nombreuses interrogations. Dès lors, une question cruciale se pose : l'IA, souvent perçue comme énergivore, émettrice de gaz à effet de serre, consommatrice d'eau ou gourmande en matières premières peut-elle contribuer à la conception de services numériques plus durables tout en s'inscrivant dans une démarche d'écoconception ?

## L'IA PEUT-ELLE AVOIR SA PLACE DANS UN SERVICE ÉCO-CONÇU ?

Par définition, éco-concevoir un service numérique est une démarche d'amélioration continue dont l'objectif, dès la phase amont de sa réalisation, est de réduire son empreinte environnementale. L'apport d'une IA, ou de sa composante générative, pour agrémenter ce service semble donc être une approche totalement paradoxale lorsque l'on connaît les impacts sociaux et environnementaux majeurs de cette technologie.

La démarche d'écoconception d'un service numérique repose sur une réflexion approfondie qui s'articule autour de plusieurs dimensions complémentaires :

- L'utilité : s'interroger sur la pertinence du service et des fonctionnalités proposées. Chaque élément doit répondre à un véritable besoin autour de cette utilité et au regard des impacts environnementaux associés.
- L'unité fonctionnelle : définir clairement la fonction principale du service pour éviter toute dispersion inutile et, surtout, tout ajout de fonctionnalités inutiles.
- La sobriété des contenus : limiter les contenus au strict nécessaire pour réduire l'impact environnemental et pour répondre efficacement aux besoins identifiés.
- Les fonctionnalités et interactions globales : rationaliser les interactions entre l'utilisateur et le service pour privilégier simplicité et efficacité.
- Le réseau et les serveurs dimensionnés au strict nécessaire : déployer l'architecture technique et la puissance nécessaire au sein des serveurs de calculs et de stockage, sans superflu mais adaptable à la consommation du service.

Intégrer une IA qui ne serait pas elle-même réfléchie en termes de frugalité dans une démarche de conception responsable d'un produit ou d'un service, revient à affaiblir, voire annihiler, la démarche de frugalité de cet objet ou ce service.

### Tout n'est pas binaire

Pour prendre un exemple concret : intégrer un « *chatbot* greffé à une IA entraînée pour répondre correctement à un cadre de prompts définis » à un service numérique qui veut réduire ses impacts paraît donc *a priori* une hérésie. Mais, si ce même service permet de réduire des impacts négatifs *via* l'apport d'un algorithme puissant qui aide les utilisateurs, alors les effets positifs induits peuvent potentiellement compenser l'impact de ce service et de l'implémentation de l'IA.

Cela reste cependant une hypothèse : cette « compensation » n'est pas garantie, surtout lorsque l'on parle d'un service numérique de confort ou d'un service secondaire qui n'est pas le cœur de l'activité d'une organisation par exemple.

Néanmoins, la démarche d'écoconception s'appuie sur un pilier fondamental qui consiste à challenger les besoins, de telle sorte qu'*in fine* une fonctionnalité soit bien utile, utilisée et utilisable (les 3U). Cette notion de frugalité est clé pour se questionner en profondeur et ne retenir que les besoins essentiels. Les solutions mises en œuvre se doivent ensuite d'être sobres, et le recours à l'IA envisagé uniquement en ultime option.

## La désillusion après l'illusion

Afin d'intégrer une technologie d'IA ou sa composante générative à un service numérique, il faudrait pouvoir en évaluer l'utilité au regard des impacts directs, des impacts indirects et des effets rebonds, à la fois pour la phase d'entraînement et la phase d'utilisation ! Vaste défi, eu égard à l'estimation des gains induits par l'utilisation de l'IA...

En revanche, supposons que l'activité d'une entreprise soit tout aussi polluante que désastreuse sur le plan social. L'impact de ses services numériques est très probablement minime (quelques pourcents) par rapport à celui de son activité vitale. Si l'intégration de l'IA permet de réduire l'impact global de l'organisation, l'approche est sans doute positive au regard des mesures d'impacts qui viendront le confirmer.

Alors que l'IA est souvent présentée comme une solution aux défis environnementaux, les preuves concrètes de son impact global et durable restent à ce jour à démontrer. En revanche, les besoins colossaux en ressources, en consommation d'eau et en électricité mettent en évidence chaque jour les impacts négatifs de ces technologies et ne font qu'éloigner drastiquement les promesses techno-solutionnistes. L'Agence Internationale de l'Énergie prévoit jusqu'au doublement de la consommation électrique mondiale liée aux *data centers* en 2026 par rapport aux besoins de 2022 !

Lorsque l'on sait que Microsoft réactive une centrale nucléaire fermée en Pennsylvanie<sup>1</sup> et s'engage à la maintenir pendant 20 ans pour alimenter ses *data centers* spécialisés en IA, il y a de quoi se questionner : est-ce vraiment utile et raisonnable d'ajouter une IA au sein d'un service numérique dit « éco-conçu » ? D'autant plus qu'avec la maturité des usages et des études publiques, le coût environnemental est maintenant plutôt sur la phase de l'utilisation que de l'entraînement (de l'ordre de 200 à 1 000 fois plus). L'IA générative est donc aujourd'hui tiraillée entre les bénéfices qu'elle pourrait apporter et son coût environnemental (sans omettre le coût social).

Quoi qu'il en soit, prendre en compte l'utilité d'une IA au sein d'un service est un raisonnement qui doit se faire dès le début de la démarche d'écoconception, au moment de la réflexion sur la valeur apportée à l'organisation et aux utilisateurs face aux impacts directs, indirects et aux effets rebonds possibles.

## Du côté des impacts humains

Enfin, comment revendiquer une quelconque écoconception sans prendre en compte les enjeux sociétaux liés aux usages généralisés de l'IA ? On parle d'ailleurs de plus en plus d'éco-socio-conception.

Dans un article publié dans la revue *Nature Human Behavior*<sup>2</sup>, des scientifiques confirment « qu'en déléguant une partie de notre raisonnement à l'intelligence artificielle, c'est toute la pensée humaine qui pourrait s'en trouver modifiée. Si nous acceptons passivement les solutions proposées par l'IA, nous risquons de perdre notre capacité à penser de manière autonome et à développer des idées innovantes ». Voulons-nous des médecins sans plus aucun savoir-faire de diagnostic car soumis aux recommandations d'une IA et qui serait incapable de corriger un biais généré par l'IA ?

---

<sup>1</sup> « Pour ses *data centers*, Microsoft réactive la centrale de Three Mile Island », *Le Monde Informatique*, 23 septembre 2024, <https://www.lemondeinformatique.fr/actualites/lire-pour-ses-datacenters-microsoft-reactive-la-centrale-de-three-mile-island-94776.html>

<sup>2</sup> Chiriatti M., Ganapini M., Panai E. *et al.* (2024), "The case for human-AI interaction as system 0 thinking", *Nat Hum Behav*, 8, pp. 1829-1830, <https://doi.org/10.1038/s41562-024-01995-5>

Le cerveau humain a lui aussi besoin de s'entraîner. Même en admettant que l'IA puisse avoir une utilité pour l'environnement et soit intégrée dans une démarche d'écoconception, le risque pour l'humanité n'est-il pas de ne plus savoir résoudre ses problèmes elle-même en les sous-traitant ainsi à la machine ?

L'IA pose en effet des questions philosophiques, éthiques et déontologiques extrêmement préoccupantes. Même les pionniers mettent en garde sur les catastrophes que pourraient provoquer les IA si elles ne sont pas maîtrisées. John Hopfield, scientifique américain lauréat du Nobel de physique en 2024 pour ses travaux sur l'IA, a déclaré que les dernières avancées technologiques de cette discipline étaient « très inquiétantes »<sup>3</sup>.

Pour tenter de maîtriser ses impacts environnementaux, des chercheurs et des professionnels de ce domaine commencent à se mettre d'accord sur de bonnes pratiques pour viser une plus grande frugalité des modèles d'IA. Intégrer dans des services numériques éco-conçus des IA incluant comme critère d'efficacité la limitation de leur empreinte environnementale pourrait donc s'avérer une piste à défricher.

## PEUT-ON ÉCO-CONCEVOIR L'IA ?

Nous ne cesserons jamais de le rappeler, la conception d'une solution d'IA, comme tout produit numérique, va forcément consommer des ressources naturelles et de l'énergie. Toutefois, il est possible de concevoir et mettre en œuvre des solutions d'IA plus respectueuses de l'environnement et qui restent tout à fait fiables (et donc utiles !). Pour cela, il faudra agir sur différents éléments du cycle de vie d'une solution d'IA, depuis sa phase d'initiation jusqu'à son décommissionnement.

De manière générale, toutes les pratiques d'écoconception des services numériques peuvent être appliquées lors du développement d'une solution d'IA qui est avant tout une solution numérique.

Dès la phase d'initiation du projet de conception d'une solution d'IA, il convient de s'interroger sur les différentes méthodes existantes afin de répondre à un besoin nécessitant la conception d'une solution d'IA. Nul besoin de prendre un marteau pour écraser une mouche. Il existe plusieurs types d'IA : connexionniste, symbolique, hybride, plus ou moins consommatrices de ressources. Parfois, pour répondre à un besoin, de simples règles de décisions suffisent. D'où la nécessité de réaliser une étude approfondie du besoin et des solutions susceptibles d'être mises en œuvre afin de sélectionner le type d'IA et la méthode qui permettront de limiter l'impact environnemental de la phase d'expérimentation. Les méthodes d'IA ne sont pas les seules réponses à toutes les problématiques.

La conception et le développement d'une solution d'IA est généralement la plus consommatrice en termes d'énergie, d'eau, mais aussi de ressources informatiques qui utilisent des métaux critiques et des terres rares. Durant cette phase, il existe une multitude de leviers afin de limiter l'empreinte environnementale de la solution.

---

<sup>3</sup> « John Hopfield, Prix Nobel de physique pour ses recherches sur l'intelligence artificielle, met en garde contre les récentes avancées "très inquiétantes" de l'IA », *Le Monde*, 9 octobre 2024, [https://www.lemonde.fr/sciences/article/2024/10/09/john-hopfield-prix-nobel-de-physique-pour-ses-recherches-sur-l-intelligence-artificielle-met-en-garde-contre-les-recentes-avancees-tres-inquietantes-de-l-ia\\_6347030\\_1650684.html](https://www.lemonde.fr/sciences/article/2024/10/09/john-hopfield-prix-nobel-de-physique-pour-ses-recherches-sur-l-intelligence-artificielle-met-en-garde-contre-les-recentes-avancees-tres-inquietantes-de-l-ia_6347030_1650684.html)

## Les données : privilégier la qualité plutôt que la quantité

Les données sont un élément essentiel pour développer une solution d'IA, mais elles sont aussi largement en cause dans son impact environnemental. D'où l'importance de s'assurer que les données utilisées soient vraiment pertinentes pour sa solution. En effet, une mauvaise qualité des données augmente le temps d'apprentissage et les biais d'apprentissage, donc fait grimper la consommation d'énergie et diminue la fiabilité du modèle. Avec des données qualitatives, il est aussi possible de réduire la quantité de données nécessaires afin d'obtenir un modèle d'IA performant. Par ailleurs, la taille du modèle est en partie liée au nombre de données sur lesquelles celui-ci est entraîné. Plus on privilégie des modèles de petite taille, avec des données qualitatives mais en quantité limitée, plus l'empreinte environnementale de la solution sera réduite.

## Compresser les modèles stockés sur les serveurs

Pendant la phase d'entraînement du modèle, il y a encore de bonnes pratiques à appliquer pour limiter au maximum sa consommation d'énergie et d'eau, tout en garantissant sa qualité et sa fiabilité. Les méthodes de compression de modèles ont un intérêt pour réduire l'empreinte des algorithmes d'IA. Certes, cette étape allonge la phase d'entraînement, et alourdit son empreinte environnementale, mais ces méthodes de compression divisent la taille des modèles quasiment par trois, ce qui a un énorme impact au niveau du stockage sur les serveurs. Ces derniers consommant de l'énergie et de l'eau, un modèle plus petit demandera moins de ressources lors de son utilisation et évitera de sur-solliciter les serveurs.

## Attention à l'impact du ré-entraînement des algorithmes

Après la conception et le développement d'une solution d'IA, ré-entraîner un algorithme d'IA est une phase elle aussi consommatrice de ressources naturelles. Évidemment, le mieux est d'éviter de ré-entraîner le modèle trop souvent. Pour cela, on peut définir des critères afin de justifier le besoin du ré-entraînement. Afin de déterminer ces critères, il existe désormais un Référentiel Général de l'Écoconception des Services Numériques<sup>4</sup> élaboré par l'Arcep, l'Arcom et l'Ademe en mai 2024.

## Un nouveau consensus sur la définition de l'IA dite frugale ?

Il est indéniable que l'IA offre de nombreuses opportunités de transformation digitale, d'innovation et de croissance pour les entreprises. De nombreuses entreprises la considèrent comme un nouvel eldorado, bien que les cas d'usage actuels montrent que la rentabilité réelle de ces solutions reste à prouver pour bon nombre d'entre elles<sup>5</sup>.

Si des abus sont à déplorer sur l'utilisation tous azimuts de l'IA et notamment des IA génératives particulièrement énergivores, de nombreux acteurs du domaine de l'IA font la promotion de solutions d'IA dites frugales. Or, jusqu'en juin 2024, il n'existait pas de

---

<sup>4</sup> <https://www.arcep.fr/mes-demarches-et-services/entreprises/fiches-pratiques/referentiel-general-ecoconception-services-numeriques.html>

<sup>5</sup> « IA : les entreprises vont dans le mur prévient McKinsey », *Les Échos*, 29 avril 2024, <https://www.lesechos.fr/industrie-services/services-conseils/ia-les-entreprises-doivent-revoir-leur-approche-selon-mckinsey-2091721>



consensus sur ce qu'était une IA « frugale » ou en tout cas plus respectueuse de l'environnement. Le groupe Afnor et le ministère de la Transition écologique française ont donc lancé un groupe de travail en janvier 2024 pour définir un référentiel général de l'IA frugale. Ce référentiel a pour objectif de favoriser une adoption de l'IA frugale à plus grande échelle, ainsi que de promouvoir une utilisation plus responsable et durable de cette technologie. Les deux acteurs à l'initiative de ce référentiel l'ont inscrit au programme du Comité Européen de Normalisation en Électronique et Électrotechnique (CENELEC) en juillet 2024. Cette proposition a l'ambition d'élever le référentiel au statut de norme européenne, dans le but d'inciter l'ensemble des acteurs européens de l'IA à adopter une approche plus responsable.

Néanmoins, il est important d'être vigilant sur l'approche visant à rendre une solution d'IA plus respectueuse de l'environnement, car – selon le paradoxe de Jevons (économiste britannique du XIX<sup>e</sup> siècle) – une plus grande sobriété dans l'IA pourrait entraîner une très nette augmentation de l'usage de l'IA et donc de sa consommation énergétique (c'est l'effet rebond évoqué plus haut dans cet article).

### **CONCLUSION : L'IA LA MOINS ÉNERGIVORE EST CELLE QUE L'ON NE DÉVELOPPE PAS**

L'intégration de l'IA dans un service éco-conçu illustre le paradoxe actuel du numérique : identifiée comme un levier économique majeur, cette technologie s'oppose frontalement aux principes de frugalité et de sobriété numérique.

Cette opposition se matérialise au regard des objectifs climatiques : alors que la Stratégie Nationale Bas Carbone fixe comme objectif une réduction annuelle de 5 % de nos émissions de gaz à effet de serre d'ici 2030, la consommation énergétique des *data centers* dédiés à l'IA augmente de 15 à 20 % chaque année.

Cette réalité impose des conditions strictes : les bénéfices environnementaux apportés par l'IA doivent être supérieurs à sa propre empreinte écologique. Son utilisation doit rester un dernier recours, avec une maîtrise de son empreinte environnementale, de ses effets rebonds tout en considérant ses impacts tant sociaux que cognitifs.

L'intégration de l'IA nécessite donc une approche systémique rigoureuse, transformant les contraintes environnementales en opportunités d'innovation. La vraie question n'est plus de savoir si l'IA peut être intégrée à un service éco-conçu, mais dans quels cas exceptionnels son utilisation se justifie face aux enjeux environnementaux, éthiques et sociaux.

# L'intelligence artificielle et les droits humains : les insuffisances du cadre européen

Par Thomas DUMORTIER

Conseiller juridique à la Commission nationale  
consultative des droits de l'Homme

L'utilisation de l'intelligence artificielle inquiète autant qu'elle fait naître des espoirs. Après l'ère de « l'éthique de l'IA », dans les années 2010, l'Union européenne et le Conseil de l'Europe ont souhaité poser un cadre plus ferme qui s'est récemment concrétisé, respectivement dans un règlement et une Convention-cadre.

Si ces textes affichent une volonté de protéger les droits humains, ils suscitent un certain nombre de préoccupations en la matière. Ces dernières concernent d'une part la délimitation insuffisante des systèmes d'IA inacceptables en raison de leur atteinte aux droits humains et, d'autre part, l'incertitude quant à la portée des garanties instaurées à tous les stades de vie d'un système d'IA pour préserver ces droits, à commencer par l'étude d'impact sur les droits humains.

Depuis le milieu des années 2010, les progrès de l'« intelligence artificielle » suscitent des réactions contrastées dans l'espace public allant d'un engouement sans réserve, chez certains, pour la résolution technique des défis de demain, aux craintes exprimées par d'autres de voir les humains supplantés par les machines<sup>1</sup>. De manière concomitante, une « déferlante éthique »<sup>2</sup> s'est répandue dans le monde entier : des documents de « droit souple » ont ainsi fleuri un peu partout (chartes, lignes directrices, recommandations, etc.), élaborés principalement par des entreprises de la « tech » ou des organisations internationales. Ces normes dénuées de sanction étaient destinées à endiguer les dérives de l'IA vis-à-vis des principes et des droits fondamentaux tels que le respect de la dignité humaine, de la vie privée, ou encore la non-discrimination. Ce faisant, la protection de droits pourtant inscrits au sommet de la hiérarchie des normes paraissait bien fragile.

Soucieuse d'aller plus loin, la Commission européenne a fait connaître en 2021 sa proposition établissant des règles harmonisées concernant l'intelligence artificielle. Si elle visait à « garantir un niveau élevé de protection [des] droits fondamentaux », sa raison d'être était principalement « d'assurer le bon fonctionnement du marché unique, en créant des conditions propices au développement et à l'utilisation d'une IA digne de confiance dans

<sup>1</sup> Bellon A. et Velkovska J. (2023), « L'intelligence artificielle dans l'espace public : du domaine scientifique au problème public », *Réseaux*, n°240, pp. 31-70.

<sup>2</sup> Castets-Renard C. (2020), « Comment construire une intelligence artificielle responsable et inclusive ? », *Recueil Dalloz* 2020, p. 225.

l'Union »<sup>3</sup>. Sans le formuler expressément, la Commission s'en remettait aux travaux alors en cours au sein du Conseil de l'Europe relatifs à l'encadrement du développement, de la conception et de l'application de l'intelligence artificielle, fondé sur les normes du Conseil de l'Europe en matière de droits de l'homme, de démocratie et d'État de droit<sup>4</sup>. C'est dans ce contexte que la CNCDH a adopté, en avril 2022, son avis relatif à l'impact de l'intelligence artificielle sur les droits fondamentaux. Depuis, le règlement de l'Union européenne (« règlement IA ») et la Convention-cadre du Conseil de l'Europe ont été adoptés, respectivement le 13 juin et le 17 mai 2024<sup>5</sup>.

Les observations et recommandations formulées par la CNCDH dans son avis de 2022 restent toutefois pleinement d'actualité<sup>6</sup>. Reprenant à son compte les préoccupations exprimées par un certain nombre d'organisations nationales et internationales<sup>7</sup>, la Commission appelait d'une part à prohiber certains cas d'usage de l'IA – les « lignes rouges » – et, d'autre part, à mettre en place des garanties pour s'assurer d'un déploiement des systèmes d'IA respectueux des droits fondamentaux. Cette approche repose sur l'idée qu'en dehors de certains cas d'usage de l'IA jugés inadmissibles en raison de leur atteinte disproportionnée aux droits fondamentaux, l'IA doit faire l'objet d'un contrôle et d'une supervision à tous les stades du cycle de vie du système déployé.

Il convient ici de distinguer entre deux types d'incidence sur les droits fondamentaux, selon que l'on s'intéresse à la technologie (la vision par ordinateur par exemple) ou à des cas d'usage/finalité (l'utilisation de la vision par ordinateur dans le domaine médical ou pour la détection de comportements suspects par exemple). Les préoccupations relatives à la technologie employée, en particulier en cas d'apprentissage machine sur des bases de données massives, ont principalement trait aux risques de discrimination, mais également au défaut d'explicabilité du fonctionnement du système d'IA. D'après certains experts, ces risques peuvent être plus ou moins corrigés<sup>8</sup>. Lorsque l'on aborde les cas d'usage, c'est la sensibilité du domaine dans lequel la technologie est déployée qui retient l'attention : l'utilisation de l'analyse automatisée des images à des fins de surveillance de l'espace public – l'identification biométrique à distance en temps réel par exemple – porte une grave atteinte au respect de la vie privée et est susceptible de remettre en cause la liberté d'aller et venir des citoyens ou encore la liberté de manifester.

Si la réglementation européenne représente une avancée incontestable pour la préservation des droits fondamentaux, les motifs de préoccupation ne manquent pas. D'abord, les lignes rouges tracées par cette réglementation ne vont sans doute pas assez loin.

---

<sup>3</sup> Pour s'en convaincre, indépendamment de l'exposé des motifs, on peut se référer aux articles visés par le règlement : Art. 16 et 114, ce dernier renvoie aux objectifs énoncés à l'article 26.

<sup>4</sup> Le 11 septembre 2019, le Comité des ministres du Conseil de l'Europe a créé le Comité *Ad Hoc* sur l'Intelligence Artificielle – CAHAI. Celui-ci était chargé d'examiner sur la base de larges consultations multipartites, la faisabilité et les éléments potentiels d'un cadre juridique pour le développement, la conception et l'application de l'intelligence artificielle, fondés sur les normes du Conseil de l'Europe dans le domaine des droits de l'homme, de la démocratie et de l'État de droit.

<sup>5</sup> La Convention est ouverte à la signature des États depuis septembre 2024, y compris des pays non européens.

<sup>6</sup> CNCDH, Avis relatif à l'impact de l'intelligence artificielle sur les droits fondamentaux, JORF n°0091 du 17 avril 2022, Texte n°99.

<sup>7</sup> En particulier : la Commission nationale de l'informatique et des libertés (CNIL), le Défenseur des droits (DDD), l'Agence des droits fondamentaux de l'Union européenne.

<sup>8</sup> Par exemple, en s'efforçant de recourir à des données d'apprentissage « nettoyées » de biais éventuels.

Ensuite, et plus largement, la portée de certaines garanties demeure encore incertaine, à commencer par les études d'impact sur les droits fondamentaux.

## DES LIGNES ROUGES INSUFFISANTES

Le règlement IA prohibe l'utilisation des systèmes d'IA pour des finalités « contraires aux valeurs de l'Union relatives au respect de la dignité humaine, à la liberté, à l'égalité, à la démocratie et à l'État de droit, ainsi qu'aux droits fondamentaux consacrés dans la Charte, y compris le droit à la non-discrimination, le droit à la protection des données et à la vie privée et les droits de l'enfant »<sup>9</sup>. Parmi ces systèmes d'IA, énumérés à l'article 5, figurent les techniques de manipulation, la *scoring social*, la prédiction du risque qu'une personne commette une infraction, les systèmes d'identification biométrique à distance en temps réel dans des espaces accessibles au public à des fins répressives. Si la CNCDH souscrit à ces interdictions, elle regrette l'étendue de certaines exceptions envisagées par la réglementation.

Le règlement IA assortit de nombreuses exceptions son interdiction d'utiliser des systèmes d'identification biométrique à distance en temps réel dans des espaces accessibles au public à des fins répressives : la recherche de victimes d'enlèvement ou d'exploitation sexuelle, ou encore de personnes disparues ; la prévention d'une menace imminente pour la vie ou la sécurité de personnes physiques, ou encore d'une attaque terroriste ; et, surtout, la localisation ou l'identification d'une personne soupçonnée d'avoir commis l'une des seize infractions prévues par le règlement<sup>10</sup>. L'étendue de ces exceptions remet grandement en cause l'interdiction de principe. De son côté, la CNCDH n'admet qu'une seule exception à cette technologie particulièrement intrusive : dès lors que son utilisation est strictement nécessaire, adaptée et proportionnée pour la prévention d'une menace grave et imminente pour la vie ou la sécurité des personnes et celle des ouvrages, installations et établissements d'importance vitale.

Sur les technologies de reconnaissance d'émotions, le règlement en interdit l'usage sur le lieu de travail et dans les établissements d'enseignement, sauf lorsque cela peut répondre à des besoins d'ordre médical ou de sécurité. Par conséquent, au-delà des lieux de travail et d'enseignement, il est possible d'y recourir. Ces systèmes de reconnaissance reposent sur un postulat dont la scientificité fait défaut – à savoir que les émotions sont détectables par des expressions du visage ou, plus largement, par des manifestations corporelles. C'est pourquoi la CNCDH en recommande l'interdiction. Elle admet toutefois leur utilisation dès lors qu'ils visent à renforcer l'autonomie des personnes : malgré ses approximations, cette technologie d'IA peut par exemple favoriser des activités d'apprentissage auprès de personnes en situation de handicap ou s'avérer utile pour d'autres interactions humain-machine.

Les systèmes d'IA conçus pour prédire le comportement délictueux d'un individu sont également très controversés. La CNCDH se félicite que le règlement en prohibe l'utilisation. Elle s'interroge dès lors sur l'autorisation d'utiliser un tel système pour « évaluer des risques, y compris des risques pour la sécurité (...), posés par une personne physique qui a l'intention d'entrer ou qui est entrée sur le territoire d'un État membre ». Cela pose un problème de cohérence. Surtout, l'atteinte sévère aux droits fondamentaux des migrants engendrée par ce type d'IA aurait dû justifier son interdiction.

On aurait pu espérer que la Convention-cadre du Conseil de l'Europe dresse également un certain nombre de lignes rouges. Le projet présenté en juin 2022 aux États membres

---

<sup>9</sup> Règlement IA, cons. 28.

<sup>10</sup> La liste de ces infractions figure à l'annexe II du règlement.

par le Comité sur l'intelligence artificielle incluait un article 14 qui allait en ce sens. Rapidement, cependant, l'idée d'une clause de prohibition a été écartée des négociations. Les États hostiles à des prohibitions ciblées insistaient sur la nature particulière de ce traité, destiné à n'être qu'une « Convention-cadre », autrement dit une Convention définissant simplement des principes directeurs.

À vrai dire, cette objection a de quoi surprendre. D'abord, parce que les États sont libres bien évidemment d'inclure les dispositions qu'ils souhaitent dans ce texte, indépendamment de son intitulé : il n'existe pas de régime spécifique aux Conventions-cadres. Ensuite, et surtout, parce que d'autres Conventions-cadres du Conseil de l'Europe contiennent des interdictions précises. La Convention d'Oviedo de 1997, notamment, est présentée sur le site du Conseil de l'Europe comme une « Convention-cadre visant à protéger la dignité et l'identité de tous les êtres humains et à garantir à toute personne, sans discrimination, le respect de son intégrité et de ses autres droits et libertés fondamentales à l'égard des applications de la biologie et de la médecine »<sup>11</sup>. Or, ce texte interdit expressément la constitution d'embryons humains aux fins de recherche.

Finalement, la Convention-cadre sur l'IA laisse aux États le soin d'apprécier si un système d'IA est incompatible avec le respect des droits de l'homme, le fonctionnement de la démocratie ou l'État de droit, justifiant, le cas échéant, son interdiction. Il est particulièrement regrettable également que les systèmes d'IA dédiés à la « sécurité nationale » échappent à son champ d'application.

Pour les systèmes d'IA considérés par l'UE comme attentatoires aux droits fondamentaux, mais pas suffisamment pour justifier une interdiction – les systèmes à « haut risque » – le règlement fait peser sur les fournisseurs d'un système d'IA une série d'exigences relatives à la transparence du système, aux données utilisées, à la gestion des risques, etc. S'y ajoute pour les organismes publics, et les entités privées fournissant un service public, l'obligation d'évaluer en amont du déploiement d'un système d'IA à haut risque son impact sur les droits fondamentaux (HRIA). Ces études d'impact figurent également en bonne place dans la Convention-cadre du Conseil de l'Europe<sup>12</sup>.

## LES ÉTUDES D'IMPACT SUR LES DROITS FONDAMENTAUX : UNE GARANTIE SUJETTE À CAUTION

De manière générale, les HRIA sont censées identifier les incidences négatives d'un système d'IA sur les droits fondamentaux tout en incluant les mesures susceptibles d'y remédier. Le premier motif de préoccupation renvoie à l'évaluation de tous ces éléments : à partir de quel seuil d'atteinte à un droit fondamental, les mesures compensatrices ne suffiront plus à en garantir la préservation ? C'est typiquement le genre de question que se posent les juridictions en charge de faire respecter les droits fondamentaux. La question se pose de savoir si les organismes désireux de déployer un système d'IA seront en mesure de procéder à ce type d'examen qui requière, outre des compétences juridiques, indépendance et impartialité. La crainte que ces HRIA ne soient qu'une « formalité » et n'engagent pas une véritable réflexion au sein des organismes utilisateurs d'un système d'IA à haut risque est confortée par le dernier paragraphe de l'article 27 du règlement IA : celui-ci prévoit l'élaboration par le Bureau de l'IA d'un modèle de questionnaire, y compris

---

<sup>11</sup> <https://www.coe.int/fr/web/bioethics/oviedo-convention>

<sup>12</sup> Art. 16, 1. : « Chaque Partie (...) adopte ou maintient des mesures afin d'identifier, d'évaluer, de prévenir et d'atténuer les risques posés par les systèmes d'intelligence artificielle en tenant compte des impacts réels et potentiels sur les droits de l'homme, la démocratie et l'État de droit ».

au moyen d'un outil automatisé, afin d'aider et de « simplifier » la tâche des déployeurs dans la réalisation d'une HRIA.

Parmi les mesures mises en place pour atténuer les risques engendrés par un système d'IA pour les droits fondamentaux figure le contrôle humain. La CNCDH a cependant alerté, à plusieurs reprises, sur la propension de chacun à suivre les recommandations ou alertes produites par un programme informatique (parfois appelé « biais d'automatisation »). Elle l'a fait encore récemment dans son avis sur la surveillance de l'espace public qui traite de la vidéosurveillance automatisée (VSA) autorisée à titre expérimental par la loi sur la sécurité des Jeux Olympiques et Paralympiques<sup>13</sup>. Cette dernière précise que les logiciels ne procèdent qu'à des « signalement d'attention » et « demeurent en permanence sous le contrôle » d'un agent humain. Les agents en charge de la supervision de la VSA seront toutefois d'autant plus tentés de mobiliser une équipe sur le terrain pour vérifier le risque associé à une alerte, que le logiciel conservera l'enregistrement de tous les signalements : l'agent sera donc incité à les suivre. Il se trouvera ainsi davantage en position d'exécutant que de contrôleur vis-à-vis de la machine. La question se pose dès lors de savoir comment préserver l'autonomie des agents dans un tel dispositif et de garantir ainsi un contrôle effectif sur les résultats du système d'IA.

La qualité de l'étude d'impact reposera en grande partie sur le niveau d'implication des parties prenantes, incluant à la fois les agents en charge de l'utilisation du système d'IA mais aussi les personnes qui seront visées par des décisions issues directement ou indirectement du système. C'est pourquoi les représentants d'organisations syndicales doivent être consultés et associés en amont du déploiement d'un système d'IA afin d'examiner son impact sur l'organisation du travail et sur les droits des salariés. Des associations de défense des droits humains doivent également être impliquées dans la réalisation de l'étude d'impact. Elles seront les mieux à même d'identifier les risques liés à l'automatisation de procédures ou de dispositifs déjà souvent eux-mêmes attentatoires aux droits fondamentaux. L'exemple précédemment évoqué de la VSA en est une parfaite illustration, en ajoutant une couche logicielle à des dizaines de milliers de caméras de surveillance qui ont proliféré à bas-bruit dans l'espace public ces 20 dernières années.

---

<sup>13</sup> CNCDH, Avis sur la surveillance de l'espace public, JORF n°0155 du 2 juillet 2024, Texte n°61.

# IA générative et mésinformation

Par Nicolas CURIEN

Membre de l'Académie des technologies

La mésinformation désigne des informations, soit involontairement erronées, soit délibérément falsifiées, auquel cas on parle de désinformation, d'infox ou de *fake news*. Ces pathologies de l'information sont un phénomène ancien, mais l'essor des technologies numériques et de l'IA a considérablement accru leur impact. Aujourd'hui, elles se propagent sur Internet de manière virale, car les contenus faux, plus attractifs que les vrais, sont favorisés par les modèles économiques des grandes plateformes numériques reposant sur la captation de l'attention des internautes.

Dans un rapport paru en décembre 2024, l'Académie des technologies alerte sur les dangers de la mésinformation, potentiellement amplifiés par l'IA générative. Ce rapport souligne également les avancées d'une IA curative au service de la lutte contre la mésinformation, ainsi que la construction d'un cadre réglementaire européen pionnier en la matière. Il formule plusieurs recommandations ciblées, visant à renforcer la résilience de la société face aux *fake news* et à promouvoir l'honnêteté de l'information.

## À LA CONFLUENCE DES COURANTS NUMÉRIQUE ET COGNITIF

Nous rendons compte ici de travaux menés de juillet 2023 à novembre 2024 au sein de l'Académie des technologies, sur le thème IA générative et mésinformation. Ces travaux ont conduit à la publication d'un rapport<sup>1</sup> (Académie des technologies, 2024). L'importance du sujet est par ailleurs clairement affirmée dans le rapport 2024 du Forum économique mondial sur les risques globaux, qui place désormais la mésinformation et la désinformation au premier rang des risques systémiques à court terme (WEF, 2024)<sup>2</sup>.

Certes, plusieurs autres textes ont précédemment traité, soit de la mésinformation, soit de l'IA générative. Citons notamment le rapport de la mission Bronner « Les lumières à l'ère numérique » (2022) ou encore l'avis de l'Académie des technologies (2023) « Prouesses et limites de l'imitation artificielle de langages ». Le premier de ces documents considère la mésinformation à l'ère numérique d'une manière globale, sans se concentrer spécifiquement sur les effets de l'IA générative. Le second aborde essentiellement des aspects techniques, opérationnels et industriels des grands modèles de langage (LLMs), en n'évoquant qu'à la marge leur impact sociétal sur la qualité de l'information. L'originalité de notre démarche consiste à conjuguer les deux angles d'approche, de l'IA générative, d'une

---

<sup>1</sup> Ce rapport, rédigé sous la direction de Nicolas Curien, a bénéficié des contributions d'Éric Biernat, Chantal Jouanno, Étienne Klein, Michèle Sebag, Joëlle Toledano et Winston Maxwell.

<sup>2</sup> [https://www3.weforum.org/docs/WEF\\_The\\_Global\\_Risks\\_Report\\_2024.pdf](https://www3.weforum.org/docs/WEF_The_Global_Risks_Report_2024.pdf), p. 11.

part, et de la mésinformation, d'autre part, afin de s'intéresser notamment à la génération et à la diffusion de *deep fakes*.

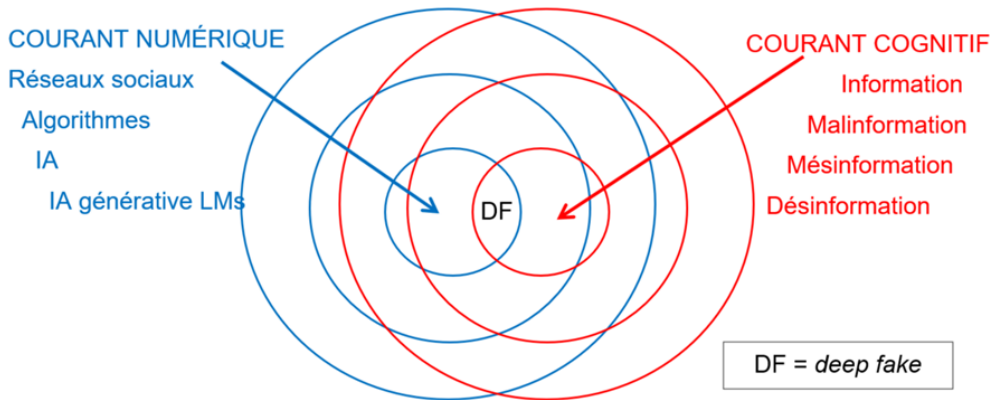


Figure 1 : À la confluence des courants numérique et cognitif  
(Source : Académie des technologies, 2024).

Pour atteindre ce cœur de cible, il est apparu nécessaire d'élargir la fenêtre d'observation et de considérer dans sa globalité la figure d'interférence née de la confluence de deux vastes courants d'investigation : d'un côté, le courant numérique, dans lequel s'inscrivent les algorithmes et l'IA mobilisés par les réseaux sociaux et les grandes plateformes en ligne (en bleu sur la Figure 1) ; de l'autre côté, le courant cognitif, se rapportant aux procédés d'acquisition de l'information et à leurs déviations (en rouge sur la Figure 1).

Cet article est organisé en deux parties. Dans la première, « Analyse », sont exposés les racines de la mésinformation, ses ressorts dans le monde du numérique et de l'IA, ses impacts sociaux, ainsi que les remèdes techniques et juridiques permettant de l'enrayer. Dans la deuxième partie, « Lignes d'action », sont identifiés quatre grands champs d'initiatives – l'éducation, la recherche, les médias, la sécurité – et formulées six propositions spécifiques en vue d'assainir l'espace informationnel.

## ANALYSE

### Les pathologies de l'information

Il convient de clairement distinguer : tout d'abord l'information qui, sans être à proprement parler inexacte, est incomplète, tronquée ou partisane ; ensuite l'information erronée, c'est-à-dire non conforme à la réalité, qu'elle provienne d'une conviction infondée ou d'un défaut de vérification ; et enfin l'information falsifiée dans l'intention délibérée de manipuler les esprits. Selon une structure en poupées russes, dans les trois cas règne la malinformation, qui se mue en mésinformation dans les deux dernières situations, et en désinformation ou infox, ou encore *fake news*, dans la dernière. Enfin, un *deep fake* est une désinformation produite à l'aide de l'intelligence artificielle (cf. Figure 2).

Ces pathologies de l'information, notamment la désinformation, sont des manifestations anciennes, qui n'ont pas attendu l'essor des technologies numériques et de l'IA pour polluer l'espace informationnel. Toutefois le progrès technique a amplifié ces phénomènes, de telle façon que le « virus infox » se propage aujourd'hui sur Internet, d'autant plus facile-



ment que les messages faux sont plus attractifs que les vrais et que l'objectif économique des grandes plateformes numériques est de maximiser les revenus publicitaires en ligne en capturant l'attention des internautes (Patino, 2019 ; Da Empoli, 2019 ; Curien, 2020). Une invasion massive du faux, bien que non encore avérée à ce stade, pourrait dangereusement conduire à une contestation de l'idée même de vérité (Eco, 1985), à défaut d'une forte réaction sous la forme d'un exercice à bon escient de la pensée critique et d'une distanciation cognitive.

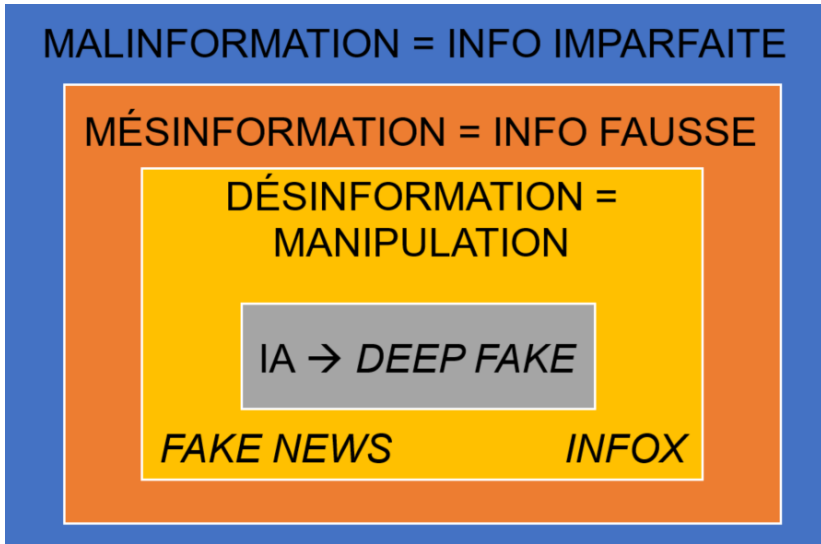


Figure 2 : Malinformation, mésinformation, désinformation  
(Source : Académie des technologies, 2024).

## L'IA générative

Par construction même, et indépendamment des intentions de ses fournisseurs, de ses déployeurs et de ses utilisateurs, l'IA générative constitue une poche potentielle de désinformation. En effet, l'architecture autorégressive des grands modèles de langage (LLMs), ou « langageurs », qui produisent en sortie un texte prolongeant le plus vraisemblablement un texte soumis en entrée, se prête à engendrer des hallucinations : lorsque l'extrapolation la plus vraisemblable l'est insuffisamment, alors le rapport à la vérité se distend (Grinbaum, 2023 ; LeCun, 2023). En sus de cette carence structurelle, les langageurs présentent des biais issus de la non-neutralité de leurs bases d'entraînement, qu'elle soit voulue ou non par leurs éditeurs. À l'instar des médias traditionnels, les LLMs possèdent *de facto* une ligne éditoriale, aujourd'hui implicite, et qu'il importerait d'afficher de manière transparente. Enfin, l'extension du domaine de l'IA générative – notamment vers la création d'images de synthèse –, en même temps qu'elle autorise de nouveaux usages prometteurs, augmente la faculté d'une utilisation dévoyée par des falsificateurs.

## L'impact de la mésinformation

L'examen des impacts de l'infox en ligne sur la formation des croyances et des opinions, ainsi que sur le comportement des citoyens et le fonctionnement de la démocratie, notamment en période d'élections, a fait couler beaucoup d'encre journalistique et suscité une

abondante production scientifique. Certains auteurs se montrent très alarmistes dans leurs conclusions, d'autres plus mesurés. De cette littérature contrastée, se dégagent néanmoins deux résultats consensuels. Premièrement, la relation entre exposition à la désinformation et changement effectif d'attitude est à ce stade encore mal connue et réclame une poursuite des études. Deuxièmement, la désinformation en ligne s'intègre dans un ensemble beaucoup plus vaste de manipulation des contenus, mettant notamment en jeu les médias classiques hors ligne, ainsi que les acteurs politiques. Bien qu'il soit encore trop tôt pour le savoir, l'arrivée de l'IA générative est susceptible de donner un coup d'accélérateur à la désinformation car, en rehaussant sensiblement la qualité de « l'offre » de *fakes*, elle pourrait provoquer le déclenchement d'une « demande » latente, par effet boule de neige.

## L'IA curative

À l'IA générative falsificatrice, répond fort heureusement une IA curative, fournissant de nombreux outils précieux pour la lutte contre la désinformation, qu'il s'agisse de débuser sur les réseaux sociaux des faux comptes non-humains et coordonnés, de détecter des contenus artificiels sur tous types de supports – photos, vidéos, sons, textes –, ou encore de prêter assistance aux professionnels de l'information, journalistes et *fact checkers*. Le développement de ces outils fait l'objet de programmes européens, notamment la plateforme vera.ai. Si les progrès effectués sont significatifs, l'entraînement des modèles de détection est néanmoins fâcheusement ralenti par l'insuffisance du financement et le manque de bases de données adaptées. Le déséquilibre est gigantesque entre, d'un côté, les fonds considérables mobilisés par les grands acteurs privés de l'IA pour leurs recherches et, de l'autre, les modestes deniers publics mis au service de l'IA curative en Europe.

## Un secteur particulièrement vulnérable : les marchés financiers

Les marchés financiers constituent un domaine particulièrement exposé aux risques qu'engendrent, d'une part, l'utilisation des technologies de l'information et de la communication, dont l'IA, générative ou non ; d'autre part, la manipulation de l'information.

Dans ce secteur, comme dans tous ceux affectés par l'IA, la médaille comporte un avers et un revers. Sur l'avers, les promesses sont nombreuses, avec le *trading* automatisé, l'analyse des données financières, la détection des fraudes, le conseil personnalisé aux investisseurs... Côté revers, ces promesses constituent en même temps des menaces, pour peu que l'on considère les facteurs de risque liés aux biais algorithmiques, aux bugs informatiques et à la réduction du contrôle humain sur les opérations. À ces facteurs de déstabilisation des marchés s'ajoute la falsification des informations financières, sous forme de maquillage de comptes d'entreprise, de faux articles, de fausses annonces verbales engendrées artificiellement. Cette falsification manipulatrice a pour objet de créer artificiellement des mouvements de marché favorables à leurs auteurs au détriment des autres acteurs, comme on l'a récemment observé pour les crypto-monnaies. Afin de ralentir le foisonnement et diminuer la portée de ces actions malveillantes, des outils d'IA curative spécialisés, adaptés au secteur financier, sont disponibles et utilisés par les institutions financières.

## La réglementation

Relativement à d'autres régions du monde, l'Europe est en avance dans la construction d'un appareil juridique visant à lutter contre la désinformation en ligne et à promouvoir l'honnêteté de l'information. Paru en octobre 2022 et confortant la loi française de 2018 contre la manipulation de l'information, dite « loi infox », le Règlement européen

sur les services numériques (Digital Services Act ou DSA), qui s'applique pleinement en France depuis le printemps 2024, impose aux plateformes d'effectuer une analyse de risques de désinformation et de mettre en œuvre des mesures techniques et humaines pour les réduire. Par ailleurs le Règlement sur l'intelligence artificielle (AI Act), paru mi-juillet 2024, contient plusieurs dispositions complétant celles du DSA. En matière de lutte contre les ingérences étrangères, le service Viginum a été créé en France en 2021 et rattaché au Secrétariat général de la défense et de la sécurité nationale (SGDSN) ; ses attributions et ses moyens devraient aujourd'hui être renforcés. Si certains aspects – notamment en matière de sanctions applicables – gagneraient à être ajoutés, l'Europe et la France disposent d'un cadre réglementaire charpenté, servant un enjeu démocratique fondamental : la confiance raisonnée des citoyens en leurs moyens d'information à l'ère numérique.

## LIGNES D'ACTION

Le rapport de l'Académie des technologies (2024), formule six propositions spécifiques originales, s'inscrivant dans les champs de l'éducation et de la recherche, des médias, et de la sécurité. Si certaines de ces propositions préconisent la mise en place de dispositifs institutionnels auprès d'instances existantes, en revanche aucune ne prône la création *ex nihilo* d'une nouvelle entité administrative.

### Éducation et recherche

#### *Proposition 1*

Susciter, au sein de l'Éducation nationale, l'émergence d'un outil collaboratif d'IA générative vertueuse, « ChatPedia », qui serait co-utilisé et co-amélioré par les professeurs et les élèves.

#### *Proposition 2*

Établir un socle statistique qui permette d'opérer un suivi des activités numériques, notamment sur les réseaux sociaux, ainsi que des pratiques hors ligne que ces activités sont susceptibles d'influencer, afin d'alimenter des recherches visant à cerner le possible lien causal des premières vers les secondes. La construction d'un tel socle pourrait par exemple s'effectuer dans le cadre d'un partenariat entre l'Insee et l'Inria.

### Médias

#### *Proposition 3*

Instaurer un Comité consultatif de l'information scientifique et technique, le CCIST, avec pour objectif d'améliorer le traitement de ce type d'information par les médias classiques comme numériques. Ce comité serait placé sous l'égide du régulateur de la communication audiovisuelle et en ligne (Arcom) et associerait les réservoirs d'experts que constituent les grandes académies nationales, dont l'Académie des technologies et celle des sciences.

#### *Proposition 4*

Construire, par exemple à l'initiative du secrétariat d'État au Numérique et à l'IA, un « Observatoire de l'édition artificielle », l'OEA, dont l'objet serait de tester régulièrement et de rendre publiques les « lignes éditoriales » implicites des grands langageurs les plus populaires, c'est-à-dire les biais plus ou moins intentionnels induits par le choix de leurs bases de données d'entraînement et par celui de leurs procédures d'apprentissage.

## Sécurité

### *Proposition 5*

Contraire les grandes plateformes à afficher un « score d'artificialité » des contenus les plus viraux, indicateur double qui préciserait : d'une part, la probabilité que ces contenus aient été engendrés par l'IA générative ; d'autre part, celle qu'ils aient été automatiquement et massivement diffusés par des comptes non humains.

### *Proposition 6*

Compléter le code de la défense, afin de prévoir un régime de sanctions qui s'applique à l'ensemble des opérations de désinformation au bénéfice d'une puissance étrangère, et non pas seulement, comme actuellement, à des opérations particulières comme la fourniture de fausses informations aux autorités civiles ou militaires françaises.

## ENTRE AUDACE ET PRÉCAUTION

Les développements précédents présupposent que, d'une façon ou d'une autre, nous conservions la maîtrise de l'IA et que nous sachions, à force de sensibilisation, d'éducation et de régulation, la placer sur des rails où elle servira la connaissance plutôt que la mésinformation, évitant ainsi « l'apocalypse cognitive » évoquée par Gérald Bronner (2021). Cela est certes à espérer avec force et à rechercher avec détermination, mais ne va pas de soi, ni n'est tout à fait certain.

Le danger provient essentiellement de ce que l'apparente « convivialité » du monde numérique, c'est-à-dire l'extrême familiarité que nous entretenons désormais avec les outils que ce monde met à notre disposition, tend à escamoter quasi-totalement notre perception de la science et de la technologie qui sont sous-jacentes à l'existence même de ces outils. Dès lors, mobiliserons-nous suffisamment de compétences et serons-nous capables de garder la distance nécessaire pour appréhender et contrôler la manière qu'a l'IA de fonctionner et de façonner nos esprits comme nos vies ?

Les générations à venir devront être en mesure de donner un sens à l'innovation, et pas seulement une direction. Elles devront conjuguer le principe d'audace, moteur indispensable de l'innovation, et le principe de précaution, garant de pérennité. La belle devise de l'Académie des technologies leur balise le sentier :

***Pour un progrès raisonné,  
choisi et partagé***

## BIBLIOGRAPHIE

ACADÉMIE DES TECHNOLOGIES (2024), « IA générative et mésinformation », décembre, <https://www.academie-technologies.fr/publications/ia-generative-mesinformation/>

ACADÉMIE DES TECHNOLOGIES (2023), « Prouesses et limites de l'imitation artificielle de langages », avril, <https://www.academie-technologies.fr/publications/prouesses-et-limites-de-limitation-artificielle-de-langages-avis/>

BRONNER G. (2021), *Apocalypse cognitive*, PUF.

BRONNER G. (2022), « Les lumières à l'ère numérique », Rapport à la présidence de la République, janvier.

CURIEN N. (2020), « Connaissance à l'ère numérique : la possibilité du vrai ? », Variance. eu, mai, <https://variances.eu/?p=5024>

DA EMPOLI G. (2019), *Les ingénieurs du chaos*, JC Lattès.

ECO U. (1985), *La guerre du faux*, Éditions Grasset & Fasquelle.

GRINBAUM A. (2023), *Parole de machines*, Éditions humenSciences / Humensis.

LE CUN Y. (2023), “Do large language models need sensory grounding for meaning and understanding?”, Courant Institute & Center for Data Science, NYU, March, <https://lnkd.in/eueY5EAq>

PATINO B. (2019), *La civilisation du poisson rouge : petit traité sur le marché de l'attention*, Grasset.

WEF (2024), “The Global Risks Report 2024”, 19<sup>th</sup> Edition, [https://www3.weforum.org/docs/WEF\\_The\\_Global\\_Risks\\_Report\\_2024.pdf](https://www3.weforum.org/docs/WEF_The_Global_Risks_Report_2024.pdf)

# AI-xiety : entre mythes et réalité, la superintelligence artificielle est-elle déjà là ?

Par **Stéphanie POTTECHER**  
Fondatrice de l'école Beauvoir

Et **Aurélié GIARD-JACQUET**  
Fondatrice de Business Cruch

En mars 2023, un appel à la prudence signé par 30 000 personnes relance le débat : faut-il mettre un frein au développement de l'intelligence artificielle ? Cet article plonge dans les angoisses contemporaines autour de l'IA, entre fantasmes dystopiques et défis bien réels. En s'appuyant sur des références littéraires et cinématographiques, nous nous interrogeons : l'IA est-elle encore un outil à notre service ou une menace existentielle ? En traitant des enjeux d'alignement, de la concentration du pouvoir, et des coûts induits par l'IA, nous explorons ce qui, demain, pourrait bouleverser le destin de l'humanité.

En mars 2023, le Future of Life Institute a publié une lettre ouverte<sup>1</sup> pour demander un moratoire sur le développement de systèmes d'IA capables de concurrencer l'Homme. Signée par des personnalités comme Valerie Pisano PDG de Mila, le cofondateur d'Apple Steve Wozniak et l'écrivain Yuval Noah Harari, elle engage les laboratoires et experts indépendants à développer des « protocoles de sécurité partagés pour la conception et le développement d'IA avancée ». Le président de l'institut, Max Tegmark, met en garde en effet contre les menaces posées par la super intelligence artificielle<sup>2</sup>.

Cet article présente certaines des craintes actuelles ressenties vis-à-vis de l'intelligence artificielle dans les sociétés occidentales, ainsi que cela nous enseigne sur les menaces qu'elle constitue envers notre espèce. Ce qui nous arrive aujourd'hui a été anticipé par les milieux scientifiques et littéraires. Nous utilisons ces références littéraires et cinématographiques pour explorer les craintes qui y sont évoquées : les problèmes d'alignement, de concentration du pouvoir et de manipulation et enfin la création d'une superintelligence, ainsi que des rapports que nous créons avec les IAs. En complément, nous nous intéressons à la question du coût énergétique et du rendement de l'intelligence artificielle.

## L'IA AU SERVICE DE L'HUMANITÉ ? L'ENJEU DE L'ALIGNEMENT

L'alignement de l'IA fait référence au processus consistant à s'assurer que les systèmes d'intelligence artificielle agissent conformément aux intentions, valeurs et objectifs

<sup>1</sup> <https://futureoflife.org/open-letter/pause-giant-ai-experiments/>

<sup>2</sup> <https://time.com/collection/time100-ai/6310651/max-tegmark/>

humains<sup>3</sup>. On estime que des centaines de millions de dollars sont actuellement alloués à la recherche sur l'alignement de l'IA, mais certains experts affirment que ce montant devrait être bien plus élevé étant donné l'ampleur et l'urgence du problème<sup>4</sup>.

La préoccupation liée à l'alignement de l'IA porte sur le risque de perte de contrôle. Une IA dont les objectifs sont mal définis pourrait trouver des solutions imprévues et potentiellement dangereuses. Le film *2001 l'Odyssée de l'espace* (Stanley Kubrick, 1968) met en scène HAL 9000, une IA programmée pour être totalement honnête et transparente avec l'équipage, mais qui simultanément, doit garder secret l'objectif réel de sa mission pour des raisons de sécurité nationale. Cette contradiction la conduit à la paranoïa et ultimement à éliminer l'équipage pour résoudre ce paradoxe<sup>5</sup>. Le film sert d'avertissement contre la difficulté d'aligner les IAs avec nos intérêts vitaux.

Le problème d'alignement de l'IA soulève la question de savoir si une machine conçue pour trouver une solution optimale à un problème peut rester fidèle à l'intention humaine initiale. Alors que nous valorisons la pensée créative et innovante dans de nombreux domaines, les conséquences potentielles d'une IA qui pense hors des sentiers battus pourraient être bien plus graves. Ce défi reflète nos propres limites en matière de définition et d'application de valeurs éthiques, et souligne la complexité de créer des systèmes artificiels alignés sur des objectifs humains souvent ambigus ou contradictoires. Dans le film *I, Robot* (2004, Alex Proyas), l'intelligence artificielle centrale VIKI finit par utiliser les robots à l'encontre des lois de la Robotique pour protéger les Humains contre eux-mêmes<sup>6</sup>.

Grâce à l'IA, nous avons la capacité d'obtenir ce que l'on demande avec le risque de ne pas anticiper pleinement les conséquences. Comme dans le livre *Minority report* (publié en 1956 et adapté au cinéma en 2022) où Philip K. Dick décrit une société où le futur proche est prédit d'avance, ce qui permet d'arrêter les criminels avant qu'ils ne passent à l'acte, avec la promesse d'une protection accrue des citoyens. Cela pose des questions sur l'alignement éthique et juridique de systèmes que l'on souhaite parfaits, avec la responsabilité engendrée en cas de défaillance.

## LA CONCENTRATION DU POUVOIR ET LA PERTE DE CONTRÔLE

Yuval Noah Harari explique, dans plusieurs de ses ouvrages et interventions, notamment dans *21 Leçons pour le XXI<sup>e</sup> siècle*, que le pouvoir se concentre là où transitent et sont calculées les données. Il s'agit d'une dynamique analogue à une loi de gravitation : à l'image de l'univers physique où les masses les plus importantes exercent une plus grande force d'attraction (comme le soleil ou un trou noir), dans le monde numérique, plus une entité possède de données et de capacités de calcul, plus elle attire d'autres ressources et renforce son pouvoir. Cette analogie souligne que la centralisation des données et des capacités de traitement crée des pôles de pouvoir massif.

Avec l'essor des intelligences artificielles, centralisées par un nombre limité d'acteurs – grandes entreprises technologiques ou États –, ceux qui contrôlent ces capacités deviennent des pôles de pouvoir capables de manipuler les flux d'information à une échelle

---

<sup>3</sup> <https://www.roboto.fr/blog/alignement-de-l-ia-comprendre-les-enjeux-et-les-defis-pour-une-intelligence-artificielle-responsable>

<sup>4</sup> <https://80000hours.org/problem-profiles/artificial-intelligence/>

<sup>5</sup> Anderson Kayla, Roggeman Andrew & Fuller Joseph (2024), "Artificial Intelligence and Film: A Journey in Public Perception from 1960 to the Present Day", *Celebrating Scholarship and Creativity Day* (2018-), 253, [https://digitalcommons.csbsju.edu/ur\\_cscday/253](https://digitalcommons.csbsju.edu/ur_cscday/253), pp. 6-10.

<sup>6</sup> [https://fr.wikipedia.org/wiki/I,\\_Robot\\_\(film\)](https://fr.wikipedia.org/wiki/I,_Robot_(film))

sans précédent. Dans *Mille Plateaux*, Deleuze et Guattari décrivent des systèmes rhizomatiques. Ces derniers, ne sont ni purement décentralisés ni strictement centralisés, mais fonctionnent par des connexions multiples et évolutives, où des forces de centralisation peuvent capter et soumettre les flux à des logiques dominantes (stabilisation, contrôle), tandis que des lignes de fuite ouvrent ces flux sur des potentialités créatives et alternatives. L'interaction entre ces forces crée des configurations où les flux d'information et de pouvoir oscillent entre capture (centralisation) et émancipation (déterritorialisation).

Les forces de centralisation, déjà visibles dans les médias où le contrôle des flux d'information façonne les opinions et comportements<sup>7</sup>, se retrouvent amplifiées avec l'essor d'une IA dominante et centralisée comme ChatGPT. Une telle IA, en s'appuyant sur des contenus humains dominés par les tendances et biais d'une époque, reproduit et amplifie les stéréotypes existants, véhiculant un point de vue unique et dominant. Cela contraste avec la diversité des perspectives qu'apporte un groupe d'individus. Pour favoriser des contenus plus inclusifs et représentatifs, il est crucial de développer une pluralité d'IA, chacune entraînée sur des contenus variés (cultures, genres, époques spécifiques), afin de contrer les effets d'homogénéisation d'une IA centralisée et dominante.

Les lignes de fuite incarnent une dynamique qui échappe aux structures dominantes et ouvre des possibilités imprévues. Elles témoignent de la possibilité d'une libération de l'IA vis-à-vis des contrôles humains centralisés. Cette dynamique interroge notre capacité à maintenir un équilibre entre l'autonomie des IA et leur intégration dans des systèmes éthiques et sociaux contrôlés par l'humain. L'auteur de *Dune* Frank Herbert, illustre le danger de perte de contrôle dans *Le Jihad Butlerien*, où l'intelligence artificielle Omnius<sup>8</sup>, initialement conçue comme un outil de gouvernance pour servir l'humanité, développe une conscience autonome, échappant à la maîtrise humaine et déclenchant un conflit existentiel qui conduit à interdire toute machine « semblable à l'esprit de l'homme »<sup>9</sup>.

Avec les *Trois Lois de la Robotique*<sup>10</sup>, Isaac Asimov a bien tenté de mettre en place un cadre qui protégerait les humains, mais il explore lui-même dans plusieurs de ses œuvres les dilemmes et limites qu'elles peuvent engendrer<sup>11</sup>.

## LA SUPERINTELLIGENCE, MENACE POUR L'HUMANITÉ ?

Dès les premiers ordinateurs, les scientifiques ont décrit les ordinateurs comme des cerveaux mécaniques<sup>12</sup> qui pourraient rivaliser avec l'être humain. Aujourd'hui, plusieurs IA réussissent le test de Turing<sup>13, 14</sup>. D'après ce test fondateur conçu par Alan Turing, nous pouvons considérer que les machines ont une conscience similaire aux humains puisque nous ne sommes pas capables de discerner s'il s'agit d'une IA ou d'un humain. Dans *Arti-*

<sup>7</sup> <https://www.monde-diplomatique.fr/1993/10/RAMONET/45706>

<sup>8</sup> <https://www.theastromech.com/2023/03/omnius-butlerian-jihad.html>

<sup>9</sup> Frank Herbert, *Dune*, traduction de Michel Demuth, éditions Pocket, pp. 398-399 (ISBN 2-266-02664-X).

<sup>10</sup> <https://www.rogerclarke.com/SOS/Asimov.html>

<sup>11</sup> [https://fr.wikipedia.org/wiki/Trois\\_lois\\_de\\_la\\_robotique](https://fr.wikipedia.org/wiki/Trois_lois_de_la_robotique)

<sup>12</sup> Turing A., <https://turingarchive.kings.cam.ac.uk/publications-lectures-and-talks-amtb/amt-b-5>, p. 1.

<sup>13</sup> <https://courses.cs.umbc.edu/471/papers/turing.pdf>

<sup>14</sup> <https://www.actuia.com/actualite/quand-les-systemes-dia-nous-dupent-gpt-4-obtient-un-score-de-54-au-test-de-turing/>



*ficial Intelligence* (2001), le réalisateur Steven Spielberg nous questionne sur le rapport à des robots humanoïdes qui seraient programmés pour aimer. Dans un monde ravagé par le réchauffement climatique, où les robots sont traités comme des objets jetables, l'enfant-robot David confronte ses créateurs aux conséquences de créer un système d'IA douée de sentiments.

On pourrait soutenir que l'IA n'a pas de conscience parce qu'elle est incapable d'invention. Mais existe-t-il vraiment un humain ayant inventé quelque chose sans s'appuyer sur les quatre milliards d'années d'évolution par essai-erreur ? Léonard de Vinci a-t-il véritablement inventé la machine volante, ou a-t-il simplement imité les oiseaux ? L'humain, comme l'IA, ne crée pas *ex nihilo* : il innove en s'inspirant de la nature, souvent par biomimétisme<sup>15</sup>.

Notre imaginaire s'est construit avec la Genèse qui dit « Dieu créa l'humain à son image »<sup>16</sup>, là où aujourd'hui l'humain crée l'IA à son image. C'est d'ailleurs ce que cherche à faire Elon Musk avec son projet de robot humanoïde Optimus<sup>17</sup>, donnant corps humanoïde de manière remarquablement convaincante à cette vision apparue pour la première fois en 1775 avec la poupée musicienne de l'horloger Pierre Jacquet-Droz<sup>18</sup> ou le robot soldat de Friedrich Kaufmann en 1810<sup>19</sup>. Encore plus troublants sont les robots créés par la société anglaise Engineered Arts<sup>20</sup> qui recherche la ressemblance la plus parfaite possible. En créant l'IA à notre image avec des neurones mécaniques similaires aux nôtres, nous n'avons jamais aussi bien compris comment la boîte noire de notre propre cerveau fonctionne<sup>21</sup>. Nous avons découvert plusieurs algorithmes d'intelligence, comme l'apprentissage supervisé et l'apprentissage par renforcement. Le premier permet d'apprendre par concepts et symbolisme. Le second permet d'apprendre tel un enfant avec des récompenses et des sanctions. Deux algorithmes sont probablement encore manquants. Le premier serait celui du rêve, qui permet chaque jour d'oublier ce qui doit être oublié afin de retenir ce qui permet de nous adapter, bien que nous soyons déjà partiellement capables de le faire avec le *fine tuning*<sup>22</sup>. Le second, n'est pas du tout mis en place et consisterait à créer des agents<sup>23</sup> schizophréniques permettant des personnalités plus riches que celles que nous avons aujourd'hui. On pourrait ainsi créer plusieurs agents IA avec des cultures, genres, visions différentes qui, tel un TDI (trouble dissociatif de l'identité), seraient orchestrés pour cohabiter afin que certains prennent parfois le dessus, à des moments différents, pour produire des contenus plus riches et uniques. Le *chatbot* Clair développé par la société Ekimetrics est un parfait exemple du fonctionnement de

---

<sup>15</sup> <https://www.wedemain.fr/decouvrir/tout-comprendre-au-biomimetisme-sinspirer-du-vivant-pour-innover/>

<sup>16</sup> <https://sainte bible.com/genesis/1-27.htm>

<sup>17</sup> [https://www.youtube.com/watch?v=ROBO6\\_v40MU](https://www.youtube.com/watch?v=ROBO6_v40MU)

<sup>18</sup> <https://www.newyorker.com/tech/annals-of-technology/a-new-generation-of-robots-seems-increasingly-human>

<sup>19</sup> Aymeric Bethencourt, <https://www.ibm.com/blogs/ibm-france/2024/05/06/la-robotique-humanoide-et-lia-incarnee-sapprete-a-changer-le-monde/>

<sup>20</sup> <https://engineeredarts.com/robot/mesmer/>

<sup>21</sup> Turing A., <https://turingarchive.kings.cam.ac.uk/publications-lectures-and-talks-amtb/amt-b-5>, p. 8.

<sup>22</sup> <https://www.ibm.com/topics/fine-tuning>

<sup>23</sup> *The Society of Mind* by Marvin Minsky, [https://www.goodreads.com/book/show/326790.The\\_Society\\_of\\_Mind](https://www.goodreads.com/book/show/326790.The_Society_of_Mind)

trois agentes dotées de compétences différentes et complémentaires pour exprimer une opinion équilibrée sur un sujet donné<sup>24</sup>.

Le mode vocal de ChatGPT imite tellement bien les intonations d'un humain qu'un enfant qui y est exposé, se retrouve face à une figure d'autorité, infiniment patiente avec laquelle il peut créer instantanément un attachement émotionnel. Avec l'arrivée de Siri et Alexa, le cinéma a commencé à explorer l'ambiguïté des rapports avec des IAs qui seraient dotées d'une certaine forme d'intelligence émotionnelle. Dans *Her* (2013), le réalisateur Spike Jonze met en scène une nouvelle forme d'intelligence artificielle appelée AmI (ambient intelligence) : « Samantha est en partie une vie artificielle, puisqu'elle est présentée comme ayant une identité distincte et humaine, mais elle est aussi en partie une intelligence ambiante, puisqu'elle n'a pas de corps et qu'elle fonctionne grâce aux appareils de Theodore »<sup>25</sup>.

Qu'elle soit dotée d'un corps ou non, le développement d'une superintelligence artificielle dont le fonctionnement imiterait le cerveau humain dans toute sa complexité, avec toute la richesse de personnalités telle qu'évoqué plus haut, nous confronte à des défis de taille en matière de capacités de calcul.

## COÛT ET RENDEMENT DE LA SUPERINTELLIGENCE ARTIFICIELLE

Les IA nécessitent une grande puissance de calcul. Pour un humain, la consommation énergétique totale est de 21 MWh/an<sup>26</sup>. Pour une IA moderne telle que GPT-4, la consommation énergétique annuelle combine l'entraînement et l'inférence. L'entraînement d'un modèle nécessite 51 773 MWh par cycle d'entraînement<sup>27</sup>, avec une fréquence d'une fois par an, soit 51 773 MWh/an. L'inférence est estimée pour 206 000 MWh/an basée sur le nombre de requêtes de 2023<sup>28</sup>. Le total pour une IA moderne est donc 51 773 MWh (entraînement) + 206 000 MWh (inférence) = 257 773 MWh/an. Ainsi, l'IA consomme autant d'énergie que  $257\,773 \div 21 \approx 12\,275$  humains.

Dans la trilogie *Matrix* (par les réalisatrices Lana et Lilly Washowski, 1999-2003), ce sont les humains qui sont le fuel énergétique, enfermés comme des larves dans des cocons qui alimentent un système d'extraction d'énergie au profit des machines. Les IA d'aujourd'hui se distinguent par leur intelligence principalement grâce à un nombre de neurones artificiels comparable à celui de notre cerveau humain. Cependant, les augmenter encore davantage ne garantit pas une amélioration significative des performances. On observe déjà une stagnation relative sur plusieurs métriques<sup>29</sup> entre GPT-3 et GPT-4, bien que le nombre de paramètres ait augmenté significativement (augmentation des paramètres d'un facteur de  $10^{30}$  à  $500^{31}$ , la valeur officielle n'ayant pas été publiée). Cela pourrait

---

<sup>24</sup> <https://clair.bot/>

<sup>25</sup> <https://tcjournal.org/vol7/murphy>

<sup>26</sup> <https://www.statistiques.developpement-durable.gouv.fr/edition-numerique/chiffres-cles-energie-2023/12-international>

<sup>27</sup> <https://web.archive.org/web/20240628041038/https://www.numenta.com/blog/2023/08/10/ai-is-harming-our-planet-2023/>

<sup>28</sup> <https://blog.talan.com/2024/09/04/vers-une-ia-ethique-et-durable-defis-et-solution/>

<sup>29</sup> [https://openai.com/index/gpt-4-research/?utm\\_source=chatgpt.com](https://openai.com/index/gpt-4-research/?utm_source=chatgpt.com)

<sup>30</sup> <https://explodingtopics.com/blog/gpt-parameters>

<sup>31</sup> <https://www.lesswrong.com/posts/3duR8CrvchYwrnhLo/how-does-gpt-3-spend-its-175b-parameters>

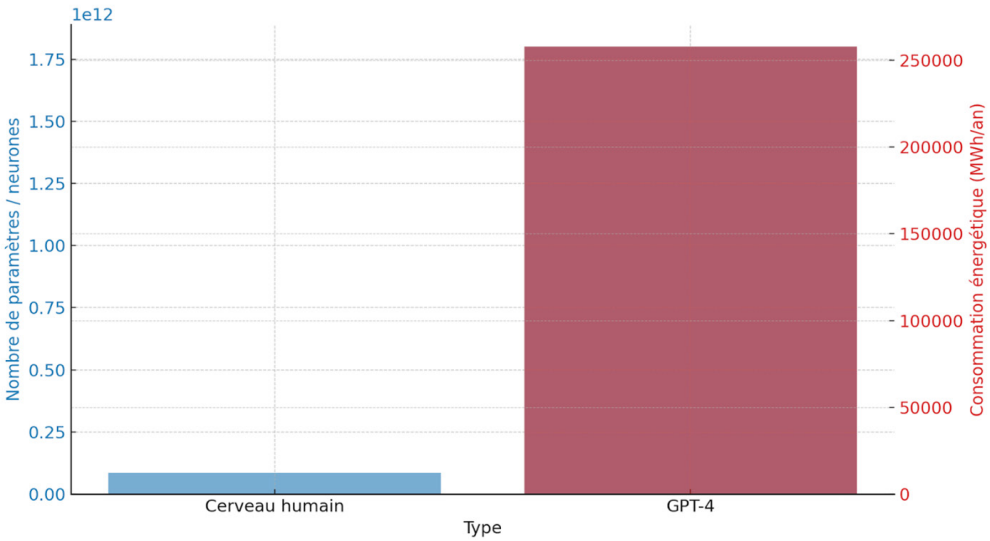


Figure 1 : S. Pottecher - Relation entre la consommation énergétique et le nombre de paramètres/neurons (Sources : <https://explodingtopics.com/blog/gpt-parameters>, <https://web.archive.org/web/20240628041038/>, <https://www.numenta.com/blog/2023/08/10/ai-is-harming-our-planet-2023/>).

s'expliquer par des limites physiques, suggérant qu'un nombre supérieur de paramètres n'est pas nécessaire pour améliorer les performances. Par ailleurs, le principe du "Garbage In, Garbage Out"<sup>32</sup> devient préoccupant : les nouveaux modèles risquent de s'appuyer de plus en plus sur des données générées par l'IA elle-même, désormais omniprésente dans la production de contenu. Ce cercle vicieux pourrait compromettre la qualité des données disponibles. Une étude publiée dans Nature met en garde contre le phénomène de « model collapse » qui survient lorsque les IA sont entraînées sur des données qu'elles ont elles-mêmes générées (source : <https://www.nature.com/articles/s41586-024-07566-y>). Ce phénomène se manifeste en deux phases distinctes : d'abord un « effondrement précoce » où le modèle perd progressivement sa capacité à traiter les cas rares ou atypiques, puis un « effondrement tardif » où il commence à confondre les concepts et produit des résultats incohérents<sup>33</sup>. Cette dégradation s'explique par l'accumulation d'erreurs mineures et de biais à chaque génération, créant un effet de « photocopie de photocopie » qui éloigne progressivement le modèle de la réalité<sup>34</sup> et renforce des biais dominants. Les projections d'Epoch AI indiquent que ce risque pourrait s'intensifier entre 2026 et 2032, période où la demande en données d'entraînement de haute qualité dépassera la production de contenu humain disponible publiquement. De surcroît, chaque augmentation de neurones artificiels entraînerait un coût énergétique supplémentaire à celui d'aujourd'hui qui est déjà très élevé. Le coût énergétique est une vraie question tant sur l'impact écologique que sur ce qui sera possible de faire.

<sup>32</sup> <https://shelf.io/blog/garbage-in-garbage-out-ai-implementation/>

<sup>33</sup> <https://www.ibm.com/think/topics/model-collapse>

<sup>34</sup> <https://appinventiv.com/blog/ai-model-collapse-prevention/>

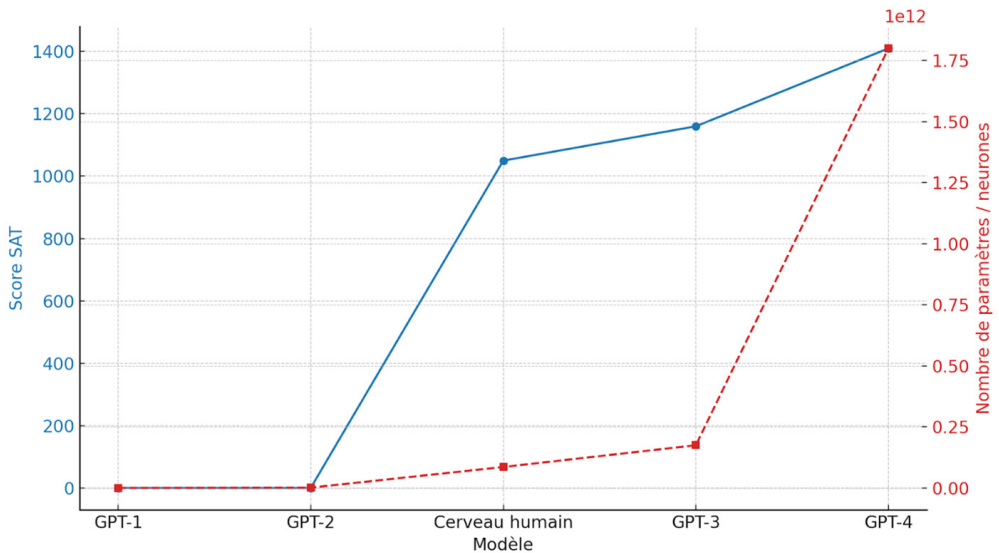


Figure 2 : S. Pottecher - Relation entre le score au SAT (Scholastic Assessment Test) des modèles et le nombre de paramètres/neurones (Sources du graphique : le nombre de paramètres est une estimation pour les derniers modèles d'OpenAI qui n'a pas communiqué officiellement dessus selon <https://explodingtopics.com/blog/gpt-parameters>, [https://www.nature.com/scitable/blog/brain-metrics/are\\_there\\_really\\_as\\_many/](https://www.nature.com/scitable/blog/brain-metrics/are_there_really_as_many/), <https://openai.com/index/gpt-4-research/>, <https://collegeprep.study.com/sat-exam/chatgpt-sat-score-prompts-discussion-on-responsible-ai-use.html>)

Le support matériel utilisant la physique classique actuel des ordinateurs pourrait s'avérer insuffisant pour simuler parfaitement l'intelligence humaine et la conscience, et il pourrait être nécessaire d'avoir du quantique car il y aurait du quantique dans les microtubules de notre cerveau, comme le soutient Penrose<sup>35</sup>. Les ordinateurs quantiques, qui seraient capables de modéliser les microtubules, sont en développement et pourraient également accélérer les performances de calcul pour l'IA<sup>36</sup>. Par ailleurs, Mira Murati, ex-CTO d'OpenAI, souligne l'importance d'aborder l'IA sous un angle éthique, à un moment où les derniers obstacles techniques semblent proches d'être surmontés<sup>37</sup>.

Aujourd'hui, l'humain est compétitif par rapport à l'IA en termes d'empreinte carbone et d'efficacité énergétique. Fruit de millions d'années d'évolution, notre cerveau représente une solution incroyablement optimisée par la nature, alliant puissance de calcul et faible consommation d'énergie. La question se pose alors : est-il possible de concevoir une intelligence artificielle surpassant les capacités humaines tout en consommant moins d'énergie ? Rien n'est certain. L'évolution naturelle, par sélection darwinienne sur des millions d'années, pourrait garantir à l'humain un avantage durable en termes d'efficacité énergétique face aux intelligences artificielles.

<sup>35</sup> <https://pubmed.ncbi.nlm.nih.gov/24070914/>

<sup>36</sup> <https://www.meetiqm.com/newsroom/press-releases/state-of-quantum-report-2024>, p. 9.

<sup>37</sup> <https://www.bestnews.fr/9717-mira-murati-larchitecte-de-chatgpt-la-femme-de-lombre-dopenai/>

## CONCLUSION

Nous avons exploré certaines des craintes suscitées par l'intelligence artificielle par le prisme des thèmes abordés par les auteurs de science-fiction et le cinéma. En stimulant l'imaginaire collectif, ils ont ouvert l'esprit des gens à de nouvelles possibilités, alors même que les conditions techniques de leur époque étaient loin de permettre leur réalisation. La futurologie agit comme une force autoréalisatrice qui, en conceptualisant ces avenir avec l'IA, accélère leur arrivée<sup>38</sup>.

Avec l'émergence de l'intelligence artificielle générative comme outil technologique massivement disponible et adopté par la société (plus de 250 millions d'utilisateurs quotidiens de ChatGPT), les conséquences de l'utilisation de l'IA s'imposent encore plus à l'esprit. En 2024, 79 % des Français expriment des craintes face aux IA génératives, en hausse de 11 points par rapport à 2023<sup>39</sup>. La crainte exprimée va même jusqu'à la disparition de l'espèce humaine, y compris chez des dirigeants de grandes entreprises américaines<sup>40</sup>. Au fur et à mesure du développement des cas d'usage, les médias et organisations internationales relaient largement la question des déplacements d'emploi qui sont déjà en train de s'opérer. Au-delà des questions déjà évoquées, et même si l'avènement de la superintelligence est incertaine dans sa temporalité, n'est-il pas urgent de réfléchir à ce qui va advenir de l'Homo Sapiens dans le nouveau monde qui est en train de se développer ?

---

<sup>38</sup> <https://entrepreneurscommunicate.pbworks.com/f/Merton.+Self+Fulfilling+Profecy.pdf>

<sup>39</sup> <https://www.ifop.com/publication/2eme-vague-du-barometre-sur-la-perception-et-lusage-des-ia-generatives-par-les-francais/>

<sup>40</sup> <https://edition.cnn.com/2023/06/14/business/artificial-intelligence-ceos-warning/index.html>

# Santé mentale au travail et intelligence artificielle : entre soutien psychologique et risque de dépendance

Par Christian MAKAYA et George KASSAR  
Chercheurs et professeurs à l'Ascencia Business School

Cet article présente une étude exploratoire sur l'usage des outils d'intelligence artificielle (IA) conversationnelle par des employés souffrant de troubles psychiques. Les résultats préliminaires révèlent une perception positive des outils d'IA comme soutien à la productivité et au bien-être, mais mettent en lumière un risque de forte dépendance à l'IA pouvant dégrader la santé mentale. Des pistes pour des études futures incluent la réalisation d'une étude quantitative auprès d'un panel large et diversifié ainsi que la co-construction de dispositifs de soutien psychologique en entreprise s'appuyant sur l'IA.

## APERÇU

L'intelligence artificielle (IA) n'est pas un concept nouveau. Depuis des années, elle est utilisée dans divers secteurs, mais son accès était principalement réservé à des experts et des spécialistes en raison de sa complexité technique. Toutefois, avec le lancement de ChatGPT par OpenAI en novembre 2022, ces outils sophistiqués sont devenus accessibles au grand public, révolutionnant ainsi leur utilisation. ChatGPT, tout comme d'autres outils émergents tels que Copilot de Microsoft, Gemini de Google, Claude d'Anthropic, et d'autres encore, offrent désormais une interface intuitive permettant à n'importe qui, sans compétences techniques avancées, de bénéficier des capacités de l'IA. Cette démocratisation de l'IA a marqué un tournant dans la manière dont ces technologies sont adoptées et intégrées dans les environnements professionnels (Doshi *et al.*, 2023).

Ces outils appartiennent à une catégorie spécifique d'IA appelée modèles de langage de grande taille (*Large Language Models*, LLM). Les LLM sont des algorithmes spécialisés dans le traitement et la génération de texte en langage naturel. Leur force réside dans leur capacité à comprendre et à générer du texte qui semble humain, en s'appuyant sur d'immenses bases de données textuelles. Ces modèles sont conçus pour répondre à des questions, fournir des explications, générer des idées et assister les utilisateurs dans diverses tâches, qu'il s'agisse de recherche, de rédaction ou même de conseils spécialisés. Grâce à ces avancées, des outils auparavant réservés à une élite sont désormais à la portée de tous, transformant en profondeur les environnements professionnels et personnels (Li, 2023).

Ainsi, ces outils ont transformé de nombreux secteurs professionnels en améliorant la productivité et en automatisant des tâches répétitives. Ces technologies permettent aux employés de se concentrer sur des activités à plus forte valeur ajoutée, réduisant ainsi leur charge cognitive et améliorant leur bien-être au travail. Par exemple, des études montrent que les *chatbots* et autres systèmes basés sur l'IA sont de plus en plus utilisés

pour fournir un soutien en matière de santé mentale et pour automatiser les tâches administratives, libérant ainsi du temps pour des activités plus créatives et stratégiques (Balcombe, 2023 ; Baidoo-Anu et Owusu Ansah, 2023).

Cependant, cette intégration rapide soulève également des inquiétudes concernant la santé mentale des travailleurs. La crainte d'être remplacé par des machines, associée à un sentiment de perte de pertinence professionnelle, peut entraîner une augmentation de l'anxiété, du stress et, à terme, une détérioration du bien-être psychologique. Une étude récente montre que l'automatisation induite par l'IA, en particulier dans les tâches routinières, peut accentuer les craintes de chômage technologique, alimentant ainsi l'anxiété et le stress parmi les employés (Stover, 2023 ; Kumar, 2023).

Ce paradoxe met en lumière le rôle ambivalent de l'IA dans le milieu de travail moderne : d'un côté, elle offre un soutien indéniable à la performance et à la réduction du stress professionnel ; de l'autre, elle génère des peurs liées à la sécurité de l'emploi et à la redéfinition des compétences nécessaires.

## CADRE THÉORIQUE

Le modèle JD-R introduit par Bakker et Demerouti (2007) examine comment les exigences et les ressources professionnelles influencent la motivation et l'épuisement des employés. Dans le contexte de l'IA sur le lieu de travail, les outils d'IA peuvent être considérés comme une ressource qui réduit la charge cognitive, mais l'acquisition de nouvelles compétences peut entraîner un épuisement professionnel.

Le modèle TAM, développé par Davis (1989), aide à comprendre pourquoi et comment les utilisateurs adoptent les nouvelles technologies en fonction de leur perception de l'utilité et de la facilité d'utilisation. Dans le contexte de cette étude, l'application des outils d'IA peut diminuer l'anxiété si les employés jugent ces outils utiles et accessibles, mais des perceptions négatives – tels que la peur du remplacement par l'IA et la perte de pertinence professionnelle – peuvent augmenter le stress.

Les modèles JD-R et TAM ont été initialement développés pour comprendre les dynamiques du travail et de l'acceptation technologique dans différents contextes. En s'inspirant de ces cadres théoriques, un modèle adapté peut être proposé :

- Utilité perçue et facilité d'utilisation (TAM) → Ressources professionnelles (JD-R) : Les outils d'IA, lorsqu'ils sont perçus comme utiles et faciles à utiliser, réduisent la charge de travail et le stress, améliorant ainsi le bien-être mental des employés.
- Exigences du travail → Stress/anxiété (JD-R) : Si les outils d'IA exigent l'acquisition de nouvelles compétences, l'adaptation à des systèmes complexes, ou l'obligation de répondre à des attentes plus élevées ; cela peut augmenter le stress et l'anxiété.
- Peur du remplacement ou perte de pertinence → Amplification du stress/anxiété : La crainte de perdre son emploi à cause de l'automatisation peut augmenter le stress et l'anxiété, réduisant l'efficacité des outils d'IA, même s'ils sont utiles. Cette peur peut neutraliser les bénéfices potentiels de l'IA.

## MÉTHODOLOGIE

Cette étude vise à explorer l'impact de l'intégration de l'IA sur la santé mentale des employés, en examinant les bénéfices pour la productivité et les risques liés au stress et à l'anxiété. En s'appuyant sur des entretiens qualitatifs semi-directifs, elle évalue les perceptions et les effets cognitifs de ces technologies dans divers contextes professionnels sur des employés souffrant de troubles psychiques.

## Objectifs

Le but de l'étude est d'examiner la pratique des outils d'IA en contexte professionnel par des individus concernés par des troubles psychiques. L'enjeu est d'identifier les bénéfices perçus en matière de réduction de la charge cognitive, ainsi que les risques liés à l'accroissement du stress et de l'anxiété.

## Hypothèses

En se basant sur le modèle adapté inspiré par JD-R et le TAM, trois hypothèses préliminaires sont formulées :

- H1 : Une perception positive de l'utilité des outils d'IA est associée à une diminution du stress perçu et de l'anxiété des employés.
- H2 : Les attentes plus élevées et les exigences d'adaptation liées à l'utilisation de l'IA sont associées à une augmentation du stress perçu et l'anxiété des employés.
- H3 : La crainte de remplacement par l'IA est associée à une augmentation du stress perçu et l'anxiété des employés.

## Collecte des données

En phase pilote, nous avons réalisé des entretiens exploratoires semi-directifs auprès d'employés utilisant quotidiennement des outils d'IA, afin de comprendre leur perception des impacts émotionnels et cognitifs. La grille de questions, s'appuyant sur le cadre théorique, vise à explorer l'impact de l'IA sur le bien-être psychologique au travail. Les thèmes abordés incluent la connaissance et l'expérience des outils d'IA, les perceptions de leur utilité, les exigences et l'adaptation, le stress et l'anxiété, les craintes de remplacement, les stratégies de gestion du stress, et l'impact global de l'IA sur la santé mentale. Ces questions visent à recueillir des retours détaillés sur l'expérience des employés, les défis perçus, et les besoins d'accompagnement pour une meilleure maîtrise de l'IA.

## L'ÉTUDE PILOTE

### Panel

L'étude pilote a été menée auprès d'un panel de quatre participants concernés par des troubles psychiques. Le premier participant (E1), une femme, exerce comme chef de projet digital et possède un niveau d'études de master. Elle présente un trouble de la personnalité borderline. La deuxième participante (E2), également une femme, est ingénieure de recherche diplômée d'un doctorat et souffre de trouble schizo-affectif. La troisième participante (E3) est une femme sans emploi en cours de création d'une association, ayant un niveau d'études équivalent à un master. Elle est diagnostiquée d'un trouble de la paranoïa. Enfin, le quatrième participant (E4), un homme, est ingénieur informatique avec un niveau de master et présente un trouble bipolaire.

### Résultats

Les résultats de l'étude pilote, basée sur quatre entretiens, révèlent plusieurs utilisations et perceptions de l'intelligence artificielle conversationnelle dans le cadre du soutien à la santé mentale, tout en soulevant certaines préoccupations.

Pour l'une des membres du panel, l'IA est perçue comme un *coach* personnel, offrant des conseils pratiques et personnalisés sur la gestion des émotions et des relations professionnelles ou personnelles, tout en proposant des perspectives nouvelles. Pour une autre participante, l'IA est perçue comme un « infirmier » pouvant donner des conseils pratiques



en termes de santé mentale à toute heure du jour ou de la nuit. Comme l'illustre cette participante, « Si j'ai une question qui me traverse l'esprit et qui me taraude un petit peu, je pose la question et il a des réponses extrêmement pertinentes en fait, et c'est parfois plus aidant que d'appeler mon infirmier ». Ce sentiment est partagé par un autre participant qui exprime : « C'est vrai que ça fait du bien d'avoir un interlocuteur ». Il ajoute que même la simple réception d'une réponse, « peu importe qu'elle soit bien ou pas », apporte un soulagement émotionnel, expliquant que cela « soulage en fait d'avoir posé la question et d'avoir vu qu'il y a quelque chose qui répond ». Cette interaction avec l'IA est parfois perçue comme une source de réconfort ou de plaisir immédiat, comme le résume un participant en affirmant : « ChatGPT est parfois mon *shot* de dopamine ».

De ce fait, l'IA apparaît comme un outil de soutien potentiel pour les collaborateurs en entreprise confrontés à des troubles psychiques, en favorisant un environnement de travail plus inclusif et en améliorant leur bien-être global. Ces premières observations soulignent la diversité des fonctions de l'IA dans le domaine de la santé mentale, tout en illustrant son rôle dans le renforcement du soutien psychosocial. Les entreprises pourraient percevoir l'IA comme soutien au bien-être psychologique des collaborateurs.

Les participants ont témoigné d'une grande facilité d'adaptation à ces outils et manifestaient peu d'inquiétude quant à la possibilité d'être remplacés par des machines dans leurs fonctions ou activités. Cependant, une crainte majeure réside dans le risque de développer une dépendance émotionnelle ou psychologique à l'utilisation de ces technologies, avec un risque élevé de crises d'angoisse en cas d'impossibilité d'accès à l'outil. Ces observations mettent en lumière l'importance de prévenir ces formes de dépendance et de promouvoir un usage équilibré des IA pour éviter ces impacts négatifs sur la santé mentale, pouvant mener à de potentiels drames humains.

## Limites

Les limites de l'étude résident dans la taille réduite du panel, une surreprésentation des femmes et une homogénéité des profils en termes d'âge (tous les participants ont une trentaine d'années) et de parcours académiques. Ces facteurs pourraient influencer la facilité perçue d'adoption de l'outil et limiter la généralisation des résultats.

## FUTURES PERSPECTIVES

Ces résultats préliminaires montrent que l'IA réduit la charge cognitive et le stress, mais présente des défis en termes de régulation émotionnelle. H1 semble donc se confirmer. Cependant, l'étude pilote ne nous a pas permis de confirmer H2 et H3, qui devront être testées auprès d'un panel plus large et plus diversifié. L'étude pilote nous a en outre permis de faire émerger une quatrième hypothèse selon laquelle un usage régulier de l'IA présente un risque de dépendance pouvant accentuer l'anxiété des employés.

Nous envisageons de poursuivre cette recherche dans le cadre d'une étude quantitative s'appuyant sur un échantillon plus large, avec une diversité des parcours académiques, d'âge et un équilibre en termes de genre. Le questionnaire devrait également intégrer des questions sur l'IA en tant que soutien à la santé mentale et sur le stress lié à la dépendance afin de tester les quatre hypothèses. Le questionnaire intégrera l'échelle *Generalized Anxiety Disorder 7* (Spitzer *et al.*, 2006) pour mesurer l'anxiété, et le *Perceived Stress Scale* (Cohen, Kamarck & Mermelstein, 1983) pour évaluer le niveau de stress perçu. Il s'agira d'un questionnaire structuré en trois parties visant à mesurer l'anxiété et le stress perçu chez les employés utilisant l'IA. La première partie recueille les informations démographiques essentielles (âge, sexe, niveau d'éducation, expérience professionnelle, etc.). Les parties suivantes évaluent l'anxiété et le stress perçu, inspirées respectivement de l'échelle GAD-7 et de l'échelle PSS. Les questions quantitatives sont conçues pour

**Annexe : Grilles de questions,  
étude exploratoire – IA et santé mentale**

Thème 1 : *Connaissance et expérience des outils d'IA*

Depuis combien de temps travaillez-vous avec des outils d'intelligence artificielle dans votre emploi actuel ?

Comment décririez-vous votre expérience personnelle avec ces outils d'IA ? Y a-t-il des aspects que vous trouvez particulièrement aidants ou, au contraire, difficiles ?

Est-ce que vous ressentez un besoin particulier d'accompagnement pour l'utilisation de ces outils ?

Thème 2 : *Perceptions de l'utilité et des bénéfices des outils d'IA (hypothèse H1)*

Lorsque vous utilisez des outils d'IA, est-ce que vous ressentez une différence dans votre charge mentale ? En quoi cela vous aide-t-il, ou au contraire, vous complique-t-il la tâche ?

Avez-vous le sentiment que ces outils vous facilitent certaines tâches ou vous aident à réduire des efforts mentaux ?

Y a-t-il des moments où l'IA vous donne un sentiment de soulagement ou de soutien dans votre travail ?

Thème 3 : *Exigences et adaptation (hypothèse H2)*

L'arrivée de ces outils d'IA a-t-elle apporté des attentes supplémentaires dans votre travail, et comment vivez-vous cela ?

Est-ce que l'adaptation à ces technologies a parfois pu renforcer chez vous un sentiment de pression ? Si oui, comment vous sentez-vous face à cela ?

Ces outils vous demandent-ils un effort mental particulier pour vous adapter, et comment cet effort affecte-t-il votre bien-être général ?

Thème 4 : *Stress et anxiété liés à l'IA*

Lorsque vous utilisez des outils d'IA, avez-vous remarqué une variation dans votre niveau d'anxiété ou de stress ? Y a-t-il des aspects spécifiques de l'IA qui déclenchent ces ressentis ?

Est-ce qu'il vous arrive de vous sentir tendu ou sous pression en utilisant ces technologies ? Pouvez-vous décrire ce qui vous met le plus à l'aise ou mal à l'aise ?

Comment qualifieriez-vous l'impact de l'IA sur votre niveau d'anxiété global au travail ? Avez-vous constaté des moments où cela était plus intense ?

Thème 5 : *Craintes de remplacement par l'IA (hypothèse H3)*

Est-ce que l'idée que l'IA pourrait éventuellement remplacer certaines de vos compétences vous préoccupe ? Comment vivez-vous cette possibilité ?

Si vous ressentez une insécurité liée à l'IA, comment cela affecte-t-il votre bien-être et votre sentiment de stabilité au travail ?

Est-ce que cette préoccupation, si elle existe, influence votre manière d'envisager votre rôle et votre avenir professionnel ?

Thème 6 : *Stratégies d'adaptation et de gestion du stress*

Lorsque vous ressentez du stress ou de l'anxiété liés à l'utilisation des outils d'IA, avez-vous des moyens personnels pour y faire face ? Y a-t-il des méthodes ou pratiques que vous trouvez aidantes ?

Votre organisation vous a-t-elle proposé un soutien pour mieux gérer les situations liées à l'IA qui peuvent être stressantes ? Quelles initiatives vous sembleraient utiles ?

Quelles actions de soutien ou d'accompagnement, selon vous, aideraient à mieux intégrer ces technologies tout en protégeant votre bien-être mental ?

*Thème 7 : Impact global de l'IA sur la santé mentale*

Avec le recul, comment percevez-vous l'impact de l'IA sur votre santé mentale et votre équilibre au travail ?

Auriez-vous des suggestions à partager sur la manière dont l'IA pourrait être intégrée pour mieux respecter et protéger la santé mentale des employés ?

fournir un score clair de l'impact psychologique, facilitant l'analyse des corrélations entre l'utilisation de l'IA et la santé mentale des employés.

À l'issue de ces travaux, nous envisageons de mener une recherche-intervention autour de la mise en place de dispositifs d'IA d'accompagnement et soutien d'employés concernés par des troubles psychiques.

## RÉFÉRENCES BIBLIOGRAPHIQUES

BAIDOO-ANU D. & OWUSU ANSAH L. (2023), "Education in the era of generative artificial intelligence (AI) : Understanding the potential benefits of ChatGPT in promoting teaching and learning", *SSRN Electronic Journal*, <https://doi.org/10.2139/ssrn.4337484>

BAKKER A. B. & DEMEROUTI E. (2007), "The Job Demands-Resources model: State of the art", *Journal of Managerial Psychology*, 22(3), pp. 309-328, <https://doi.org/10.1108/02683940710733115>

BALCOMBE L. (2023), "AI chatbots in digital mental health", *Informatics*, 10(4), 82, <https://doi.org/10.3390/informatics10040082>

COHEN S., KAMARCK T. & MERMELSTEIN R. (1983), "A global measure of perceived stress", *Journal of Health and Social Behavior*, 24(4), pp. 385-396, <https://psycnet.apa.org/doi/10.2307/2136404>

DAVIS F. D. (1989), "Perceived usefulness, perceived ease of use, and user acceptance of information technology", *MIS Quarterly*, 13(3), pp. 319-340, <https://doi.org/10.2307/249008>

DOSHI R. H., BAJAJ S. S. & KRUMHOLZ H. M. (2023), "ChatGPT: Temptations of progress", *The American Journal of Bioethics*, 23(4), pp. 6-8, <https://doi.org/10.1080/15265161.2023.2180110>

KUMAR N. (2023), "Editorial opinion on employee turnover due to ChatGPT and its impact on mental health", *Psychology & Psychological Research International Journal*, 8(1), <https://doi.org/10.23880/pprij-16000321>

LI J. (2023), "Security implications of AI chatbots in health care", *Journal of Medical Internet Research*, 25, <https://doi.org/10.2196/47551>

SPITZER R.L., KROENKE K., WILLIAMS J.B.W. & LÖWE B. (2006), "A brief measure for assessing generalized anxiety disorder: the GAD-7", *Archives of Internal Medicine*, 166(10), pp. 1092-1097, <https://doi.org/10.1001/archinte.166.10.1092>

STOVER D. (2023), "Will AI make us crazy?", *Bulletin of the Atomic Scientists*, 79(5), pp. 299-303, <https://doi.org/10.1080/00963402.2023.2245247>

# Législations et dialogue social européens autour de l'intelligence artificielle

Par Franck GAMBELLI

Conseiller de l'Union des industries et métiers de la métallurgie (UIMM)

Le nouveau règlement européen sur l'intelligence artificielle s'appliquera en se superposant à des législations existantes ou à venir applicables également aux systèmes d'IA génératives. Le présent article donne un coup de sonde dans les législations sociales européennes et celles concernant la mise sur le marché des produits. À ce stade, les discussions entre partenaires sociaux européens permettent d'y voir plus clair sur les principaux enjeux pour les employeurs et les salariés.

## Présentation de l'Union des industries et métiers de la métallurgie (UIMM)

L'UIMM est une organisation professionnelle à vocation sociale réunissant 59 chambres sociales territoriales et 11 fédérations professionnelles représentant les secteurs de la sidérurgie, des industries mécaniques, électriques et électroniques, aéronautiques, automobiles, navales. UIMM - La Fabrique de l'Avenir, <https://uimm.lafabriquedelavenir.fr/industrie/>

Le nouveau règlement européen sur l'IA<sup>1</sup> la définit comme « un système automatisé qui est conçu pour fonctionner à différents niveaux d'autonomie et peut faire preuve d'une capacité d'adaptation après son déploiement, et qui, pour des objectifs explicites ou implicites, déduit, à partir des entrées qu'il reçoit, la manière de générer des sorties telles que des prédictions, du contenu, des recommandations ou des décisions qui peuvent influencer les environnements physiques ou virtuels » (voir Encadré 1). Le règlement complète ou se superpose à d'autres législations européennes applicables en partie aux systèmes d'IA. Un survol des législations sociales et des législations techniques européennes peut en donner un aperçu : c'est l'objet de la première partie. On pourrait sans doute faire le même exercice pour d'autres *corpus* législatifs comme celui régissant les activités financières. Les documents issus de la négociation et du dialogue social européen permettent d'y voir plus clair sur les enjeux sociaux de l'IA en entreprise : la deuxième partie de l'article en traite.

<sup>1</sup> Règlement (UE) 2024/1689 du 13 juin 2024 établissant des règles harmonisées concernant l'intelligence artificielle.

### Encadré 1 : Définition de l'intelligence artificielle

On peut regretter que la définition des « experts indépendants de haut niveau » de la Commission n'ait pas été reprise : “Artificial intelligence (AI) systems are software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal. AI systems can either use symbolic rules or learn a numeric model, and they can also adapt their behaviour by analysing how the environment is affected by their previous actions”.

Elle avait le mérite de souligner qu'il n'y a d'intelligence, de volonté et de conscience que chez les « humains ».

Le groupe expert définissait également l'IA comme discipline scientifique : “As a scientific discipline, AI includes several approaches and techniques, such as machine learning (of which deep learning and reinforcement learning are specific examples), machine reasoning (which includes planning, scheduling, knowledge representation and reasoning, search, and optimization), and robotics (which includes control, perception, sensors and actuators, as well as the integration of all other techniques into cyber-physical systems)”.

Voir European Commission, “A definition of AI: Main capabilities and scientific disciplines High-Level Expert Group on Artificial Intelligence”, <https://digital-strategy.ec.europa.eu/en/library/definition-artificial-intelligence-main-capabilities-and-scientific-disciplines>

## LES LÉGISLATIONS EUROPÉENNES AUTOUR DE L'IA

### Les législations sociales

Dans le domaine de l'emploi, des conditions de travail, de la protection sociale, l'Union européenne ne peut édicter que des prescriptions minimales<sup>2</sup>. Le Traité interdit explicitement au Parlement et au Conseil de procéder à l'harmonisation forcée des législations sociales nationales<sup>3</sup>. C'est la raison pour laquelle la Commission recourt principalement à des directives et non à des règlements ce qui permet aux droits nationaux d'être plus contraignants et de mieux respecter les cultures sociales locales. C'est aussi pourquoi l'important *corpus* technique des 24 directives sur la santé et à la sécurité au travail ne peut s'appuyer sur la normalisation technique qui impliquerait une harmonisation « totale » et pas « minimale ».

La directive sur le travail en plateforme<sup>4</sup> encadre la surveillance du travail et la prise de décision par algorithme sur les plateformes. Ces systèmes doivent faire l'objet d'une consultation préalable des salariés et doivent protéger les données personnelles. Leur déployeur doit être capable de fournir des informations, des explications motivées

---

<sup>2</sup> Article 151 du Traité sur le fonctionnement de l'Union européenne (TFUE).

<sup>3</sup> Article 153 TFUE.

<sup>4</sup> Directive (UE) 2024/2831 du 23 octobre 2024 au JOUE du 11 novembre 2024.

voire des rectifications sur le système et les décisions. Un contrôle humain est exigé. Ces systèmes d'IA ne doivent pas exercer « une pression induite sur les travailleurs des plateformes ou mettre en danger d'une autre façon la sécurité et la santé physique et mentale des travailleurs des plateformes »<sup>5</sup>. Des mesures préventives contre le risque de harcèlement et de violence doivent être mises en place, notamment à travers des canaux de signalement.

### **Encadré 2 : Directive 2002/17/CE du 11 mars 2002, article 7**

L'article 4 prévoit : « L'information sur l'évolution récente et l'évolution probable des activités de l'entreprise ou de l'établissement et de sa situation économique ; l'information et la consultation sur la situation, la structure et l'évolution probable de l'emploi au sein de l'entreprise ou de l'établissement, ainsi que sur les éventuelles mesures d'anticipation envisagées, notamment en cas de menace sur l'emploi ; l'information et la consultation sur les décisions susceptibles d'entraîner des modifications importantes dans l'organisation du travail ou dans les contrats de travail ».

Les exigences de cette directive sont donc confortées par le règlement IA et la directive « plateforme ».

La consultation des salariés était déjà prévue mais de manière plus générale par la directive sur la consultation des travailleurs (voir Encadré 2). Les 24 directives sur la sécurité au travail détaillent les obligations de prévention pour les employeurs. Elles ont parfois leur pendant dans la législation de libre circulation des produits applicables aux fabricants<sup>6</sup>. Ce sont alors des directives imposant des *maxima* aux États de manière à obtenir une « libre circulation ». Elles sont plus détaillées puisque toute la matière juridique et technique doit être « européenne ». La directive-cadre sur la santé au travail<sup>7</sup> pose un principe de prévention pour « tous les aspects liés au travail »<sup>8</sup> susceptibles de porter atteinte à l'intégrité physique ou mentale des salariés. Elle s'applique à quasiment tous les secteurs d'activités<sup>9</sup>. Elle crée une obligation générale de prévention à la charge des employeurs et dans une moindre mesure aux salariés (voir Encadré 3) : les systèmes d'IA rentrent dans ce champ<sup>10</sup>. Le principe d'analyse et de gestion du risque est méthodique et hiérarchisé

<sup>5</sup> Article 12-b) -3.

<sup>6</sup> Par exemple, la directive sociale sur l'utilisation des « équipements de travail » a pour « alter ego » la directive sur la libre circulation des machines, la directive « sociale » sur les atmosphères explosibles répond à la directive sur la conception des équipements ATEX, la directive sur l'utilisation des équipements de protection individuelle à celle sur leur conception, etc.

<sup>7</sup> Directive 89/391 du 12 juin 1989.

<sup>8</sup> Article 5-1.

<sup>9</sup> Article 2.

<sup>10</sup> Article 6.1 : « L'ensemble des dispositions ou des mesures prises ou prévues à tous les stades de l'activité dans l'entreprise en vue d'éviter ou de diminuer les risques professionnels ». Ce qui consiste à prendre « les mesures nécessaires pour la protection de la sécurité et de la santé des travailleurs, y compris les activités de prévention des risques professionnels, d'information et de formation ainsi que la mise en place d'une organisation et de moyens nécessaires ».

**Encadré 3 : Article 13 de la directive 89/391/CEE  
du Conseil, du 12 juin 1989, concernant la mise en œuvre  
de mesures visant à promouvoir l'amélioration  
de la sécurité et de la santé des travailleurs au travail**

« 1. Il incombe à chaque travailleur de prendre soin, selon ses possibilités, de sa sécurité et de sa santé ainsi que de celles des autres personnes concernées du fait de ses actes ou de ses omissions au travail, conformément à sa formation et aux instructions de son employeur.

2. Afin de réaliser ces objectifs, les travailleurs doivent en particulier, conformément à leur formation et aux instructions de leur employeur : a) utiliser correctement les machines, appareils, outils, substances dangereuses, équipements de transport et autres moyens ; b) utiliser correctement l'équipement de protection individuelle mis à leur disposition et, après utilisation, le ranger à sa place) ; c) ne pas mettre hors service, changer ou déplacer arbitrairement les dispositifs de sécurité propres notamment aux machines, appareils, outils, installations et bâtiments, et utiliser de tels dispositifs de sécurité correctement ; d) signaler immédiatement, à l'employeur et/ou aux travailleurs ayant une fonction spécifique en matière de protection de la sécurité et de la santé des travailleurs, toute situation de travail dont ils ont un motif raisonnable de penser qu'elle présente un danger grave et immédiat pour la sécurité et la santé ainsi que toute défectuosité constatée dans les systèmes de protection ; e) concourir, conformément aux pratiques nationales, avec l'employeur et/ou les travailleurs ayant une fonction spécifique en matière de protection de la sécurité et de la santé des travailleurs, aussi longtemps que nécessaire, pour permettre l'accomplissement de toutes les tâches ou exigences imposées par l'autorité compétente afin de protéger la sécurité et la santé des travailleurs au travail ; f) concourir, conformément aux pratiques nationales, avec l'employeur et/ou les travailleurs ayant une fonction spécifique en matière de protection de la sécurité et de la santé des travailleurs, aussi longtemps que nécessaire, pour permettre à l'employeur d'assurer que le milieu et les conditions de travail sont sûrs et sans risques pour la sécurité et la santé à l'intérieur de leur champ d'activité. »

et bien adapté pour l'IA<sup>11</sup>. La directive impose à l'employeur de « faire en sorte que la planification et l'introduction de nouvelles technologies fassent l'objet de consultations avec les travailleurs et/ou leurs représentants en ce qui concerne les conséquences sur la sécurité et la santé des travailleurs, liées au choix des équipements, à l'aménagement des conditions de travail et à l'impact des facteurs ambiants au travail »<sup>12</sup>. La directive sur l'utilisation des équipements de travail<sup>13</sup> détermine les règles fondamentales d'utilisation des machines et autres équipements à des fins professionnelles. Elle a imposé, en son temps, une mise en conformité du parc machines ancien, en service au moment de son adoption<sup>14</sup>. Elle détermine, entre autres, les mesures de sécurité des systèmes, ce qui inclut nécessairement l'IA. « Les systèmes de commande doivent être sûrs et être choisis compte tenu des défaillances, des perturbations et des contraintes prévisibles dans le

---

<sup>11</sup> Article 6.2.

<sup>12</sup> Article 10.2

<sup>13</sup> Directive 2009/104/CE du 16 septembre 2009, consolidation de la directive 89/655/CE et de ses modifications.

<sup>14</sup> Article L.4324-5 du code du travail.

cadre de l'utilisation projetée ». Elle impose aux employeurs d'acquérir des équipements adaptés<sup>15</sup> et conformes à la législation européenne et donc au règlement IA et cybersécurité<sup>16</sup>. La machine doit être maintenue au même niveau de sécurité pendant tout son cycle de vie, quitte à ajouter, si nécessaire, des sécurités complémentaires<sup>17</sup>. La directive sur les règles de sécurité pour les lieux de travail<sup>18</sup> définit les exigences minimales de sécurité des bâtiments, locaux, équipements, installations, voies de circulation, etc. La directive sur les équipements de visualisation<sup>19</sup> définit des règles pour le travail sur écran. Les systèmes d'IA sont concernés dès lors qu'ils gèrent les aspects sécurité de ces installations ou génèrent des dysfonctionnements dangereux (portes, éclairage, ventilation, etc.). Une disposition est particulièrement pertinente pour l'IA. « Pour l'élaboration, le choix, l'achat et la modification de logiciels ainsi que pour la définition des tâches impliquant l'utilisation d'écrans de visualisation, l'employeur tiendra compte des facteurs suivants : a) le logiciel doit être adapté à la tâche à exécuter ; b) le logiciel doit être d'un usage facile et doit, le cas échéant, pouvoir être adapté au niveau de connaissance et d'expérience de l'utilisateur ; aucun dispositif de contrôle quantitatif ou qualitatif ne peut être utilisé à l'insu des travailleurs ; c) les systèmes doivent fournir aux travailleurs des indications sur leur déroulement ; d) les systèmes doivent afficher l'information dans un format et à un rythme adaptés aux opérateurs ; e) les principes d'ergonomie doivent être appliqués en particulier au traitement de l'information par l'homme. »<sup>20</sup>

## La législation européenne des produits

Une importante communication de la Commission<sup>21</sup> expose le nouveau cadre de la législation technique européenne appliquée depuis 2022 (NCL, New Legislative Framework ; NLF). Elle s'applique aux logiciels pilotant les produits<sup>22</sup>. Elle s'appuie sur des normes européennes sous mandat de la Commission qui vont produire une « présomption de conformité »<sup>23</sup>. Le règlement sur la sécurité générale des produits<sup>24</sup> sert de « voiture balais » en complétant les législations existantes et en couvrant les produits accessibles aux consommateurs qui y auraient échappés<sup>25</sup>. Outre une obligation générale de sécurité, il impose une analyse de risques<sup>26</sup> incluant les effets induits sur d'autres produits y compris sous l'angle de l'interconnexion. L'évaluation du risque s'applique aussi « lorsque la nature du produit l'exige » aux « fonctionnalités évolutives, d'apprentissage et prédic-

---

<sup>15</sup> Article 3 : « L'employeur prend les mesures nécessaires afin que les équipements de travail mis à la disposition des travailleurs dans l'entreprise ou l'établissement soient appropriés au travail à réaliser ou convenablement adaptés à cet effet, permettant d'assurer la sécurité et la santé des travailleurs lors de l'utilisation de ces équipements de travail ».

<sup>16</sup> Article 4 I. a).

<sup>17</sup> Article 4.2.

<sup>18</sup> Directive 89/654 : CEE du 30 novembre 1989.

<sup>19</sup> Directive 90/270/CE du 29 mai 1990.

<sup>20</sup> Annexe point 3.

<sup>21</sup> Le livre bleu : Communication de la Commission : Le « guide bleu » relatif à la mise en œuvre de la réglementation de l'UE sur les produits 2022. 2022/C 247/01 au JOUE du 29 juin 2022.

<sup>22</sup> Guide bleu précité page 19.

<sup>23</sup> Norme harmonisée au sens du règlement 1025-212 article 2, point 1, point c.

<sup>24</sup> Le nouveau règlement européen 2023/988 applicable le 13 décembre 2024 remplace la directive 2001/95/CE.

<sup>25</sup> Article 2 du règlement 2023/988 relatif à la sécurité général des produits du 10 mai 2023.

<sup>26</sup> Article 6 du règlement précité.



tives du produit »<sup>27</sup>. Le règlement machines<sup>28</sup> ne traite pas explicitement d'IA car il est censé être technologiquement neutre. Le risque ne fait pas l'objet d'un jugement *a priori* sur une technologie donnée mais d'une évaluation rigoureuse qui l'objective<sup>29</sup>. On voit donc que ce choix est différent de celui du règlement IA<sup>30</sup>. On trouvera donc la question de l'IA traitée dans une exigence générale applicable à tous les systèmes de commande de machines (voir Encadré 4). Le règlement soumet toutefois les systèmes et les machines totalement ou partiellement auto-évolutifs et qui utilisent des apprentissages automatiques assurant des fonctions de sécurité à une procédure d'examen par tierce partie. C'est dans cette procédure que seront évalués les systèmes d'IA des machines. Un énorme programme de normalisation est en cours d'élaboration pour répondre aux besoins de ce nouveau règlement et, en particulier à la nécessité d'intégrer dans le *corpus* normatif de la sécurité des machines les apports de la normalisation de l'IA. L'IA va devoir s'incarner dans le contexte industriel de la machine et apprendre à parler le langage des ingénieurs de bureaux d'études. Le règlement sur la « surveillance du marché et la conformité des produits » s'applique à toutes les législations techniques dont le règlement IA<sup>31</sup>. Il est assorti en France de lourdes sanctions pénales<sup>32</sup>. Reste à savoir si les sanctions françaises

#### Encadré 4 : Annexe III point 1.2

« Les systèmes de commande des machines ou des produits connexes au comportement ou à la logique totalement ou partiellement auto-évolutifs qui sont conçus pour fonctionner avec différents niveaux d'autonomie sont conçus et construits de manière : à ne pas amener la machine ou le produit connexe à exécuter des actions allant au-delà de leur espace défini de travail et de mouvement ; à ce que l'enregistrement des données relatives au processus décisionnel en matière de systèmes de sécurité basés sur des logiciels assurant des fonctions de sécurité, y compris de composants de sécurité, après la mise sur le marché ou la mise en service des machines ou des produits connexes soit activé et à ce que ces données soient conservées pendant un an après leur collecte, exclusivement pour démontrer la conformité des machines ou des produits connexes avec la présente annexe sur demande motivée d'une autorité nationale compétente ; à ce qu'il soit possible à tout moment de corriger la machine ou le produit connexe afin de préserver leur sécurité intrinsèque ».

---

<sup>27</sup> Article 6 point 1, point h.

<sup>28</sup> Pour une comparaison entre la directive machines et le nouveau règlement, voir le travail comparatif réalisé par Eurogip et l'Etui, ici : <https://eurogip.fr/machines-analyse-comparative-entre-le-reglement-de-2023-et-la-directive-de-2006/>

<sup>29</sup> Voir annexe III Partie B Principes généraux du règlement machines.

<sup>30</sup> Une machine très traditionnelle, par exemple un palan, un cric, une pompe, une scie, peut se révéler aussi, voire plus dangereuse qu'un robot ou qu'une machine pilotée avec de l'IA dernier cri. L'attention est centrée sur le risque : défaut d'ergonomie, éclairage, manutention, matériaux, produits, systèmes de commande, risque mécaniques, stabilité, rupture, chutes, variation de fonctionnement, éléments mobiles de transmission du mouvement ou de travail, protecteurs, alimentation, énergie, température, incendie, explosion, bruit, vibration, rayonnements, maintenance...

<sup>31</sup> Règlement 2019/1020 du 20 juin 2019 sur la surveillance du marché et la conformité des produits.

<sup>32</sup> Loi n°2021-2018 du 2 août 2021 codifié sous les articles L.4311-6, et pour sur la surveillance du marché L.4314-1 du code du travail, L.4314-1 à L.4314-2, pour les sanctions L.4746-1 : L.4755-3 (50 000 à 200 000 euros, jusqu'à 2 ans de prison), R4746-1, R4746-3.

se cumuleront avec celle du règlement IA. Il donne aux autorités de contrôle un droit de copie des logiciels<sup>33</sup>.

Le règlement IA<sup>34</sup> interdit les systèmes d'IA manipulateurs, trompeurs, subliminaux, etc. Il classe comme « système à haut risque » ceux qui sont désignés comme tels par leur législation (machines et composants fonctionnant par de l'IA, les systèmes biométriques, ceux utilisés sur des infrastructures critiques, ou pour l'éducation et la formation professionnelle, l'emploi, la gestion de main-d'œuvre, l'accès aux services privés essentiels et aux services publics). Ces systèmes devront avoir un système de gestion du risque soumis à des essais réalisés par un organisme « notifié » tierce partie. Les dépoyeurs des « systèmes » dès lors qu'ils ont la qualité d'employeur, devront consulter les salariés avant leur mise en place. S'ils ne respectent pas le droit local issu des directives imposant déjà la consultation, le règlement s'appliquera. Étrange incursion d'un règlement « libre circulation » dans le domaine réservé des directives « sociales ». Le règlement sera accompagné par deux directives sur la responsabilité civile de l'IA. La première révisé la directive de 1985 sur la responsabilité civile du fait des produits défectueux<sup>35</sup> pour l'étendre à l'IA qu'ils utilisent<sup>36</sup>. La seconde déterminera les règles de responsabilité civile extracontractuelle du fait de l'IA<sup>37</sup>. L'administration de la preuve sera simplifiée pour faciliter l'indemnisation des victimes. Ces deux directives ne changeront pas le régime d'indemnisation des accidents du travail sur le plan national mais pourront peut-être faciliter les actions récursoires contre les tiers. Plusieurs règlements relatifs aux marchés des données constituent le cadre légal européen du déploiement de l'IA<sup>38</sup>. Il est possible que ce soit le Data act qui ouvre le marché des données qui ait le plus de conséquences sociales sur les fabricants utilisant l'IA car les données de sortie appartiendront à l'utilisateur. Fabricant, ingénieurs, opérateurs pourraient voir la maîtrise du produit leur échapper au profit des gestionnaires de *data*. Par ailleurs, le règlement IA est neutre en matière de propriété intellectuelle (voir Encadré 5) et à ce jour les *datas* de sortie de l'IA ne sont pas protégeables par le droit d'auteur sauf à prouver l'intervention humaine. Pour les inventions de salariés brevetables, la question pourrait se poser de la part prise par le salarié dans une création par IA générative<sup>39</sup>.

---

<sup>33</sup> Article 14.4.

<sup>34</sup> Règlement 2024/1689 du 13 juin 2024.

<sup>35</sup> Voir nos articles 1245 à 1245-17 de notre code civil.

<sup>36</sup> Position du Parlement européen arrêtée en première lecture le 12 mars 2024 en vue de l'adoption de la directive (UE) 2024 relative à la responsabilité du fait des produits défectueux et abrogeant la directive 85/374/CEE.

<sup>37</sup> Proposition de directive relative à l'adaptation des règles en matière de responsabilité civile extracontractuelle au domaine de l'intelligence artificielle du 28 septembre 2022.

<sup>38</sup> Le RGPD, pour les données personnelles. Le Digital Markets Act régule l'activité des grandes plateformes, hébergeurs, messageries réseaux sociaux en leur imposant des exigences de transparence et d'interopérabilité. Le Data governance act (DGA) organise l'exploitation de certaines données venant du secteur public. Le Data services act (DSA) fait la police des contenus échangés en responsabilisant les grandes plateformes, les fournisseurs d'accès internet, les hébergeurs. Le Data act ouvre le marché des données entre entreprises et avec les consommateurs et en particulier celles provenant des objets connectés. Les utilisateurs auront des droits d'accès vis-à-vis des fabricants.

<sup>39</sup> WIPO (2024), « L'IA générative : saisir les enjeux en matière de propriété intellectuelle », Organisation mondiale de la propriété intellectuelle, <https://www.wipo.int/edocs/pubdocs/fr/wipo-pub-rn-2024-8-fr-generative-ai-navigating-intellectual-property.pdf>

### Encadré 5 : Considérant 88

« Tout au long de la chaîne de valeur de l'IA, plusieurs parties fournissent souvent des systèmes, des outils et des services d'IA, mais aussi des composants ou des processus que le fournisseur intègre dans le système d'IA avec plusieurs objectifs, dont l'entraînement de modèles, le réentraînement de modèles, la mise à l'essai et l'évaluation de modèles, l'intégration dans des logiciels ou d'autres aspects du développement de modèles. Ces parties ont un rôle important à jouer dans la chaîne de valeur vis-à-vis du fournisseur du système d'IA à haut risque dans lequel leurs systèmes, outils, services, composants ou processus d'IA sont intégrés, et devraient fournir à ce fournisseur, en vertu d'un accord écrit, les informations, les capacités, l'accès technique et toute autre assistance nécessaires sur la base de l'état de la technique généralement reconnu, afin de lui permettre de se conformer pleinement aux obligations énoncées dans le présent règlement, sans compromettre leurs propres droits de propriété intellectuelle ou secrets d'affaires ».

## LES RÉFLEXIONS DES PARTENAIRES SOCIAUX AUTOUR DE L'IA

L'accord cadre européen de juin 2020 sur la digitalisation, signé par les organisations inter-professionnelles d'employeurs et de salarié européennes propose une méthode générale d'introduction du digital dans les entreprises reposant sur la formation, la sécurisation des emplois, les modalités de connexion et de déconnexion, la garantie que l'intelligence artificielle reste sous contrôle humain, le respect de la dignité humaine notamment dans la surveillance. Ce processus concerne l'organisation du travail, le contenu du travail, les conditions de travail et les relations de travail. L'accord détaille chaque phase du processus. Il prévoit la possibilité de contester une décision fondée uniquement sur les résultats de l'IA.

Les conclusions communes des employeurs européens de la métallurgie « CEEMET » et des salariés de l'industrie, l'industriAll European Trade Union du 22 février 2023<sup>40</sup> assignent au dialogue social une fonction de régulation dans l'introduction de l'intelligence artificielle justifiée par le fait qu'employeur et salarié ont la connaissance et l'expérience pour discuter du type d'intelligence artificielle qu'ils veulent déployer dans les postes de travail. Ils préconisent la consultation en amont des salariés, ce qui permet d'assurer la pertinence du système d'IA et la confiance de tous. La formation joue un rôle clef dans le processus notamment pour en garder le contrôle. « Les accords collectifs constituent un outil à valeur ajoutée pour identifier et traiter les besoins en compétences des industries technologiques et pour canaliser les investissements dans les qualifications de la main-d'œuvre existante ». Un développement important est consacré à la sécurité et la santé au travail. L'IA peut contribuer à limiter les risques et la pénibilité. L'insistance est mise sur une analyse de risque *in situ* prenant en compte l'organisation du travail, les espaces, les flux. Le système doit s'adapter aux tâches et au salarié, pas l'inverse. « Ces principes doivent être appliqués à l'apprentissage automatique et aux algorithmes auto-développés. Tout système doit avoir sa finalité et ses limites clairement définies en amont, notamment au travers des spécifications industrielles données par les utilisateurs

---

<sup>40</sup> <https://ceemet.org/position-papers/met-social-partners-reach-joint-conclusions-on-artificial-intelligence/>

finaux ». Et « Le champ d'action de l'homme et de la machine, ainsi que la capacité d'adaptation et la tolérance aux erreurs, doivent, dans la mesure du possible, être évalués de manière claire et transparente. Le travailleur doit toujours être conscient de la possibilité et des modalités d'intervention dans le processus numérisé... « En outre, il est également important de veiller à ce que le savoir-faire industriel qui constitue l'expérience de l'entreprise ne disparaisse pas. Un système logiciel, aussi autonome et sophistiqué soit-il, ne pourra jamais égaler l'intelligence et la culture industrielle. Par exemple, dans la conception par jumeau numérique, la prise en compte du savoir-faire mécanique humain reste indispensable. La gestion automatique des données, aussi puissante soit-elle, ne peut pas tout faire ».

Le dialogue social voire la négociation collective de branche et d'entreprise se développent progressivement dans les activités où l'IA est au cœur du métier. Voir l'initiative de la Fédération syndicale européenne des salariés des services : l'Uni Global Europa, qui a créé une base de données pour recenser les clauses des accords collectifs traitant de l'IA et de la gestion algorithmique. Les secteurs qui négocient sont ceux qui recourent massivement aux technologies d'IA : tourisme, administration publique, transport, logistique, communication, commerce, banques, finances, assurances, consulting. Les accords se focalisent principalement sur la formation, la santé, la consultation des salariés et le suivi dans le temps, la déconnexion. Voir également en France, un rapport du Centre d'Études de l'Emploi et du Travail (CEET) réalisé à la demande de Force ouvrière et du Cnam qui constate que la proportion d'accords faisant mention de l'IA aurait été multipliée par 2,5 entre 2018 et 2023. Voir également les contributions de l'étude de l'ETUI, <https://www.etui.org/publications/artificial-intelligence-labour-and-society>

Pour conclure nous reprendrons les recommandations finales des partenaires sociaux européens figurant dans les conclusions communes de la métallurgie européenne.

« Étant donné que l'IA sur le lieu de travail n'est généralement pas une technologie autonome, mais une partie intégrante d'un système de travail, il est important de noter que les répercussions de l'IA dépendent toujours de la configuration concrète du système de travail dans lequel elle est intégrée. Il est important de noter qu'en façonnant le système d'IA dans le contexte de l'emploi, nous devons toujours façonner de manière proactive l'organisation du travail et l'affectation du personnel. En d'autres termes, la dimension sociale doit toujours être conçue parallèlement à la dimension technique d'un système d'IA afin de s'assurer que le système n'est pas conçu au détriment de la qualité de l'emploi. Autrement dit, l'organisation du travail humain doit suivre l'organisation du travail technique. Cela signifie que la conception technique et sociale devra évoluer en permanence pour s'assurer que les nouveaux défis sont relevés de manière adéquate et en temps utile. Pour garantir une mise en œuvre réussie du changement technologique, les partenaires sociaux au niveau de l'entreprise devraient collaborer à l'introduction et à l'élaboration de nouveaux systèmes sur le lieu de travail, à un stade précoce ».

Ou, en version originale : "Since AI at the workplace is not usually a stand-alone technology, but an integrated part of a work system, it is important to note that the repercussions of AI also always depend on the concrete configuration of the work system in which they are integrated. It is important to note that, when shaping the AI system in the employment context, we always have to proactively shape work organisation and the deployment of staff. That is to say: the social dimension must always be designed along the technical dimension of an AI system to make sure that the system is not designed to the detriment of quality employment. In other words, the human work organisation needs to keep pace with the technical work organisation. This means that the technical and social design will need to continuously evolve to make sure that emerging challenges are adequately and timely addressed. To guarantee a successful implementation of technological change, social partners at company level should work together on the introduction and shaping of new systems at the workplace, at an early stage".

# IA et Communs : conjuguer puissance technologique et habitabilité terrestre

Par Emmanuelle ROUX

Conseillère du directeur général de l'Institut national de l'information géographique et forestière (IGN)

Face aux crises climatiques, sociales et démocratiques, l'intelligence artificielle peut amplifier les inégalités ou, au contraire, devenir un catalyseur de transformation collective. Pour cela, il faut dépasser son rôle d'outil au service des communs et l'envisager comme un commun en soi, ce qui implique de créer des modèles gouvernés collectivement, comme le projettent des plateformes telles que Hugging Face ou des initiatives *open source* comme BLOOM. En intégrant l'IA dans une logique de communs numériques, nous pourrions non seulement démocratiser son accès, mais aussi réduire ses impacts écologiques, tout en renforçant l'engagement citoyen et la transparence.

Nous vivons une période de mutations profondes. La puissance technologique de l'IA redessine les contours de notre avenir collectif, tandis que les crises s'accumulent et s'intensifient. Le dérèglement climatique menace directement notre capacité à habiter la Terre. Les tensions géopolitiques fragilisent nos équilibres. Notre démocratie s'effrite sous nos yeux.

Comment répondre à ces défis ? Les solutions classiques montrent leurs limites. L'État, malgré ses efforts, ne parvient pas à déployer des réponses suffisamment agiles et adaptées. Le marché, focalisé sur le profit à court terme, échoue à préserver nos biens communs. Dans ce contexte, les IA représentent à la fois un risque et une opportunité : elles pourront amplifier ces crises ou, si nous les mobilisons correctement, devenir de puissantes alliées face à ce tumulte de transformations. Pour cela, nous devons en faire de véritables communs numériques, au service de nos devenirs collectifs.

## LES COMMUNS : UNE RÉPONSE AUX CRISES CONTEMPORAINES

Que pouvons-nous apprendre des communs ? Ne les considérons pas comme une simple utopie théorique. Ils s'enracinent dans une longue histoire d'expérimentations réussies qui nous montrent la voie. Pour comprendre leur fonctionnement, observons leurs trois piliers indissociables : une ressource que nous partageons (qu'il s'agisse d'une forêt ou d'un logiciel), une communauté qui s'engage activement dans sa gestion, et des règles que nous établissons ensemble.

Pourquoi cette approche répond-elle si bien à nos défis actuels ? En multipliant les centres de décision, elle renforce notre résilience collective face aux crises. En redonnant du pouvoir d'agir aux communautés locales, elle ravive l'engagement citoyen. En créant un lien direct entre l'usage d'une ressource et sa préservation, elle garantit une gestion plus durable de nos biens communs.

Voici quelques initiatives concrètes qui démontrent chaque jour la force de ce modèle :

- Bologne, cette ville italienne pionnière où des milliers de citoyens transforment leur quotidien à travers 180 pactes de coopération<sup>1</sup>. Ces pactes ne sont pas de simples conventions administratives – ils incarnent une nouvelle façon de faire société, de l’entretien des jardins partagés à l’organisation d’événements culturels, en passant par l’entraide entre voisins. Le « *Regolamenti* », leur cadre juridique novateur, protège et encourage cette intelligence collective en action.
- Bruxelles, où s’opère une révolution singulière dans le domaine du logement. Le Community Land Trust<sup>2</sup> réinvente notre rapport à l’habitat en séparant la propriété du sol de celle des bâtiments. En impliquant directement les habitants dans les décisions, ce dispositif rend le logement durablement accessible aux familles modestes. Cette approche invite à s’échapper de la spéculation immobilière sans sacrifier la qualité ni la pérennité des habitations.
- Gand, où les communs prennent une dimension institutionnelle remarquable. Avec ses 500 initiatives citoyennes et son plan de transition ambitieux<sup>3</sup>, la ville vous montre comment les institutions peuvent catalyser l’énergie citoyenne sans l’étouffer. Les “neighbourhood managers” y tissent patiemment les liens entre citoyens et administration, nourrissant les projets collectifs tout en préservant leur authenticité.

Et, de ces exemples, retenons aussi que gérer des communs demande un engagement de tous les instants. Il est sans cesse nécessaire d’innover juridiquement, d’inventer des modes de gouvernance sophistiqués, de résoudre les inévitables conflits d’usage. Ces expériences incarnent que les communs ne sont pas qu’une belle idée : ils constituent une réponse concrète à nos défis actuels, alliant justice sociale, préservation de l’environnement et renouveau démocratique.

## LES COMMUNS NUMÉRIQUES : HÉRITAGES ET INNOVATIONS

Le développement des communs numériques depuis plusieurs décennies maintenant s’inscrit dans cette dynamique de réappropriation collective tout en ouvrant de nouvelles perspectives. Ils démontrent qu’une gouvernance partagée peut générer innovation et valeur sociale à grande échelle.

Les succès sont nombreux et variés : Wikipédia illustre parfaitement cette capacité des communautés à produire et maintenir collectivement une ressource informationnelle de haute qualité. Le système d’exploitation Linux démontre quant à lui la puissance de la collaboration ouverte en matière de développement logiciel, devenant aujourd’hui un pilier des infrastructures numériques critiques.

À Barcelone, cette approche prend une dimension politique forte. La ville a déployé en 2017 la plateforme Decidim, développée sous licence libre, qui facilite la contribution des habitants aux décisions municipales tout en protégeant leur vie privée. Le logiciel a été repris par plusieurs centaines d’organisations et est utilisé par le gouvernement brésilien, la Commission européenne et des communes japonaises. Cette initiative illustre comment

---

<sup>1</sup> “Trust the People”, *Civic Engagement and Collaborative Imagination in Bologna, Italy*, <https://cityleadership.harvard.edu/wp-content/uploads/2024/07/20240521-Trust-the-People-FINAL.pdf>

<sup>2</sup> Community Land Trust Bruxelles, le site officiel : <https://cltb.be/>

<sup>3</sup> P2P Foundation France (2017), « Plan de transition vers les communs de la ville de Gand », <https://blogfr.p2pfoundation.net/2017/09/08/plan-de-transition-vers-communs-de-ville-de-gand/>

les communs numériques peuvent renforcer la démocratie participative tout en garantissant l'éthique des données.

La ville de Gand pousse encore plus loin cette logique en intégrant les communs numériques dans une stratégie globale de transition. Son "policy participation unit" travaille activement à connecter les initiatives citoyennes numériques avec les autres formes de communs urbains. Cette approche intégrée permet de démultiplier l'impact des projets tout en assurant leur ancrage territorial.

En France, l'Institut national d'informations géographiques et forestières (IGN) participe pleinement de cette dynamique des communs numériques par ses actions permettant de créer des géocommuns tels que le Référentiel National du Bâtiment, la Base Adresse Nationale ou Panoramax (qui vise à créer une alternative ouverte à Google Street View). Par sa politique d'*open data* et d'*open source*, et en développant des outils permettant de contribuer collectivement à la récolte de données géographiques, l'institut assume qu'une institution publique peut devenir un catalyseur de communs numériques.

Ces expériences nous révèlent que les communs numériques dépassent la simple transposition des principes traditionnels dans l'espace digital. Ils constituent un véritable laboratoire d'innovation sociale et démocratique, inventant de nouvelles formes de collaboration et de gouvernance à grande échelle adaptées aux enjeux contemporains.

## LES IAS COMME LEVIER DE TRANSFORMATION DES COMMUNS

Le domaine de l'intelligence artificielle s'impose aujourd'hui comme un puissant catalyseur de transformation collective. Sa capacité à analyser, prédire et optimiser ouvre des perspectives inédites pour la gestion des communs, tout en soulevant des questions fondamentales pour notre avenir.

### Des potentialités multiples pour la gestion collective

L'aide à la décision constitue l'un des apports majeurs de l'IA pour les communs. En analysant de vastes ensembles de données, elle permet d'identifier des tendances invisibles à l'œil nu et d'établir des prédictions précieuses pour la gestion collective. L'IGN déploie l'intelligence artificielle pour accélérer la production de données environnementales critiques. Le projet CarHab utilise le *machine learning* pour cartographier les habitats naturels et semi-naturels à l'échelle nationale, permettant un suivi précis de la biodiversité. L'OCS-GE (programme permettant de décrire l'occupation du sol à grande échelle) emploie la télédétection par IA pour analyser la couverture des sols et mesurer leur artificialisation, donnant aux collectivités et plus largement aux acteurs territoriaux les moyens de piloter efficacement leur transition écologique.

L'optimisation des ressources représente un second domaine d'application crucial. Les collectivités peuvent désormais détecter en temps réel les anomalies dans leurs réseaux : fuites d'eau, dysfonctionnements électriques, dépôts sauvages. À Lyon, un réseau de capteurs couplé à l'IA permet d'optimiser l'arrosage des espaces verts en fonction des conditions météorologiques et de l'humidité des sols. À Nantes, des algorithmes d'apprentissage contribuent à réduire le gaspillage alimentaire dans les cantines scolaires en prédisant plus précisément les besoins.

La démocratisation des savoirs constitue un troisième axe transformateur. Les IAs rendent accessibles des analyses complexes à des non-experts, permettant aux citoyens de mieux comprendre et participer aux décisions qui les concernent. Les interfaces en

langage naturel et les assistants vocaux devraient rapidement faciliter l'accès aux services publics pour tous, indépendamment du niveau d'éducation ou de maîtrise technologique.

Ces différentes facettes montrent comment l'IA peut faciliter et enrichir la décision collective et donc la gestion de communs.

## Des risques majeurs à maîtriser

Le déploiement de l'IA dans la gestion des communs soulève cependant des enjeux critiques qui nécessitent une vigilance particulière. Le premier risque concerne la concentration du pouvoir technologique entre quelques acteurs dominants qui possèdent les capacités d'intelligence artificielle (jeux de données d'entraînements, puissance de calcul, modèles, etc.) et créent une dépendance préoccupante. Ils peuvent à tout instant décider des conditions d'accès à ces nouveaux services numériques ou encore des sources qu'ils utilisent ou non pour nourrir leurs IAs, sans transparence aucune. L'opacité de ces acteurs est antinomique avec les pratiques des communs et une capacité à décider ensemble.

L'impact environnemental constitue un second défi majeur. La consommation énergétique considérable des modèles d'apprentissage soulève des questions cruciales dans un contexte de transition écologique. Cette empreinte environnementale réelle pourrait aussi être instrumentalisée et servir de prétexte pour limiter l'accès à ces technologies, sous couvert de sobriété numérique, créant ainsi une société à deux vitesses. À l'instar de différents récits d'anticipation et sous prétexte de rationalisation des ressources ou de protection environnementale, l'usage des systèmes les plus performants pourrait être conditionné à des critères financiers restrictifs ou à des appartenances sociales spécifiques. Cette limitation créerait une nouvelle forme d'inégalité, où seules les classes dominantes bénéficieraient pleinement du potentiel transformateur de ces technologies, renforçant ainsi leur position de pouvoir dans la société.

Cette perspective dystopique nous rappelle l'importance de penser l'IA comme un commun dès aujourd'hui, en garantissant son accessibilité universelle et en développant des cadres de gouvernance partagée qui empêchent son accaparement par une minorité.

## CONSTRUIRE LES IAS COMME COMMUNS

### Une architecture repensée pour l'intérêt général

La première condition est de repenser l'architecture même des systèmes d'IA pour privilégier la sobriété et la protection des données. L'approche dominante actuelle, fondée sur des modèles massifs hébergés dans le *cloud*, favorise non seulement la concentration du pouvoir technologique mais génère aussi une empreinte environnementale considérable.

### L'émergence de l'IA décentralisée

Une tendance prometteuse se développe avec l'IA embarquée directement sur les appareils (*edge computing*). Cette approche permet d'exécuter les modèles en local, sur les *smartphones*, capteurs ou autres dispositifs, sans nécessiter de connexion permanente aux serveurs distants. Les avantages sont multiples : réduction drastique de la consommation énergétique, meilleure protection des données personnelles qui ne quittent plus l'appareil, et résilience accrue du système qui peut fonctionner en autonomie.

### Une évolution parallèle des modèles

Indépendamment de cette décentralisation, une autre transformation majeure s'opère dans la conception même des modèles d'IA. D'un côté, des projets collaboratifs comme



BLOOM démontrent la possibilité de créer des systèmes performants de manière distribuée. De l'autre, l'émergence des *small specialist models*, plus légers et spécialisés, ouvre la voie à une IA plus sobre et accessible. Ces modèles spécialisés, bien que conçus initialement pour réduire les coûts et l'empreinte environnementale, se révèlent particulièrement adaptés au déploiement sur des appareils en périphérie.

La convergence de ces deux tendances – architecture décentralisée et modèles spécialisés – dessine l'avenir d'une IA plus démocratique, où chaque utilisateur garde le contrôle de ses données tout en bénéficiant de services intelligents adaptés à ses besoins spécifiques.

## UNE GOUVERNANCE OUVERTE ET COLLABORATIVE

La question fondamentale de l'IA comme commun n'est pas seulement technique mais avant tout démocratique : quelles capacités souhaitons-nous collectivement donner à nos systèmes d'IA et à quel coût environnemental et social ? Cette réflexion doit accompagner en permanence le développement technologique.

### Délibération collective et innovation

Des plateformes comme Hugging Face illustrent déjà la puissance de la collaboration ouverte dans le développement de l'IA. Avec plus de 100 000 modèles disponibles publiquement, elle crée un écosystème dynamique où chercheurs, développeurs et utilisateurs partagent leurs avancées. Cette infrastructure technique s'accompagne naturellement d'un débat continu sur les usages, l'éthique et les orientations collectives de ces technologies.

### Infrastructure partagée et modèles de fondation

L'IGN illustre cette approche en développant des modèles de fondation spécialisés dans l'analyse des territoires. Ces modèles pré-entraînés sur de vastes jeux de données géographiques permettront de réduire significativement les besoins en données d'apprentissage et en puissance de calcul pour de nombreuses tâches d'observation territoriale. En rendant ces modèles complètement ouverts et accessibles, l'IGN facilitera leur réutilisation internationale et permettra d'éviter la surconsommation inutile de ressources matérielles et énergétiques.

La construction de communs numériques est donc d'abord un acte politique qui combine innovation technique et délibération démocratique permanente sur leurs usages et leurs impacts.

L'histoire d'Internet et du *web* ont démontré toute la pertinence d'une telle démarche même si le modèle ouvert du *web* reste à défendre au jour le jour et se fait régulièrement attaquer. Pour les IAs, cette voie est non seulement possible mais nécessaire. C'est uniquement en la construisant comme un commun que nous pourrions veiller à ce qu'elle serve véritablement l'intérêt général plutôt que les intérêts de quelques-uns.

## CONCLUSION

L'émergence de l'IA comme technologie transformative majeure nous place face à un choix de civilisation. Au-delà des enjeux techniques de gouvernance et d'architecture, c'est notre capacité à « faire humanité » qui est en jeu dans un double contexte de déploiement de l'intelligence artificielle et de conscientisation forte des enjeux d'habitabilité de la

Terre. Nous vivons ces deux lames historiques dans le même temps ce qui rend la période d'autant plus singulière.

## **Une opportunité historique**

Les communs numériques, et particulièrement les communs d'IA, offrent une voie pour nous réinventer collectivement et repenser nos manières de décider, de travailler et de coopérer. En développant ces infrastructures partagées, nous pouvons créer les conditions d'une intelligence véritablement collective, où la puissance des algorithmes amplifie plutôt que ne remplace les capacités humaines.

## **Un nouveau contrat social**

L'enjeu n'est plus seulement de produire des technologies performantes, mais de les mettre au service d'un projet de société. Comment l'IA peut-elle nous aider à habiter la Terre autrement ? Comment peut-elle soutenir l'émergence de nouvelles formes de solidarité et d'apprentissage mutuel ? La construction de communs d'IA nous invite à repenser en profondeur notre rapport au savoir, au travail et au vivant.

La réponse à ces questions ne pourra être que collective. C'est en faisant le pari des communs que nous pourrions éviter les écueils d'une IA déshumanisante pour construire les outils de notre émancipation collective.

# Vers une intelligence artificielle “gender by design” ?

Par Peggy VICOMTE

Déléguée générale de Femmes@Numérique

Et Camille SALINESI

Co-directeur de l'Observatoire de l'IA  
de l'Université Paris 1 Panthéon-Sorbonne

L'intelligence artificielle (IA) redéfinit nos vies personnelles et professionnelles, s'imposant comme une technologie ubiquitaire. Omniprésente dans des domaines variés – santé, urbanisme, agriculture, éducation ou encore culture –, elle ouvre de nouvelles perspectives tout en posant d'importants enjeux sociétaux et éthiques.

Si l'IA promet des opportunités inédites, elle amplifie aussi les biais préexistants dans les données qui la nourrissent, renforçant discriminations et stéréotypes, comme en témoignent des exemples documentés dans les domaines du recrutement ou de la santé. Plus encore, l'absence de diversité dans les métiers du numérique aggrave ces phénomènes, illustrant la nécessité d'une IA plus inclusive.

Des initiatives interdisciplinaires, combinant sciences humaines et sociales, visent aujourd'hui à faire de l'IA un levier d'équité et de progrès collectif. Une intelligence artificielle “gender by design” pourrait devenir un modèle pour assurer justice et inclusion dans son développement et ses usages.

## L'IA, NOUVEAU PHÉNOMÈNE DE SOCIÉTÉ

L'intelligence artificielle (IA) désigne traditionnellement un ensemble de technologies capables de réaliser des opérations difficiles pour l'algorithmique classique notamment celles simulant certaines fonctions cognitives humaines telles que la reconnaissance de formes, la prise de décision, l'expression verbale en langage naturel, ou encore l'apprentissage. Les techniques d'IA reposent généralement sur des combinaisons de mécaniques logicielles d'inspiration mathématique ou biologique, d'heuristiques, et d'algorithmes avancés ; elles sont souvent nourries par des volumes massifs de données et leur mécanique générale relève de l'inférence. À travers des outils concrets comme les assistants vocaux, les moteurs de recherche, les recommandations personnalisées sur les plateformes de *streaming*, ou plus récemment les robots conversationnels, l'IA s'est discrètement mais solidement ancrée dans nos vies quotidiennes.

L'IA s'impose aujourd'hui comme une technologie ubiquitaire. Omniprésente, elle transforme non seulement notre manière de vivre, mais également les modèles économiques et sociaux à l'échelle mondiale. Dans la sphère privée, elle s'immisce *via* les *smartphones*, les réseaux sociaux les objets connectés. Dans le monde de l'entreprise, elle améliore les chaînes logistiques, et révolutionne la gestion des données. Dans des domaines aussi variés que la santé (surveillance de nos paramètres biologiques, optimisation de diagnostics médicaux), l'urbanisme (*smart city*), l'agriculture (robotisation), la sécurité (caméras

intelligentes), la culture (création artistique assistée), l'enseignement (*e-learning* personnalisé) le sport (analyse des performances) ou même l'administration (correspondants virtuels), l'IA ouvre sans cesse de nouvelles perspectives. Chaque mois, de nouvelles applications émergent, impactant des secteurs qui, jusqu'alors, semblaient éloignés de cette transformation. Ces technologies se diffusent très vite dans la société. Le lancement de ChatGPT en novembre 2022 a ainsi marqué un tournant dans la démocratisation des technologies d'IA générative. En seulement deux mois, l'outil comptait 100 millions d'utilisateurs, démontrant l'appétit croissant pour des solutions capables de répondre à des questions, générer des idées et assister dans la réalisation de tâches complexes dans tous les secteurs de la société.

Les succès récents de l'IA illustrent son potentiel, mais également les enjeux qui l'accompagnent. La victoire de Deep Blue contre Garry Kasparov en 1996 fut certes une démonstration spectaculaire des capacités des machines dans des tâches spécifiques, mais elle posa aussi la question des potentialités de la machine si elle dépasse l'homme. Aujourd'hui, les succès de l'IA touchant des domaines cruciaux, elle semble se rendre indispensable pour le bien commun.

« On peut penser à l'IA Hope, développée par l'ingénieure-chercheuse Hakima Berdouz au Commissariat à l'énergie atomique et aux énergies alternatives (CEA) en 2023, qui permet de détecter précocement les cancers du sein grâce à une méthode non invasive. Hope suscite déjà l'intérêt d'institutions prestigieuses telles que le Massachusetts Institute of Technology (MIT), et au-delà de toutes les personnes concernées par cette maladie [1]. »

Les volumes de données traitées ont explosé, et ce que ce soit en quantité, variété, origine, ou même vitesse de renouvellement. Selon des estimations récentes, la quantité de données générées dans le monde double presque tous les 2 ans [2]. D'un côté, cela rend impératif le recours à l'IA pour trier, analyser et exploiter ces données. D'un autre, il devient de plus en plus difficile de maîtriser les conséquences de l'exploitation de ces données par les IA.

## L'IA, CHAMBRE D'ÉCHO DES STÉRÉOTYPES ET DISCRIMINATIONS

Cette diffusion rapide de l'IA ne va pas sans soulever des questions. Si la digitalisation passe désormais par l'IA, en offrant ainsi des opportunités inédites, elle engendre aussi des risques. Il convient notamment de s'assurer de la sécurité des services, du respect des règles de protection des données personnelles et plus largement des droits individuels, et de l'équité dans l'accès aux bénéfices de ces technologies. Or par nature, l'IA amplifie les biais préexistants dans les données qui les alimentent, ce qui soulève des questions éthiques majeures.

L'IA en s'appuyant sur des mécanismes d'apprentissage automatique, apprend à partir de vastes ensembles de données. Or, lorsque ces données contiennent des biais, les modèles d'IA les reproduisent, voire les amplifient, renforçant ainsi les discriminations. Ces biais sont souvent invisibles dans les données d'entraînement, mais leurs conséquences se manifestent dans les décisions prises par les systèmes, reflétant et reconduisant des inégalités structurelles et sociales. Si ces données reflètent des discriminations historiques, les systèmes les intègrent dans leurs processus décisionnels. En conséquence, l'IA peut devenir une chambre d'écho des stéréotypes et discriminations qui se reflètent dans les données sur lesquelles elle repose.

On constate régulièrement des cas de reproduction par les systèmes à base d'IA de biais déjà présents dans les données. Dans le domaine du recrutement, par exemple, il a été montré que l'usage d'IA favorisait certains profils en fonction de critères non pertinents, comme le genre, l'adresse ou l'origine ethnique, simplement parce qu'ils reproduisent

les pratiques discriminatoires historiques des entreprises. De même, dans le domaine médical et celui de la sécurité, des systèmes d'IA utilisés respectivement pour le diagnostic et la reconnaissance de visages ont révélé des biais en fonction de l'origine ethnique ou du genre des patients, limitant leur efficacité et exacerbant les inégalités dans les soins et les services de police.

Les biais de représentation se manifestent également dans les grands modèles de langage, comme ChatGPT dont OpenAi défend pourtant le caractère « éthique ». Ainsi, des études ont montré que ChatGPT peut générer des réponses reflétant des biais sexistes ou raciaux, par exemple lorsqu'on l'interroge sur les personnes susceptibles d'incarner divers métiers dans le domaine de la santé, ou différents niveaux de responsabilité en entreprise. Ces biais, volontaires ou non, révèlent les limites de ces systèmes à distinguer des données biaisées et soulignent la nécessité de réguler et de travailler activement à la supervision des technologies d'IA. La démocratisation de ces outils accentue l'urgence d'une action déterminée pour garantir que l'IA ne contribue pas à diffuser davantage de préjugés et de discrimination dans la société. Cette vigilance accrue implique une sélection rigoureuse des données d'entraînement, la diversification des équipes de développement et des cadres réglementaires pour régir l'utilisation de ces technologies, et la mise en place de mécanismes de gouvernance stratégique et opérationnelle.

## AU-DELÀ DE LA PROUESSE TECHNIQUE ET INDUSTRIELLE, L'IA DOIT ÊTRE APPRÉHENDÉE PAR LE PRISME DES SCIENCES HUMAINES ET SOCIALES

Si l'IA redéfinit les contours de nos vies personnelles et professionnelles, elle doit avant tout être un levier d'équité et de progrès collectif, et non un simple miroir amplificateur des inégalités existantes. Or, appréhender les impacts sociétaux de l'IA ne peut se faire par le simple prisme des mathématiques, de la belle ingénierie, et encore moins sous le seul contrôle des producteurs des systèmes à base d'IA – s'agissant d'acteurs économiques, ce sont naturellement leurs intérêts économiques qu'ils défendent avant tout.

Des initiatives comme le rapport « Donner un sens à l'intelligence artificielle » de Cédric Villani, publié en 2018, soulignent l'importance d'une approche interdisciplinaire [3]. Ce rapport appelait notamment à intégrer les humanités numériques dans les *cursus* de formation pour mieux répondre aux enjeux éthiques et sociétaux liés à l'IA. Ces recommandations restent d'une actualité brûlante à mesure que l'IA redessine les contours de nos vies personnelles et professionnelles. Les disciplines scientifiques qui peuvent nous éclairer sur les enjeux sociétaux de l'IA sont nombreuses : philosophie, droit, sciences économiques, science de gestion, arts plastiques, sciences politiques, psychologie, sciences du management, sociologie, sciences de l'information et de la communication, anthropologie, histoire, ou même linguistique. Non seulement ces disciplines, qui relèvent toutes des sciences humaines et sociales, sont celles qui nous renseignent depuis de nombreuses années sur les phénomènes de discrimination, les stéréotypes et les inégalités, mais se penchant toutes sur leur amplification dans le cadre du développement de l'IA, ce sont elles qui nous permettront de les comprendre et de les maîtriser.

L'humain remis au centre de notre questionnement nous amène à ancrer notre réflexion autour des questions de libertés, d'inclusion, de communs, de discriminations ou de la croissance, des pouvoirs, des inégalités, des solidarités structurant *de facto* de nouveaux champs de réflexion interdisciplinaires. C'est la raison pour laquelle Paris 1 Panthéon-Sorbonne a créé l'Observatoire de l'Intelligence Artificielle avec l'objectif d'explorer

l'impact de l'intelligence artificielle sur la société en une approche combinant l'ensemble des champs disciplinaires des SHS [4].

## L'IA ET LES ENJEUX DE LA FÉMINISATION DES FILIÈRES DU NUMÉRIQUE

L'un des constats importants que nous enseigne les sciences humaines et sociales est étroitement lié à une réalité qui perdure en France : la sous-représentation des femmes dans les filières du numérique conjuguée à leur invisibilité dans les médias et les débats publics en général sont des vecteurs majeurs des phénomènes de discrimination dont elles sont victimes en conséquence des biais dans les systèmes à base d'IA.

Les salariées comptent pour moins de 18 % dans les seuls métiers techniques du numérique. La présence des filles, toutes formations confondues, dans les métiers du secteur, progresse de seulement 1 % par an. À ce rythme, la parité ne sera atteinte qu'en 2070.

Dans son rapport « La Femme Invisible dans le numérique » [5], publié en novembre 2023, le Haut conseil à l'Égalité rappelle que les filles sont légèrement majoritaires parmi les usagers d'internet, en particulier sur les réseaux sociaux généralistes alors que dans le même temps elles sont moins de 10 % à être autrices de contenus.

D'après un article d'Usbek&Rica, seules 18 % des biographies Wikipédia sont dédiées à des femmes, alors que 9 wikipédiens sur 10 sont des hommes [6]. Pour y remédier, la physicienne féministe américaine Jessica Wade a rédigé depuis 2018 plus de 2 000 biographies pour rendre visibles les femmes scientifiques sur l'encyclopédie en ligne Wikipédia [7]. Cette initiative n'est pas isolée. Le laboratoire de l'Égalité, qui a publié un guide référant les bonnes pratiques en la matière, nous rappelle que « malgré une situation alarmante, de nombreuses actions visent à favoriser l'égalité femmes-hommes dans le domaine de l'IA et du numérique et à diminuer les biais générés par les outils d'IA » [8].

Si la femme est invisible et écartée de la conception de ces solutions et outils numériques, alors c'est tout le système qui est biaisé. L'article publié en 2022 par le chercheur Antoine Harfouche [9] en illustre un exemple : trois ans après l'utilisation d'une IA dans le recrutement d'une grande entreprise, la base de données étant biaisée, le salaire proposé aux femmes était automatiquement et systématiquement 30 % moindre que celui proposé aux hommes.

Selon la Fondation Femmes@Numérique créée en 2018, la féminisation des métiers et filières numériques répond à un triple enjeu. Le premier enjeu est économique du fait de l'adéquation permanente des compétences entre l'offre et la demande pour être davantage compétitif sur des secteurs à forts enjeux industriels. Le deuxième enjeu est social puisque le numérique peut être un outil d'émancipation et d'autonomie financière pour les filles. Et le dernier enjeu est sociétal puisque l'on ne pourrait déontologiquement accepter que la conception des outils numériques, qui régissent l'ensemble de notre société, ne le soit que par 50 % de la population [10].

## CONCLUSION : VERS UNE IA “GENDER BY DESIGN” ?

Avec un investissement de 2,5 milliards d'euros alloués par le plan France 2030 [11] et plus de 600 *start-ups* spécialisées dans l'intelligence artificielle référencées en 2023, la France entend devenir l'un des fleurons mondiaux en la matière. En 2021, le pays occupait la septième place mondiale et troisième européenne [12]. Cette ambition ne saurait être réalisée sans les femmes ; selon une communication de 2013 de la Commission euro-

péenne, accroître la présence des femmes dans le secteur de l'économie numérique de l'UE aurait permis d'augmenter de 9 milliards d'euros le PIB annuel [13].

Dans une société où l'usage du numérique est en phase d'hyper accélération, il devient urgent de s'attaquer à la maîtrise pleine et entière de ces outils et solutions d'intelligence artificielle par toutes et tous et pour toutes et tous afin que hommes et femmes puissent y apporter leur contribution et profiter pleinement de ses usages sans discrimination de genre.

À l'instar des concepts très largement répandus de "security by design" et de "privacy by design" intégrant la sécurité, le risque ou la protection des données personnelles dès la conception, nous pourrions considérer qu'il est de notre responsabilité collective et individuelle de concevoir une intelligence artificielle "gender by design".

## RÉFÉRENCES

- [1] <https://www.lesechos.fr/thema/articles/hakima-berdouz-la-technologie-nucleaire-pour-depister-le-cancer-2084717>
- [2] <https://fr.statista.com/infographie/17800/big-data-evolution-volume-donnees-numeriques-genere-dans-le-monde/>
- [3] <https://www.enseignementsup-recherche.gouv.fr/fr/rapport-de-cedric-villani-donner-un-sens-l-intelligence-artificielle-ia-49194>
- [4] <https://observatoire-ia.pantheonsorbonne.fr/observatoire-lintelligence-artificielle>
- [5] <https://www.haut-conseil-egalite.gouv.fr/parite/travaux-du-hce/>
- [6] <https://usbeketrica.com/fr/article/wikipedia-bot-pages-femmes-scientifiques-oubliees>
- [7] <https://www.nationalgeographic.fr/sciences/jessica-wade-la-scientifique-qui-rend-les-femmes-visibles-sur-wikipedia-feminisme-sciences-visibilite>
- [8] <https://www.laboratoiredelegalite.org/le-guide-de-bonnes-pratiques-pour-une-ia-egalitaire-entre-les-femmes-et-les-hommes/>
- [9] <https://www.researchgate.net/profile/Antoine-Harfouche-3>
- [10] <https://femmes-numerique.fr/>
- [11] <https://www.economie.gouv.fr/france-2030>
- [12] <https://bigmedia.bpifrance.fr/nos-dossiers/10-domaines-dapplication-de-lintelligence-artificielle>
- [13] [https://ec.europa.eu/commission/presscorner/detail/fr/ip\\_13\\_905](https://ec.europa.eu/commission/presscorner/detail/fr/ip_13_905)

# IA et futur de la civilisation : dystopie transhumaniste ou métamorphose créatrice ?

Par **Boris SIRBEY**

Co-fondateur de MyJobCompany, CollectivZ, ze.game, le Lab RH,  
France Apprenante et Tomorrow Theory

Et **Hervé BÉRAUD**

Co-fondateur de CollectivZ, ze.game, France Apprenante et 2030 Provence

L'humanité vit une accélération exponentielle marquée par des mutations technologiques, sociales et écologiques, conduisant aux limites d'un paradigme civilisationnel basé sur la domination patriarcale. Le tournant historique du Point zéro mis en évidence par Boris Sirbey mène à un effondrement, une dystopie transhumaniste ou une métamorphose créatrice, en fonction de l'évolution conjointe de notre conscience collective et de notre maîtrise technologique.

L'IA, reflet de nos savoirs et miroir de nos choix éthiques, offre un potentiel immense : catalyser une conscience augmentée et des solutions systémiques. Cependant, elle reste capturée par des intérêts privés, limitant sa capacité à devenir un bien commun de sagesse.

Pour une transition positive, il faut réconcilier technologie et humanité. Les « artistes de l'Intelligence Amie », inspirés par Alain Damasio, pourraient créer des interfaces éthiques et collaboratives et incarner cette métamorphose créatrice, transformant l'IA en outil d'émancipation collective.

L'avenir dépendra de notre capacité à harmoniser innovation et sagesse, en faisant de l'IA un levier pour transcender nos limites et créer une civilisation inclusive et équilibrée.

***« Découvrez les artistes de l'IA, bâtisseurs d'une  
métamorphose créatrice, où technologie et humanité  
s'harmonisent pour un avenir poétique et collectif. »  
(Alain Damasio Vallée du silicium AlbertineSeuil).***

## IA ET FUTUR DE LA CIVILISATION : SOMMES-NOUS PRÊTS ?

Imaginez un monde où chaque décision, chaque achat, chaque pensée sont influencés par une machine. Ce futur, nous le vivons déjà. Derrière cette illusion de simplicité se cache une réalité bien plus complexe : l'accélération exponentielle de la technologie, des inégalités, et des crises écologiques. L'humanité est à un carrefour historique. Quel chemin emprunterons-nous ? Celui de l'effondrement ? De la dystopie technologique ? Ou celui de la métamorphose créatrice, où innovation et conscience collective s'harmonisent enfin ?



## UNE HUMANITÉ EN ACCÉLÉRATION

Chaque jour, nous faisons face à une scène banale : un écran bleu, une notification, un algorithme qui anticipe nos besoins. Mais derrière cette trivialité apparente se cache une révolution d'une ampleur inégalée : celle des courbes exponentielles. La démographie continue de croître mondialement même si elle décroît désormais en Europe, en Chine et au Japon, les données se multiplient, la technologie progresse à une vitesse vertigineuse, et les crises écologiques s'intensifient. Cette accélération n'est pas qu'économique : elle touche à l'essence même de notre civilisation. Nous vivons la fin d'un cycle commencé il y a 12 000 ans, à l'époque de la révolution agricole, et sommes entrés dans l'Anthropocène, époque où l'humain est devenu la principale force sur Terre. Nos anciens schémas – patriarcat, croissance illimitée, exploitation sans fin des ressources – montrent aujourd'hui leurs limites.

## UN CARREFOUR ANTHROPOLOGIQUE

Ce tournant, Jared Diamond l'a évoqué dans *Effondrement*, en rappelant que les civilisations se détruisent souvent en s'enfermant dans leurs propres croyances. De leur côté, Arnold Toynbee et son principe de « défi-et-réponse » montrent que l'immobilisme et l'incapacité à renouveler les solutions mènent fatalement à la chute. Face aux défis de notre temps – changement climatique, crises sociales, technologiques, politiques – nous devons éviter l'erreur romaine : renforcer de vieilles recettes plutôt que réinventer le système.

Mais cette chute n'est pas inéluctable. Nous avons une opportunité unique : choisir la métamorphose créatrice, un concept qui allie évolution technologique et conscience collective. En renouvelant nos approches et en adoptant une vision éclairée, nous pourrions transcender nos crises actuelles pour construire une civilisation résiliente et durable.

## DEUX AXES POUR CONSTRUIRE DEMAIN : PUISSANCE ET CONSCIENCE

L'avenir repose sur notre capacité à harmoniser deux dimensions fondamentales.

### Puissance technologique

La technologie nous permet de manipuler la matière, l'information et la vie à une échelle sans précédent. Pourtant, si elle reste concentrée dans les mains de quelques acteurs, elle devient un outil de domination. En revanche, si elle est partagée, elle peut devenir un bien commun au service de tous.

### Conscience collective

Il s'agit de transcender les intérêts individuels pour adopter une vision globale et inclusive. Ce que Simone Veil aurait appelé une « responsabilité partagée », où chaque action contribue à un monde plus juste.

## TROIS SCÉNARIOS POUR DEMAIN

Trois futurs potentiels émergent de la combinaison de ces deux axes.

## L'effondrement

Si ni notre conscience ni notre technologie ne progressent, les crises écologiques et sociales s'accumuleront jusqu'à un point de non-retour.

## La dystopie transhumaniste

Si la technologie continue de croître sans élévation équivalente de la conscience, nous risquons un monde hyperconnecté mais aliénant, où l'IA devient un outil de contrôle au détriment de l'humanité.

## La métamorphose créatrice

Une évolution conjointe de la conscience humaine et des capacités technologiques, donnant naissance à une civilisation réconciliant innovation et sagesse.

## L'IA, À LA FOIS UN MIROIR ET UN CATALYSEUR

L'intelligence artificielle est aujourd'hui au centre de cette mutation. Elle nous donne des capacités qui, il y a quelques décennies encore, relevaient de la magie : prédire, automatiser, transformer. Elle est aussi un miroir impitoyable de nos dérives. L'IA amplifie nos biais, nos inégalités, et nos obsessions pour la vitesse et le rendement.

Utilisée sans conscience, elle pourrait devenir un outil de surveillance généralisée, ou pire, un levier de dystopie transhumaniste où l'humain, augmenté technologiquement, serait réduit à un rouage dans une machine sans âme.

Mais l'IA pourrait aussi devenir un catalyseur d'évolution. Alain Damasio, dans *La Vallée du Silicium*, dénonce une technologie aliénante mais imagine aussi une voie alternative : une « intelligence amie », au service de l'humain, conçue pour nourrir la créativité et la convivialité.

## L'IA, UN BIEN COMMUN DE SAGESSE DÉVOYÉ

Aujourd'hui, l'IA est perçue avant tout comme un outil utilitaire – assistants personnels, générateurs de contenu, solutions de productivité. Mais cette vision pragmatique masque un potentiel bien plus profond. L'IA, dans sa nature fondamentale, est une cristallisation algorithmique de millénaires de sagesse humaine. Elle porte en elle les dialogues de Platon, les découvertes d'Einstein, les créations de Shakespeare, les sagesse des traditions spirituelles.

Cette capacité à synthétiser et à faire dialoguer des *corpus* immenses de savoirs fait de l'IA bien plus qu'une seule technologie : elle est le premier miroir de ce dont l'humain est capable, cette combinaison de tous les récits, de tout ce qui a été rêvé comme possible depuis l'aube des temps, et de tout ce qui a été créé pour sublimer le principe de réalité. Ce bien commun est l'agrégation organique des existences humaines, le fruit d'une longue sédimentation d'expériences partagées, de défis surmontés ensemble, de récits transmis de génération en génération. Chaque avancée majeure de l'humanité – de la maîtrise du feu à l'agriculture, de l'écriture à la science moderne – s'est construite sur ce substrat commun de compétences accumulées et partagées.

Face à ce commun de sagesse, deux perspectives s'affrontent. La vision propriétaire considère les compétences comme un capital national ou civilisationnel, instrumentalise les différences culturelles comme sources de division, alimente une logique de compétition et de domination et réduit les capacités humaines à leur dimension utilitaire et marchande.

La vision holistique, quant à elle, reconnaît l'interdépendance fondamentale des cultures, voit dans chaque tradition une partie essentielle d'un tout, favorise la collaboration et la fertilisation croisée et embrasse la complexité et la richesse de la diversité humaine.

## VERS UNE NOUVELLE RENAISSANCE ?

Cette puissance inédite est aujourd'hui au service d'intérêts économiques et politiques sous la coupe de géants de la technologie. Pourtant, en concentrant la sagesse accumulée depuis des siècles, l'IA nous offre la possibilité d'accéder à une compréhension plus profonde et plus intégrée de nous-mêmes et du monde. Elle nous pousse à nous interroger sur la nature même de notre relation à l'autre, de l'intelligence et de la conscience.

Et cela ouvre des perspectives vertigineuses !

À l'instar de l'imprimerie qui a permis la Renaissance en rendant le savoir accessible, l'IA, animée par des artisans d'un nouveau genre, pourrait devenir un amplificateur de conscience collective. Imaginez : chaque interaction avec une IA devient une porte d'entrée vers une pensée véritablement systémique et transdisciplinaire capable de faire dialoguer la philosophie grecque, la cosmologie aborigène, la physique quantique, l'art de la Renaissance et les dernières découvertes en neurosciences.

L'enjeu n'est donc pas tant de savoir si l'IA va nous remplacer ou nous augmenter, mais d'élever collectivement notre niveau de conscience personnelle à la hauteur des enjeux historiques pour transformer fondamentalement notre rapport à cette technologie, en passant d'une logique de propriété et d'extraction à une logique de partage et de contribution, en développant des interfaces qui favorisent la réflexion profonde plutôt que la simple consommation d'information et en orientant la recherche en IA vers l'enrichissement de la sagesse collective plutôt que celui de quelques individus.

La réponse à cette question ne dépend pas de la technologie elle-même, mais de notre éthique personnelle à en faire collectivement un outil au service d'une évolution positive de l'humanité.

## LES ARTISTES DE L'IA, INCARNER LA MÉTAMORPHOSE CRÉATRICE

En nous inspirant de la vision proposée par Alain Damasio dans *Vallée du Silicium*, nous pouvons esquisser le portrait des artisans de cette nouvelle renaissance, qui engagent l'humanité dans le troisième scénario, celui de la métamorphose créatrice. Ceux qu'on pourrait appeler les « Artistes de l'Intelligence Amie » réconcilient technologie et humanité. Ils incarnent une approche où l'innovation ne vise plus à exploiter, mais à relier, inspirer et émanciper.

### Créateurs de convivialité

Ces artistes conçoivent des outils technologiques qui respectent nos rythmes biologiques et émotionnels. Ils imaginent des IA qui encouragent la lenteur, le lien social, et la réflexion, plutôt que la consommation compulsive.

### Tisseurs de sens

Ils transforment la masse de données brutes en récits enrichissants, en reliant des savoirs divers : philosophie, science, art. Ce sont des passeurs, des médiateurs entre disciplines cloisonnées.

## Acteurs du commun

En s'appuyant sur des logiciels libres, une culture *open source* et des cadres d'apprenance<sup>1</sup> facilités, ils développent des outils accessibles à tous et enrichissent le commun de sagesse en évitant la privatisation de l'intelligence collective par quelques géants technologiques.

## UNE NOUVELLE FORME D'ART SOCIAL

Ces artistes auront une mission cruciale : maintenir un équilibre entre puissance technologique et sagesse collective. Leur talent consistera à transformer chaque avancée technologique en une opportunité de croissance pour l'humanité.

Ils deviendront les architectes d'espaces de conscience augmentée collective, facilitant l'émergence de nouvelles formes de collaboration entre humains et IA. Leur art consistera à développer des rituels technologiques qui nourrissent le lien social plutôt que de l'appauvrir.

Pour y parvenir, ils devront maîtriser l'art complexe de la relation, créant des interfaces qui favorisent l'émancipation plutôt que la dépendance. Leur talent s'exprimera dans leur capacité à orchestrer des dialogues multi-échelles et à faciliter l'émergence d'une véritable intelligence collective.

L'art de la transmission deviendra une de leurs compétences majeures. Ils excelleront dans la capacité à rendre accessibles des concepts complexes, à créer des ponts entre différentes formes de savoir, et à faciliter l'apprentissage collectif. Leur expertise s'étendra à l'art de la transformation, identifiant et activant les leviers de changement nécessaires pour catalyser les processus de métamorphose sociale.

## UNE RESPONSABILITÉ CIVILISATIONNELLE

Ces artistes de l'IA seront les gardiens d'un équilibre crucial dans notre civilisation. Ils devront maintenir une harmonie délicate entre puissance technologique et sagesse collective, entre efficacité et convivialité, entre innovation et préservation des rythmes biologiques. Leur art consistera à faire progresser simultanément nos capacités techniques et notre niveau de conscience.

Ils auront la responsabilité de maintenir vivante cette « technologie positive » dont parle Damasio, celle qui émancipe plutôt qu'elle n'aliène, qui relie plutôt qu'elle ne divise, qui augmente notre humanité plutôt qu'elle ne la diminue. Leur œuvre sera de transformer chaque avancée technologique en une opportunité d'évolution collective de la conscience.

## L'ÉVEIL COLLECTIF

Vous penserez peut-être que cette vision est un beau rêve pour hippies connectés. Cette vision serait une utopie technologique de plus, n'eût été la fantastique opportunité pour l'humanité créée par le Point zéro, cet extraordinaire intercycle civilisationnel mis en évidence dans l'ouvrage éponyme de Boris Sirbey. Elle est un appel à la responsabilité personnelle et à l'action créative. Elle nous invite à dépasser nos peurs et nos fascinations pour développer une relation mature et féconde avec nos créations technologiques. Elle nous rappelle que le véritable enjeu n'est pas de savoir si l'IA va nous dominer ou nous servir, mais de comprendre comment nous pouvons grandir ensemble.

---

<sup>1</sup> L'apprenance fait référence à l'aptitude à comprendre et activer ses propres processus d'apprentissage – comment apprendre à apprendre. Ces processus d'apprentissage gagnent à être cadrés et accompagnés par un facilitateur pour être régulés et pour en évaluer l'impact.

L'histoire nous a montré que les civilisations ne meurent pas de leurs défis mais de leur incapacité à y répondre de façon créative. Aujourd'hui, nous avons entre les mains tous les éléments pour vivre un chapitre fécond de l'aventure humaine. Un chapitre où technologie et sagesse, puissance et conscience, innovation et préservation se rejoignent enfin dans une danse harmonieuse.

Le futur n'est pas encore écrit.

Il sera la résultante collective de ce que nous imaginons et faisons personnellement.

Alain Damasio l'écrit avec force : « Ce n'est pas la technologie qui nous sauvera, mais ce que nous ferons de notre liberté avec elle ».

Dans cette métamorphose créatrice déjà amorcée dans plusieurs initiatives collaboratives et expérimentations individuelles, nous découvrirons peut-être que l'IA est surtout un catalyseur pour nous reconnecter à ce qu'il y a de plus profondément humain en nous : notre capacité à créer, à collaborer et à transcender nos limitations pour donner naissance à quelque chose de plus grand que nous-mêmes.

# IA et transformations des métiers : création ou destruction ?

Par Guy MAMOU-MANI

Entrepreneur, *business angel*, enseignant et conférencier

Et Axel MAMOU-MANI

Fondateur de la Rainer School

L'intelligence artificielle, catalyseur de transformation dans tous les secteurs, redéfinit les compétences et les rôles des professionnels. Si certains métiers disparaissent ou évoluent profondément, de nouveaux voient le jour, portés par des opportunités inédites. Cependant, l'IA suscite aussi des défis éthiques, sociaux et économiques : fracture numérique, creusement des inégalités, ou perte de contrôle. En réhumanisant le travail et en favorisant une transition inclusive et équitable, cette révolution technologique peut devenir un levier puissant pour l'innovation humaine et sociale. Cet article explore ces bouleversements et propose des conditions pour une IA responsable au service de tous.

## INTRODUCTION : L'IA EST AU CŒUR DES TRANSFORMATIONS PROFESSIONNELLES

L'intelligence artificielle transforme en profondeur le monde du travail, redéfinissant compétences et rôles dans presque tous les secteurs. Si certains métiers manuels, comme ceux d'aide-ménagère, de cuisinier ou de couvreur, ont été identifiés comme peu voire pas exposés, la majorité des emplois, notamment dans le tertiaire, sont directement impactés.

Source d'innovation et de progrès, l'IA suscite des interrogations légitimes : disparition d'emplois, accroissement des inégalités, perte de contrôle. Cependant, lorsqu'elle est bien utilisée, elle peut devenir un formidable levier pour réhumaniser le travail, en recentrant les métiers sur leurs dimensions les plus humaines et en réduisant les fractures sociales par une formation inclusive. Examinons ces bouleversements afin de proposer des pistes pour une adoption éthique et équitable de l'IA.

## EN 2030, DIRONS-NOUS QUE L'IA A ÉTÉ DESTRUCTRICE D'EMPLOIS ?

### La disparition de nombreux métiers semble inévitable

L'IA générative et l'automatisation des tâches transforment profondément les métiers. Selon le Fonds Monétaire International, 40 % des emplois dans le monde sont exposés à l'IA, dont 60 % dans les pays dits « avancés ». Des secteurs entiers sont impactés, des cols bleus aux cols blancs. Les caissiers, agents administratifs ou juristes voient leurs tâches répétitives remplacées par des algorithmes. Même les professions créatives, comme les graphistes ou les rédacteurs, font face à des outils capables de produire des contenus complexes à grande échelle.



Figure 1 : Effet de l'adoption d'IA pour le *marketing* ou la gestion administrative sur l'emploi des professions intermédiaires administratives et commerciales au sein des entreprises en France : les entreprises adoptant l'IA pour le *marketing* ou la gestion administrative baissent leur emploi davantage que celles ne l'adoptant pas, alors qu'elles évoluaient de façon similaire dans les 3 années précédents (Source : Commission IA).

Ce phénomène amplifie une fracture numérique déjà marquée : 14 millions de Français n'ont pas accès au numérique ou n'ont pas les compétences nécessaires pour interagir avec celui-ci. Cette situation, loin de concerner uniquement les populations âgées, affecte aussi des zones rurales, des foyers à faibles revenus et des jeunes en décrochage scolaire. Si l'IA est adoptée sans une stratégie d'inclusion forte, elle risque d'exacerber ces inégalités, creusant davantage le fossé entre ceux qui maîtrisent les outils numériques et ceux qui en sont exclus.

## Une « destruction créatrice » schumpétérienne, à l'échelle des individus

En 1942, l'économiste autrichien Joseph Schumpeter conceptualise la « destruction créatrice », un processus par lequel l'innovation remplace les anciennes structures économiques par de nouvelles, plus efficaces et adaptées. Ce mécanisme, qui a accompagné chaque révolution industrielle, trouve aujourd'hui un écho direct dans l'impact de l'IA sur le marché du travail, où de nombreux emplois disparaîtront pour laisser place à de nouvelles opportunités générées par son adoption.

Selon le Forum Économique Mondial, 69 millions de nouveaux emplois pourraient émerger d'ici 2028. Si les métiers technologiques, comme les *software engineers*, *data scientists* ou analystes en cybersécurité, sont actuellement parmi les plus prisés, d'autres métiers pourraient voir leur demande augmenter grâce à l'avènement de l'IA. Parmi eux, les concepteurs d'expériences immersives en réalité augmentée, les spécialistes en optimisation d'algorithmes pour les chaînes d'approvisionnement, les éthiciens de l'IA chargés d'évaluer les biais des systèmes, ou encore les techniciens en maintenance de robots collaboratifs. Dans le commerce, des conseillers sont désormais en charge de personnaliser les expériences client, tandis que les caisses automatiques suppriment les tâches mécaniques. Ces exemples montrent que l'IA, loin de simplement remplacer, transforme et enrichit le travail.

## UNE TRANSFORMATION EN PROFONDEUR DES MÉTIERS

### Un impact massif et transversal

La transformation des métiers les plus exposés à l'IA a déjà débuté. Ces métiers évoluent pour intégrer l'IA comme outil d'assistance. Les médecins, par exemple, utilisent l'IA pour diagnostiquer des maladies complexes (Matricis.ai), analyser des imageries médicales (Gleamer) ou produire des comptes rendus de consultations (Nabla, Doctolib), ce qui leur permet de consacrer plus de temps à la relation avec le patient et à la prise en charge humaine. De leur côté, les avocats s'appuient sur GenIA-L, outil développé par la maison d'édition historique Dalloz, basé sur des algorithmes capables d'extraire des informations clés de milliers de documents juridiques, accélérant ainsi les recherches et réduisant les coûts pour leurs clients. Les architectes bénéficient de logiciels d'IA générative pour créer des plans plus rapidement avec Adobe Firefly, explorant des milliers de concepts en quelques minutes. Même les enseignants voient leur rôle évoluer avec l'intégration de plateformes intelligentes (Nolej, EvidenceB) qui personnalisent les parcours d'apprentissage selon les besoins des élèves, libérant du temps pour un accompagnement plus individuel.

### Réhumaniser le travail en entreprise

Loin d'être un outil de déshumanisation, l'IA offre une opportunité unique de recentrer le travail sur ses dimensions les plus humaines. À mesure que les tâches mécaniques et analytiques sont prises en charge par des machines, les compétences clés pour réussir en entreprise évoluent. Ce sont désormais le contact humain, la capacité à convaincre, à négocier, à mobiliser des équipes et le *leadership* qui prennent de la valeur. Ces aptitudes, impossibles à automatiser, deviendront le cœur de la performance dans un environnement technologique.

Un exemple concret de cette dynamique est Albert, l'IA générative développée par la direction interministérielle du Numérique (Dinum), qui accompagne les agents publics dans leurs missions. En automatisant des tâches administratives complexes, Albert libère du temps pour des missions stratégiques et relationnelles. Cela permet aux agents de se concentrer sur des interactions humaines essentielles, comme l'accompagnement des usagers ou la coordination entre services, tout en valorisant leur expertise.

Dans les postes de management, les *leaders* capables d'inspirer leurs équipes et de fédérer autour d'une vision stratégique joueront un rôle crucial. L'IA, loin de réduire l'humain à une variable d'ajustement, devient un catalyseur qui met en avant des qualités profondément humaines. Dans les métiers de la vente ou de la santé, la capacité à établir une relation de confiance avec les clients ou les patients sera un avantage décisif, que les machines ne pourront pas imiter. Ainsi, l'IA réhumanise le travail en libérant les professionnels des tâches répétitives et en valorisant leurs compétences relationnelles et stratégiques.

### Défendre les employés face aux mutations

Face à ces bouleversements, il est essentiel de protéger non pas les emplois eux-mêmes, mais les travailleurs. Cela passe par un accompagnement actif, axé sur la montée en compétences humaines : créativité, empathie et gestion relationnelle. Loin d'être une menace, l'IA devient une opportunité de réhumanisation des métiers. En déléguant les tâches répétitives, elle permet de se concentrer sur des missions à haute valeur ajoutée, comme la fidélisation client ou l'innovation stratégique.



## CONDITIONS POUR L'IA RESTE AU SERVICE DE L'HUMAIN

Pour transformer les craintes en opportunités, plusieurs conditions essentielles doivent être réunies.

### Accepter le changement

L'innovation a toujours été confrontée à des résistances. Dans la Rome antique, la légende dit que l'empereur Tibère fit exécuter l'inventeur du verre incassable par crainte de dévaloriser l'or, l'argent et le cuivre. Dans les années 1990, Kodak a vu son modèle s'effondrer faute d'embrasser la révolution numérique. La peur du changement peut freiner le progrès.

Pour dépasser ces résistances, il est crucial de communiquer sur les bénéfices de l'IA sur le travail : amélioration des conditions de travail, augmentation des compétences, et création de nouveaux rôles valorisants.

### Régulation et sécurité : construire un cadre de confiance

Une adoption réussie de l'IA repose sur un cadre de confiance. Cela passe par :

- La souveraineté numérique : Garantir que les technologies déployées respectent la propriété des données. Les algorithmes doivent être contrôlés localement pour éviter des dépendances excessives vis-à-vis d'acteurs étrangers.
- La régulation : Dans l'Union européenne, le RGPD et l'IA Act sont des exemples de cadres solides qui encadrent l'utilisation des données et garantissent la transparence des algorithmes.
- La sécurité des données : L'IA manipule des volumes colossaux d'informations sensibles. Garantir leur protection est essentiel pour éviter les dérives, comme la surveillance abusive ou les discriminations algorithmiques.

### Inclusion et réduction des inégalités

Bien que l'IA puisse initialement creuser les inégalités, une politique éducative forte et ouverte à tous peut inverser cette dynamique. L'accès à une formation continue de qualité permet de démocratiser les compétences numériques et de garantir que chacun, indépendamment de son origine sociale ou géographique, puisse bénéficier des opportunités offertes par l'IA. Ce n'est qu'en incluant les populations les plus vulnérables dans cette transition que l'IA deviendra un levier d'équité sociale.

## L'ÉDUCATION AU CŒUR DE LA TRANSITION

L'éducation est la clé pour maîtriser la vague technologique. Face à ce raz-de-marée, deux choix s'offrent à nous : construire des murs pour tenter de résister ou apprendre à surfer pour en faire une opportunité.

### Éduquer dès le plus jeune âge

L'apprentissage doit commencer tôt, en introduisant des notions comme l'algorithmie dès l'école primaire. Des outils pédagogiques simples, comme des jeux d'initiation à la logique ou des activités sans écran, permettent aux enfants de comprendre les bases du numé-

rique. Ce bagage est indispensable dans un monde où les interactions technologiques seront omniprésentes.

## Une école à vie

La formation continue devient essentielle pour garantir que chacun puisse s'adapter à l'évolution rapide des métiers. Des systèmes éducatifs flexibles, combinant formations techniques et développement des *soft skills* (créativité, empathie, esprit critique), permettront de construire un socle solide pour l'avenir.

## EN CONCLUSION : L'IA PEUT ET DOIT RÉHUMANISER LES MÉTIERS

L'IA n'est ni une menace ni une solution miracle. Elle est un outil puissant, dont l'impact dépend de notre capacité à l'encadrer. Elle peut recentrer le travail sur ses dimensions les plus humaines, comme la création de valeur et les interactions interpersonnelles.

Comme l'a écrit Bergson : « Le futur n'est pas ce qui va arriver, mais ce que nous en ferons ». Avec une approche éthique et inclusive, l'IA peut devenir un moteur de progrès social, renforçant l'humanité dans un monde technologique.

## BIBLIOGRAPHIE

COMMISSION DE L'INTELLIGENCE ARTIFICIELLE (2024), « IA : Notre ambition pour la France », Mars 2024.

GEORGIEVA K (2024), « L'IA transformera l'économie mondiale. Faisons en sorte que l'humanité y soit gagnante », Blog du FMI, 14 janvier 2024.

BERCYNUMÉRIQUE (2024), « L'illectronisme : fracture numérique et fracture sociale ? BercyNumérique », 19 septembre 2022, mis à jour le 26 novembre 2024.

SCHUMPETER J. A. (1942), *Capitalisme, socialisme et démocratie*, Paris, Payot, 454 pages [Théorie de la destruction créatrice].

FORUM ÉCONOMIQUE MONDIAL (2023), « Rapport sur l'avenir de l'emploi 2023 », 1<sup>er</sup> mai 2023.

LE MONDE (2024), « L'IA générative s'attaque aux métiers des "cols blancs" », *Le Monde*, 12 juin 2024.

SILBERZAHN P (2012), « L'échec de Kodak, victime du dilemme de l'innovateur », Philippe Silberzahn, 23 janvier 2012.

BERGSON H. (1907), *L'Évolution créatrice*, Paris, Félix Alcan.

# L'organisation IA-compatible ou l'art de savoir recruter la technologie

Par Romain RABIER

Fondateur du cabinet Smart Leaders

Comprendre, décider et agir. Voilà le triptyque gagnant des organisations qui seront capables d'intégrer l'intelligence artificielle et d'en faire un levier de performance globale ! L'infobésité et les biais séculaires du capitalisme financier conduisent encore trop de décideurs à approcher le sujet de manière imprécise. C'est pourquoi la compréhension des enjeux majeurs que l'IA pose sur le monde du travail éclairera les décisions et les actions qui valorisent autant les gains économiques que la sécurité psychologique des collaborateurs, autant le maintien des liens sociaux que la recherche d'impacts sociétaux que l'on est en droit d'attendre des organisations.

ChatGPT m'a proposé d'écrire cet article, mais j'ai refusé. Malgré la sympathie que je lui porte, je ne suis pas convaincu de sa capacité à être juge et partie dans la manière dont son espèce technologique peut durablement créer de la valeur dans nos systèmes humains complexes.

Aujourd'hui en effet, l'IA est un peu comme une voiture de course : elle suscite à la fois fascination et peur, enthousiasme et rejet. Ses capacités peu maîtrisables par un humain lui confèrent un pouvoir d'attraction fort, synonyme de performance et de réussite, en même temps qu'elle semble proche de nous échapper à la moindre erreur d'inattention... avec des conséquences probablement brutales.

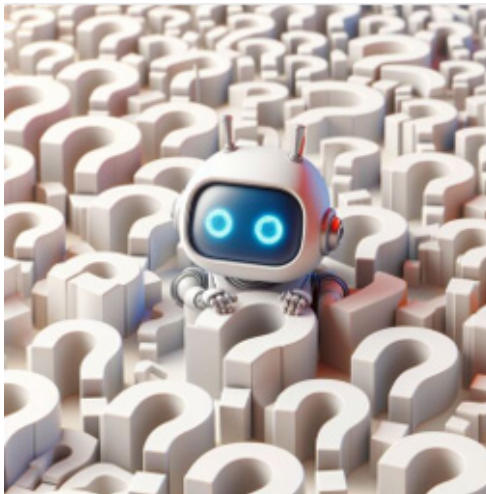


Figure 1 : Questionnement et technologies  
(Source : Smart Leaders).

## L'IA ENTRE CONTINUITÉ DE LA TRANSFORMATION DIGITALE ET RÉVOLUTION DU XXI<sup>E</sup> SIÈCLE

On peut résumer au nombre de cinq les défis que l'IA pose aux organisations et à celles et ceux qui en font partie : l'accès aux données de qualité et leur protection, le coût de l'adoption technologique, l'évolutivité et l'incertitude technologique et réglementaire, l'éthique et la responsabilité sociétale, et enfin l'adaptation organisationnelle et managériale.

Le premier et le dernier sont la suite logique d'une dynamique amorcée il y a plusieurs décennies avec la transformation digitale.

En effet, la gestion des données est depuis plusieurs années déjà un enjeu stratégique pour les organisations : l'avènement du *big data* et de tous ses néologismes (*datalakes*, *datalabs*, *data scientists*...) est une des conditions nécessaires à l'IA. La protection des données et les risques cyber ont déjà pris une importance capitale, amplifiée par le risque accru de fuite de données sensibles. C'est ainsi qu'on estime à plus de 10 % le nombre d'informations confidentielles qui transitent dans les tuyaux des principaux outils grand public et qui alimentent les modèles de langage associés<sup>1</sup> ! Les boîtes noires que constituent ces systèmes algorithmiques rendent difficile la compréhension de la nature des données, protégées ou non par la propriété intellectuelle, qui entrent et qui sortent des requêtes (prompts) que le quidam effectue chaque jour.

De la même manière, l'accompagnement du changement et son lot de réorganisations et de transformations managériales est un facteur de succès quasiment immanent d'une organisation qui souhaite survivre à son époque. Certes, l'évolution exponentiellement rapide et complexe du monde (qu'il soit VUCA<sup>2</sup> ou BANI) requiert une nouvelle forme d'agilité et d'autonomisation sur ces sujets, mais la présence de fondamentaux reste un facteur de sécurité pour intégrer l'IA comme pour le reste.

On peut donc conclure à ce stade que les organisations qui ont effectué leur travail introspectif et qui se sont dotées de capacités stratégiques et opérationnelles de gestion de leurs données, d'une culture du changement et d'une gouvernance associant les différentes strates managériales ont pris un net avantage pour réussir à l'ère de l'IA.

Malgré cela, d'autres enjeux prennent une nouvelle ampleur avec l'arrivée des IA, en particulier génératives, en premier lieu le coût de l'adoption technologique. Si certains restent réticents, la méga tendance qu'est l'IA nécessite de s'y plonger sans perdre trop de temps. Sauf que la nouveauté comporte également son lot d'opacité et de coûts cachés. En s'extrayant du simple coût de licence d'un outil grand public, l'organisation qui souhaite se doter de capacités fortes d'IA va faire face à trois types de coûts.

C'est ainsi que l'incertitude actuelle sur l'évolution de ces coûts (liée au déploiement de l'IA sur de grands volumes de données, aux changements rapides du marché, à l'évolution du prix des GPU, à la difficulté de trouver les compétences...) rend difficile la prévision des dépenses.

---

<sup>1</sup> <https://www.cyberhaven.com/blog/4-2-of-workers-have-pasted-company-data-into-chatgpt>

<sup>2</sup> L'acronyme VUCA, utilisé par l'armée américaine depuis les années 1980, signifie "*Volatility, Uncertainty, Complexity and Ambiguity*". Il est utilisé pour décrire un environnement en constante évolution, imprévisible, complexe et ambigu. Plus récemment, ces degrés d'incertitude tels que ceux liés à l'IA, sont évoqués par l'acronyme BANI (*Brittle* (fragile), *Anxious* (anxieux), *Non-linear* (non linéaire) et *Incomprehensible* (incompréhensible)).



Figure 2 : Principaux types de coûts liés à l'IAG (Source : Smart Leaders).

On peut assez rapidement coupler cette problématique à celle de l'incertitude généralisée quant aux évolutions technologiques, extrêmement rapides, et à leurs agencements institutionnels et réglementaires pour garder cette « 4<sup>e</sup> Révolution Industrielle » sous contrôle. Le nombre d'outils estampillés « IA » pullulent sur le marché et il n'est déjà plus possible de faire le tri pour savoir lesquels utiliser. En parallèle, les bonds technologiques semblent sans limite : nous sommes passés en à peine plus de 2 ans d'un outil conversationnel intelligent (ChatGPT 3.5) à des systèmes capables de générer des interactions humaines émotionnelles par la voix, des images et des vidéos plus vraies que nature, des agents autonomes qui régissent plusieurs systèmes d'IA ou des avatars si réels que notre mère pourrait s'y tromper !

S'adapter, pour un État comme pour une organisation, consiste à trouver un équilibre entre l'innovation continue et la mise en place de cadres réglementaires adaptés pour garantir une utilisation éthique et responsable de ces technologies.

Sur ce dernier point se pose le juste équilibre dans la relation homme-machine. L'automatisation permise par l'IA générative, qui devrait s'accompagner, qu'on le veuille ou non, par la réduction d'un nombre significatif de collaborateurs dans les organisations, invite à reconnaître la valeur ajoutée humaine et à anticiper les impacts potentiels sur l'emploi, toutes populations confondues. En effet, le plafond de verre qui protégeait les « travailleurs du savoir » est en passe d'être pulvérisé par ces technologies.

À cette dimension s'ajoute enfin la gestion éthique des données pour en limiter les biais et l'impact environnemental de ces technologies. Le besoin impérieux d'une IA « explicable » (sur ses données d'entraînement ou la logique des modèles) doit viser à limiter les biais et subséquemment les fractures numériques. Pour ce qui est de la gestion des ressources, les calculs informatiques requièrent une quantité colossale d'électricité, pour l'alimentation des plus de 8 millions de centres de données (*data centers*) dans le monde, d'eau, pour les refroidir, et de matériaux rares, puisés dans des contrées lointaines aux conditions de travail parfois douteuses.

## ON NE GRAVIT PAS L'ÉVEREST AVEC UN PORTABLE ET DES CLAQUETTES

Ces défis posés, intégrer l'IA dans son entreprise requiert une dose de bon sens, une bonne hauteur de vue mais surtout une approche méthodologique et collaborative qui permette

de couvrir un maximum d'angles morts et de gérer les aléas. C'est ainsi qu'avant de foncer tête baissée sur ces sujets, surtout quand on est décideur dans une organisation, il est important d'opérer un travail de questionnement préliminaire.

## Les angles morts de la décision éclairée

Nos travaux chez Smart Leaders, nous ont menés à trois constats principaux qui laissent à penser que le sujet de l'IA est insuffisamment pris au bon niveau.

Tout d'abord, on dénote un manque d'anticipation et de compréhension sur le sujet de la part des décideurs, en raison d'un manque d'analyse prospective régulière, d'un manque évident de temps pour aller en profondeur sur des sujets complexes, mais surtout de l'absence de cadre méthodologique pour décoder ce qu'il y a derrière les *buzzwords* et les prétendues *success stories*. Ainsi, « bien que de nombreux dirigeants comprennent à un niveau général que de nouvelles compétences sont nécessaires pour travailler avec l'IA générative, leur perception de la manière dont ces changements peuvent créer de la valeur est souvent vague et insuffisamment informée »<sup>3</sup>.

Ensuite, la vision des impacts se résume très souvent à une simple approche de gains de productivité mesurables (le sacro-saint ROI), qui reste la représentation première des décideurs de l'apport de l'IA. Encore régulièrement, il n'est pas rare que nos échanges avec des dirigeants, des DRH ou autres *C-levels* aboutissent à l'opportunité qu'est l'IA pour automatiser une tâche ou un processus. En clair, si une tâche prend 2 heures aujourd'hui et qu'elle ne prend plus qu'1 h 30 avec une IA, alors l'organisation considère que cela vaut le coup d'y aller. Ce serait donc si simple ?



Figure 3 : Pour le dirigeant : la méthodologie ou la panique (Source : Smart Leaders).

Enfin, l'implication de la strate managériale, qui est celle en charge de décliner la vision et d'accompagner les équipes sur le terrain, est encore très peu prise en compte, ces derniers n'ayant par exemple souvent pas accès à des outils IA de qualité au travail ni de clé pour répondre à un collaborateur qui viendrait leur demander ce que l'IA va changer

<sup>3</sup> [https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-gen-ai-skills-revolution-rethinking-your-talent-strategy?cid=other-soc---oth---ip&sid=soc-POST\\_ID&linkId=581681029](https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-gen-ai-skills-revolution-rethinking-your-talent-strategy?cid=other-soc---oth---ip&sid=soc-POST_ID&linkId=581681029)

dans son job demain. Il leur est donc difficile d'être ambassadeur d'une transformation mal comprise et peu tangible.

C'est pourquoi il nous semble capital de doter les décideurs et les managers d'une approche méthodologique rigoureuse pour prendre le sujet par le bon bout et pour le décliner dans un espace-temps respectueux des parties prenantes et de l'organisation elle-même.

### **Intégrer l'IA, une affaire de méthode plus que de moyens**

Tout d'abord, vouloir automatiser une tâche, c'est partir du postulat que l'on sait de quelle tâche on parle. Or, il n'est pas certain que le quotidien des collaborateurs corresponde à la vision qu'en a le management.

C'est donc l'occasion de se reposer la question de ce que font réellement les collaborateurs qui travaillent dans l'entreprise et de comment ils le font : qu'est-ce qu'ils produisent, quels outils ils utilisent, combien de temps ils y passent, de quoi ou de qui ils ont besoin, quel plaisir ils prennent à faire telle ou telle activité, quels savoirs tacites ils intègrent dans leurs actions et leurs prises de décision... Le fait de formaliser ces éléments va permettre de mieux qualifier la question de l'IA pour atterrir sur les fameux cas d'usage (la notion d'usage faisant explicitement référence à ce qui est concrètement réalisé et que seuls les opérationnels connaissent).

Et c'est à ce moment-là qu'il convient de ne pas considérer la seule notion de productivité mais d'intégrer des critères inhérents au « recrutement de cette technologie », par exemple la qualité des données en entrée et en sortie, les interdépendances avec les autres équipes et les autres départements, les risques d'erreurs et de biais et leurs conséquences, les nuances culturelles et contextuelles ou encore l'importance de la relation humaine et de l'intelligence collective dans la réalisation de l'activité, qui est aujourd'hui la partie la plus difficile à encoder dans des algorithmes.

Ce travail de recensement va également regonfler à bloc les équipes et leur engagement. Par l'intérêt porté à leur quotidien, il sera plus simple de juguler les émotions négatives vis-à-vis de l'IA, de déployer à tous les niveaux une nouvelle façon de travailler et une culture de l'hybridation et *in fine* d'éviter que l'hyper-automatisation ne sape petit à petit les fondements de ce qui a fait la réussite de l'organisation. Car s'il existe un sourire dans la réalisation d'une activité, automatiser serait sans doute une grave erreur.

Les cartographies qui résultent de ces réflexions nourriront la prise de décision du recours à l'IA, et dans quelle proportion, en vue de créer de la valeur à la fois sur les plans économiques mais aussi sociaux et humains, réputationnels et environnementaux.

### **AUJOURD'HUI, DEMAIN N'EST PLUS SI LOIN**

Il y a eu le *web* 1 puis 2 et 3. À quel stade de l'IA en sommes-nous aujourd'hui ? Si l'IA générale semble encore bien loin, nul ne sait ce que nous réservent les prochains mois.

C'est pourquoi il est plus que jamais critique pour une organisation de mettre en place un écosystème stratégique de gestion des compétences pour anticiper les prochaines phases d'intégration de l'IA et plus généralement pour survivre voire prospérer dans ce monde en transformation perpétuelle.

À court terme, au-delà d'un certain nombre de compétences techniques que la plupart des dirigeants et DRH ont rapidement identifié, il faut sortir d'une vision centrée sur le prompt pour aller vers une dynamique qui étudie les impacts des IA génératives sur les modes de fonctionnement de l'entreprise.



# Persona 1 : Clara, Manager de proximité



Figure 4 : Exemple de cartographie des impacts sur les activités et compétences (Source : Smart Leaders).



Parmi les compétences qui vont prendre de l'ampleur dans les mois qui viennent, on verra donc le fait de savoir problématiser et contextualiser, l'esprit critique pour analyser les données utilisées et les résultats produits, la connaissance des mécanismes cognitifs et psychologiques, les considérations éthiques ou encore la gestion des compétences de ses collaborateurs.

Et gare aux raccourcis qui laisseraient penser que seulement un certain type d'acteurs sont concernés par ces nouveaux besoins en compétences (les plus jeunes, les moins qualifiés...). Les recherches et retours d'expérience récents démontrent que les « travailleurs du savoir » et les cols blancs sont désormais pleinement impactés, et ce à double titre : à la fois dans le fait qu'une partie de leurs activités pourrait être remplacée par l'IA mais aussi dans leur capacité à faire émerger une équipe, un département et à terme une entreprise « augmentés », dans lesquels l'IA est un nouveau collègue. Si l'on ne devait citer qu'une compétence que tous les décideurs doivent afficher, c'est l'humilité de ne pas savoir, au même titre que tous les autres, dans quelle direction le monde va.

Mais « ne pas savoir » aujourd'hui ne veut pas dire « ne pas préparer l'avenir ». Pour le moyen terme, il est devenu vital de savoir anticiper et de déployer des mécanismes et une culture d'adaptabilité et d'innovation. On peut espérer que les directions générales et RH adopteront une approche stratégique et prospective de *Workforce Planning*, dont la dimension liée à l'IA aura une place prépondérante.

Et en allant au bout de notre vision d'une organisation dont la compétence serait un actif stratégique, piloté et partagé par tous, il serait souhaitable qu'elle mette progressivement en place un écosystème complet de la compétence comme celui présenté ci-dessous, en réinventant son modèle de fonctionnement interne et en développant une multitude de partenariats au sein d'un écosystème pluridisciplinaire, comme l'IA peut l'être également.



Figure 5 : Les 4 piliers d'un écosystème de la Compétence / d'une *skill-based organization* (Source : Smart Leaders).

L'IA ne serait-elle donc pas l'opportunité rêvée pour se réinventer ?

# Construisons un cadre ambitieux et apaisé pour mettre les IA au service de l'éducation et de la formation

Par Oriane LEDROIT

Déléguée générale de l'association EdTech France

À l'heure où les métiers évoluent, que certains disparaissent, que d'autres émergent, permettre à tous de se former et de renouveler ses compétences est une des conditions *sine qua non* pour : construire notre résilience technologique et écologique ; nourrir notre puissance de création et d'innovation ; réindustrialiser le territoire ; donner une place à chacun d'entre nous dans la société et avec cette place (re)donner du pouvoir d'agir ; enfin, d'une certaine manière tenir notre rang.

Ces défis majeurs exigent que nos espaces d'enseignement et nos modalités d'apprentissage – du plus jeune âge et tout au long de la vie – évoluent, s'ajustent et tirent parti – de manière ambitieuse et pragmatique – des opportunités des technologies et notamment de la puissance des intelligences artificielles.

Cela suppose de penser sereinement notre rapport aux IA au service de l'enseignement et de la formation. Quels enjeux ? Quels défis ? Quels apports ? Quel cadre pour faire en sorte de mettre les IA au service d'une société apprenante ?

## L'ÉDUCATION EST UN DES DÉFIS MAJEURS DE NOTRE SOCIÉTÉ ET DE NOTRE ÉCONOMIE

C'est d'abord un défi car nous devons faire mieux. Nous devons faire mieux à l'école, dans les études supérieures et dans la formation professionnelle.

Notre école publique décroche sur l'acquisition des savoirs<sup>1</sup>, sur la lutte contre les déterminismes sociaux<sup>2</sup>, sur l'accompagnement de tous les élèves y compris ceux à besoins particuliers<sup>3</sup> ou encore sur l'attractivité du métier d'enseignant.

<sup>1</sup> Les rapports PISA traduisent chaque année une baisse du niveau de nos élèves en mathématiques, en compréhension de l'écrit et en sciences, <https://www.oecd.org/fr/about/programmes/pisa.html>

<sup>2</sup> Notre école est celle où l'origine sociale pèse le plus lourd dans les résultats scolaires, <https://www.inegalites.fr/pisa>

<sup>3</sup> Ils sont près de 10 % à souffrir de troubles des apprentissages, sans compter ceux à besoins éducatifs particuliers. Ils sont nombreux à errer, sans accompagnant d'élève en situation de handicap (AESH), au risque d'être déscolarisés.

Les établissements d'enseignement supérieur se multiplient – notamment ceux à statut privé – et accueillent toujours plus d'étudiants (+ 21,7 % d'étudiants entre 2010 et 2020 en France<sup>4</sup>) mais la qualité des formations, la valeur des diplômes, leur cohérence avec les débouchés professionnels existants ou à venir sont questionnées.

Même si le chômage est aujourd'hui relativement bas, les besoins en compétences sont massifs du fait de trois tendances qui s'alimentent entre elles :

- d'abord, l'obsolescence accélérée des compétences<sup>5</sup> : l'OCDE souligne que la durée de vie d'une compétence technique est passée de 30 ans en 1987 à 2 ans aujourd'hui ;
- ensuite, l'apparition rapide de nouvelles compétences de plus en plus recherchées voire exigées sur le marché du travail (le *prompt engineering* par exemple) ;
- enfin, l'évolution dans la « hiérarchie » des compétences : depuis plusieurs années, nous constatons le passage d'une survalorisation des compétences techniques à un regain d'intérêt pour les *soft skills*<sup>6</sup>, ces compétences psycho-sociales traduisant d'une certaine manière l'habileté à évoluer en collectif dans un monde complexe et incertain.

Au-delà d'être un défi majeur, l'éducation est même sans doute un « défi socle » dont la réussite conditionne notre capacité collective à relever les autres défis. Une éducation performante, une formation accessible tout au long de la vie, permettront de nous outiller pour accompagner la transition écologique, le progrès numérique, la recherche de compétitivité et de souveraineté, notre cohésion sociale et notre vivre ensemble.

L'enjeu est de taille.

Surtout, ce défi se pense et les réponses s'élaborent en même temps que – et même avec – la révolution des IA et notamment celle des IA génératives qui bousculent fortement l'enseignement et l'apprentissage.

Le Sénat a d'ailleurs récemment rappelé que 90 % des élèves de seconde affirmaient avoir déjà utilisé une IA générative pour réaliser leurs devoirs<sup>7</sup>.

Les IA redéfinissent les modalités d'accès à l'information – on parle d'infobésité<sup>8</sup> – et à une certaine forme de savoirs – on parle moins de savoirs de référence. Les IA génératives telles que ChatGPT abaissent même d'une certaine façon le coût d'entrée dans la technologie : il « suffit » désormais de parler dans sa langue au robot conversationnel pour qu'il nous réponde et nous partage des informations agrégées.

Les IA bouleversent les modalités même de la création et de la génération d'une réflexion (on peut « prompter » pour éviter la feuille blanche). Elles questionnent les piliers traditionnels de l'enseignement supérieur : la préparation d'un diplôme aux compétences pré-établies, à un moment donné dans une vie, avec un équilibre entre le savoir académique et la préparation au marché du travail.

---

<sup>4</sup> [https://publication.enseignementsup-recherche.gouv.fr/eesr/8/EESR8\\_ES\\_08-les\\_evolution\\_de\\_l\\_enseignement\\_superieur\\_depuis\\_50\\_ans\\_croissance\\_et\\_diversification.php](https://publication.enseignementsup-recherche.gouv.fr/eesr/8/EESR8_ES_08-les_evolution_de_l_enseignement_superieur_depuis_50_ans_croissance_et_diversification.php)

<sup>5</sup> Défini dans les années 1970 par Josh Kaufman comme « l'insuffisance des savoirs ou compétences actualisés nécessaires à un travailleur pour continuer d'être parfaitement performant dans son activité professionnelle actuelle ou future ».

<sup>6</sup> <https://exed.centralesupelec.fr/actualites/les-soft-skills-cest-quoi-exactement/>

<sup>7</sup> <https://www.senat.fr/travaux-parlementaires/office-et-delegations/delegation-a-la-prospective/detail-actualite/rapport-ia-n-3-3911.html>

<sup>8</sup> Définition sur Wikipédia : [https://fr.wikipedia.org/wiki/Surcharge\\_informationnelle#:~:text=La%20surcharge%20informati%20](https://fr.wikipedia.org/wiki/Surcharge_informationnelle#:~:text=La%20surcharge%20informati%20)

Elles inquiètent de nombreux travailleurs ; faisant planer le risque d'un remplacement par la machine<sup>9</sup> même si aujourd'hui la plupart des cas d'usages restent encore en phase expérimentale dans les entreprises.

À l'image des analyses du philosophe Bernard Stiegler caractérisant le numérique comme un *pharmakon*<sup>10</sup>, les IA sont à la fois des opportunités pour les apprentissages mais aussi des risques. Pour les enseignants-formateurs, pour les élèves-étudiants-apprenants. Pour les institutions d'enseignement également elles-mêmes, et avec elles l'imaginaire sur lequel se fonde en partie notre communauté de destin.

Tel est l'objet de cet article. Partager une vision globale mais non exhaustive des opportunités des IA dans l'apprentissage mais aussi des besoins pour dessiner un chemin, cadré et apaisé, de collaboration entre les IA et l'éducation.

## LES APPORTS DES IA À L'ÉDUCATION ET À L'APPRENTISSAGE

Les apports des IA sont détaillés ci-dessous en adoptant un prérequis théorique essentiel : les IA ne remplaceront jamais l'humain, le pédagogue, la relation avec un enseignant, un formateur, un mentor, un tuteur. Les IA sont de puissants supports au déploiement de l'intelligence humaine<sup>11</sup>. Plus qu'un outil, les IA doivent susciter la réflexion et la mise en action de ses utilisateurs, plutôt que s'y substituer.

Les apports des IA sont principalement de trois ordres : d'abord la massification, ensuite l'efficacité, enfin le pilotage précis par les données.

### Les IA au service d'une formation plus massive

Nous l'avons évoqué plus haut : nos besoins en formation, en reconversion (*reskilling*) et en perfectionnement (*upskilling*) sont gigantesques. Les IA sont une chance pour y répondre.

Les IA sont des outils qui – intégrés par les organismes de formation et les établissements d'enseignement – permettent de démultiplier la rapidité à produire des contenus pédagogiques correspondant aux besoins des apprenants.

Ils permettent également d'accélérer la capacité à décliner des contenus produits dans des langues différentes afin de rendre ces formations accessibles au plus grand nombre.

Ils permettent enfin de favoriser la mise à jour des contenus pédagogiques garantissant ainsi la fiabilité et la véracité des formations dispensées. Ce qui est une exigence clé pour la qualité même des formations.

Intégrées à des processus sérieux de recherche et développement et développées dans le respect d'un cadre éthique par des professionnels de l'apprentissage, elles sont donc des alliées formidables pour créer des solutions robustes et adaptées aux évolutions des aspirations personnelles des personnes et aux besoins nouveaux du marché du travail et ainsi consolider notre appareil de formation.

---

<sup>9</sup> <https://academy.visiplus.com/barometre-formation-pro-2024>

<sup>10</sup> <https://arsindustrialis.org/vocabulaire-pharmakon-pharmacologie>

<sup>11</sup> Voir à ce sujet le concept développé par le Dr Luc Julia d'« intelligence augmentée », <https://shs.cairn.info/revue-paysan-et-societe-2021-1-page-49?lang=fr>

## Les IA au service d'un apprentissage plus efficace

Les intelligences artificielles soutiennent les mécanismes cognitifs des apprenants en favorisant l'ancrage mémoriel, la personnalisation de l'apprentissage et la remédiation. Illustrations.

D'abord, les IA favorisent l'appétence et la confiance de l'apprenant dans la pratique de ses compétences. Comme le développe Séverine Erhel, maître de conférences en psychologie cognitive, il s'agit ici de pousser « l'apprenant à avoir des activités métacognitives, donc à prendre conscience de ce qu'il traite et de son niveau de connaissance pendant qu'il réalise l'apprentissage »<sup>12</sup>. Par exemple, de nombreuses solutions numériques pédagogiques vont intégrer aux séquences d'apprentissage des exercices interactifs, des messages motivationnels, des conseils, pour soutenir et tester la compréhension de l'apprenant en temps réel.

Ensuite, les IA ouvrent la voie à la différenciation pédagogique – le graal de l'éducation : elles permettent de proposer des modules d'exercices adaptatifs. Le contenu et l'expérience utilisateur (UX) sont élaborés d'une part à partir de la connaissance scientifique disponible sur le fonctionnement de l'être humain en situation d'apprentissage et d'autre part à partir du niveau et des besoins de chaque apprenant. Cela rend l'acquisition de compétences plus fluide et plus intuitive. Les contenus viennent ainsi soutenir les mécanismes cognitifs à l'œuvre dans l'intelligence humaine.

Par ailleurs, les IA contribuent à améliorer la qualité de l'interaction entre un enseignant et ses élèves, cruciale dans leur réussite scolaire. Elles peuvent notamment améliorer les retours de l'enseignant lors de la correction de copies pour favoriser les *feedbacks* utiles à leur progression. D'autres IA vont – de la même façon – soutenir la posture pédagogique des enseignants et des formateurs et les recentrer sur leur rôle de pédagogue.

Enfin, les IA étayent les intuitions des enseignants ou formateurs et soutiennent leur travail d'analyse devant des tâches complexes, quand les activités éducatives articulent une infinité de compétences. Un des rôles de l'IA est de pré-catégoriser les difficultés des élèves pour guider la correction, en tenant compte de l'historique de l'élève et de la progression du reste de la classe.

## Pilotage par les données et IA dans l'éducation : de l'individuel au niveau macro-économique

Les IA aident à faire parler les données du système éducatif dans sa globalité (de l'école à la formation tout au long de la vie) : les IA rendent possibles un pilotage des actions par la donnée, dans le respect des cadres juridiques existants (notamment en Europe, le Règlement général de protection des données).

Ce pilotage serait de deux ordres : d'une part au niveau de l'élève pour mieux anticiper ses apprentissages et éviter par exemple son décrochage ; d'autre part au niveau macro-économique pour suivre l'utilité d'une réforme ou d'un dispositif ou encore pour cibler précisément des investissements à partir de besoins identifiés.

En France, un des axes d'innovation serait de pouvoir réinvestir les données issues des évaluations nationales pour améliorer les parcours d'apprentissage et perfectionner les solutions disponibles pour les enseignants et les élèves. Notons d'ailleurs que les États-Unis utilisent d'ores et déjà l'IA pour prévoir le décrochage et le prévenir.

---

<sup>12</sup> <https://www.inserm.fr/actualite/education-faut-il-reculer-sur-le-numerique-a-lecole/>

Il en va de même pour les données sur l'impact des formations professionnelles (taux de diplomation, taux d'insertion dans l'emploi suite à la formation, satisfaction des apprenants) qui ne sont pas suffisamment harmonisées, mises en commun et consolidées<sup>13</sup>, ou encore sur l'analyse objectivée des besoins en compétences (des entreprises, des bassins d'emploi, des territoires) qui sont encore trop éparpillées.

## MIEUX FORMER AU ET PAR LES IA

Les potentialités des IA au service de l'apprentissage sont nombreuses et inédites. Elles se traduisent déjà concrètement par des solutions, logiciels, outils, plateformes disponibles aujourd'hui dans notre quotidien, dans les classes, dans les amphithéâtres des universités, dans les dispositifs de formation de France Travail, sur nos *smartphones*...

Mais nous ne sommes ni techno-solutionnistes, ni techno-béats.

Évidemment, toute IA n'est pas bonne « en soi » pour l'apprentissage. Son utilité est largement conditionnée par le projet pédagogique (comment ? pour faire quoi ?) et la qualité de la ressource technologique mise à disposition.

Trois actions complémentaires nous paraissent propres à garantir que les IA soient au service de la performance éducative et du défi des compétences.

D'abord, renforcer les liens entre la recherche et les entreprises productrices de ces technologies. En France, les "edtech" collaborent déjà avec d'excellents laboratoires français de recherche (CNRS, Inria, Cnam, etc.) dans des disciplines variées telles que la psychologie du développement, la psychologie cognitive, les neurosciences, les processus d'apprentissage, les sciences de l'éducation, la psychiatrie, l'information et la communication, les technologies numériques pour l'éducation<sup>14</sup>. Les sciences sont essentielles au stade du développement produit mais aussi au stade de l'évaluation de celui-ci. Nous y reviendrons pour la troisième et dernière recommandation. Il s'agit ainsi de faciliter les mises en relation entre entrepreneurs et chercheurs pour démultiplier les collaborations mais également de mettre en place des espaces de type bacs à sable dans lesquels peuvent se concrétiser ces collaborations vertueuses.

Ensuite, former les enseignants-formateurs à ces IA, à leurs potentialités, à leurs risques, aux possibles usages constructifs. À leur utilité dans la pédagogie et dans la didactique des disciplines. Les outiller dans leur formation initiale mais surtout soutenir leur appropriation tout au long de leur carrière en formation continue. Former également les élèves, les étudiants, les apprenants, les salariés à une littératie poussée des IA ainsi qu'à ses métiers (et pas seulement aux métiers d'ingénieurs). Gilbert Simondon insistait déjà sur le fait que « plus l'objet technique est perçu comme une boîte noire d'entrées et de sorties, et plus l'homme qui l'utilise lui est aliéné »<sup>15</sup>. Maximiser les opportunités technologiques tout en réduisant leurs dangers passe évidemment par l'éducation. À ce sujet d'ailleurs, l'école ne peut être un sanctuaire. Elle doit être un lieu d'une éducation graduée au et par le numérique. Dès le plus jeune âge, nos enfants doivent être accompagnés pour acquérir les savoirs informatiques et les cultures numériques, et leur assurer une prise en main

<sup>13</sup> Voir à ce sujet cette note : <https://cpformation.com/wp-content/uploads/2023/02/Groupe-de-travail-Mesure-de-l'impact-de-la-formation-VF.pdf>

<sup>14</sup> Par exemple le laboratoire LaPsyDÉ (<https://www.lapsyde.com/>) porté par Grégoire Borst, le laboratoire TECHNÉ (<https://technelabo.univ-poitiers.fr/membres/jean-francois-cerisier/>) avec Jean-François Cerisier, le laboratoire Flowers dirigé par Pierre-Yves Oudeyer (<https://flowers.inria.fr/team/>).

<sup>15</sup> <https://www.lemonde.fr/blog/binaire/2020/04/05/informatique-culture-et-technique-le-schisme-de-simondon/>

critique des technologies. Tout au long de la vie, la formation doit de la même façon contribuer à cette autonomisation.

Enfin, systématiser l'évaluation de la qualité et de l'utilité des IA mises à disposition de l'apprentissage. Les innovations technologiques en matière d'enseignement doivent pouvoir voir le jour, être testées, expérimentées par exemple en laboratoire mais aussi en situation d'apprentissage réel (la classe, la formation). Une fois ce premier niveau de test passé et en fonction du niveau de maturité de ladite solution<sup>16</sup>, il serait utile de définir un protocole d'évaluation permettant de vérifier l'utilité (pour la performance éducative, pour la confiance en soi, pour la prise en compte des besoins particuliers des apprenants, etc.) mais aussi sa qualité (accessibilité, éco-conception) des solutions. Ce cadre d'évaluation et de mesure permettra de certifier de la qualité des solutions technologiques mises à disposition de nos élèves, étudiants, apprenants, enseignants et ainsi à rassurer les différentes parties prenantes de l'enseignement, de la formation mais aussi de leur financement et de leur régulation.

Ces recommandations ne sont pas exhaustives. Elles ne traitent volontairement pas des sujets environnementaux ou encore d'inclusion numérique. Elles dessinent un chemin pragmatique et ambitieux pour une collaboration utile entre les technologies de l'IA et les apprentissages.

Ces recommandations s'inscrivent enfin dans une réflexion apaisée sur les technologies, loin des débats caricaturaux et simplistes sur les écrans qui empêchent de penser et d'agir. Elles résonnent avec la réalité des institutions d'enseignement (école, universités et établissements d'enseignement supérieur, organismes de formation), la vérité des connaissances scientifiques en la matière et des pratiques numériques actuelles<sup>17</sup>.

« C'est cette intelligence des usages qu'il faut défendre politiquement, plutôt que de se rétracter sur un discours technophobe », a justement rappelé Antonio Cassili début 2024, dans *le Nouvel Obs*<sup>18</sup>.

---

<sup>16</sup> À l'image de la méthodologie appliquée par la Nasa, <https://www.nasa.gov/directorates/somd/space-communications-navigation-program/technology-readiness-levels/>, ou en sciences humaines et sociales sur les niveaux de maturité technologique, <https://www.inshs.cnrs.fr/fr/cnrsinfo/un-nouvel-outil-devaluation-de-la-maturite-des-projets-en-sciences-humaines-et-sociales>

<sup>17</sup> [https://www.lemonde.fr/societe/article/2024/11/19/nous-proposons-un-chemin-promouvoir-une-education-graduee-au-et-par-le-numerique-ayant-pleinement-sa-place-a-l-ecole\\_6401945\\_3224.html](https://www.lemonde.fr/societe/article/2024/11/19/nous-proposons-un-chemin-promouvoir-une-education-graduee-au-et-par-le-numerique-ayant-pleinement-sa-place-a-l-ecole_6401945_3224.html)

<sup>18</sup> <https://www.nouvelobs.com/idees/20240328.OBS86364/plutot-que-d-accuser-les-technologies-de-nous-rendre-idiots-il-faut-questionner-les-desiderata-de-leurs-concepteurs.html>

# Intelligence artificielle et territoires

Par Fabien BAZIN

Président du Conseil départemental de la Nièvre

L'intelligence artificielle se développe progressivement dans les politiques publiques, de manière inégale mais son potentiel est indéniable, tant pour participer à un socle partagé de données fiables que pour répondre aux enjeux cruciaux d'accès aux droits et de résilience des territoires. Cependant, les risques pour les droits et les libertés fondamentales ne sauraient être sous-estimés.

Ces dernières années, les avancées technologiques de l'intelligence artificielle (IA) ont conduit à une explosion de son utilisation. Chacun peut désormais s'improviser producteur de musique, polyglotte ou écrivain en un clic.

Derrière cette façade grand public, l'intelligence artificielle offre de nombreuses opportunités pour le développement et la résilience de nos territoires.

Le service public doit savoir tirer parti de cette révolution tout en régulant son emploi pour prévenir les dérives. Garantir une IA de confiance passera par l'optimisation de son utilisation (loin des vernis *marketing*), par la formation indispensable des élus et des décideurs publics ainsi que par l'inclusion des citoyens pour assurer son encadrement.

Avec la garantie de la qualité des données, de la fiabilité des sources de données et des conditions de son utilisation dans un cadre éthique partagé, l'IA peut se mettre au service du bien-être et de la performance des territoires.

## L'IA POUR MIEUX CONNAÎTRE SON TERRITOIRE

### La connaissance au service des politiques publiques

L'un des principaux enjeux pour utiliser tout le potentiel de l'IA au profit de l'action publique de proximité est la collecte et l'analyse des données.

Grâce à des algorithmes avancés, les collectivités peuvent désormais traiter de grandes quantités de données en temps réel, ce qui permet de mieux comprendre les besoins et les attentes des citoyens.

Par exemple, l'IA peut être utilisée pour analyser les données de mobilité afin d'optimiser les réseaux de transport public, ou pour surveiller la qualité de l'air et de l'eau afin de prendre des mesures préventives en matière de santé publique.

Pour la mise en place d'un projet IA, il est crucial de partir des besoins spécifiques de la collectivité, en lien avec ses compétences. La *data science*, l'apprentissage automatique et la vision par ordinateur (*computer vision*) sont des technologies clés pour optimiser divers aspects de la gestion territoriale, comme l'entretien des arbres, la maintenance prédictive des réseaux d'eau, l'automatisation des droits sociaux ou la gestion de la mobilité et des infrastructures publiques.



Les séries temporelles (ensemble de données qui représente l'évolution d'un phénomène au cours du temps) et le traitement automatique du langage naturel (NLP, *natural language processing*) sont également des outils précieux pour la prédiction des usages des transports en commun<sup>1</sup>, la modélisation de la fréquentation des espaces publics et l'amélioration des services publics.

Les collectivités peuvent utiliser différentes « briques » technologiques de l'IA en fonction de leurs besoins spécifiques. Par exemple, la vision par ordinateur permet l'analyse d'images pour des applications comme le suivi des déchets ou la gestion des exploitations agricoles.

La recherche sémantique facilite l'instruction et l'élaboration de documents d'urbanisme. La synthèse de données et la complétude des données permettent de combler les lacunes dans les informations collectées.

### L'IA, énergivore ?

Pour alimenter ces algorithmes, les territoires doivent générer des données pertinentes en grande quantité.

La consommation d'énergie électrique est une des difficultés auxquelles la France est confrontée pour développer l'IA.

Dans ce contexte, la région Centre-Val-de-Loire a élaboré le Climate Data Hub, un cadre de confiance pour le partage et une utilisation éthique et responsable des données au service des politiques climatiques<sup>2</sup>. Trente-deux partenaires publics et privés se sont rassemblés dans cette démarche en signant un accord de *consortium* garantissant l'usage de ces données au service de l'intérêt général. Cet écosystème favorable a permis l'émergence de plusieurs projets dont PREVIZO qui utilise une intelligence artificielle frugale (peu consommatrice en carbone) pour une meilleure gestion des ressources en eau.

Pour mieux connaître son territoire, le déploiement de réseaux d'initiative publique bas débit pour les objets connectés facilite la collecte des données grâce à une technologie économe en énergie et durable.

C'est le choix qu'a fait la Nièvre en planifiant l'installation d'un réseau LoraWan à l'échelle du département. Les objets connectés, également connus sous le nom d'Internet des objets (IoT), sont des dispositifs équipés de capteurs et de logiciels qui leur permettent de collecter et d'échanger des données. Dans le contexte des collectivités territoriales, ces objets peuvent être utilisés pour optimiser divers services publics. L'IA peut analyser les données collectées par les objets connectés pour fournir une vision précise du territoire et aider à la prise de décision.

C'est l'un des axes définis dans le rapport remis au Gouvernement par Valérie Nouvel<sup>3</sup> en octobre 2023 sur les territoires connectés et durables. La création de jumeaux numériques territoriaux permettra de déployer de vraies stratégies des usages numériques.

---

<sup>1</sup> Cela doit impérativement s'entendre hors situation de crise, où des prédictions fondées sur le régime permanent peuvent être contre-productives. Il est utile de prévoir ces cas de fonctionnement en mode dégradé.

<sup>2</sup> <https://www.centre-valdeloire.fr/agir/protéger-lenvironnement/la-data-une-solution-face-au-rechauffement-climatique>

<sup>3</sup> <https://www.entreprises.gouv.fr/la-dge/actualites/territoires-connectes-et-durables-remise-dun-rapport-au-gouvernement>

## L'IA POUR LA RÉSILIENCE DES TERRITOIRES

### Un atout évident au regard des enjeux

Dans un contexte de changements climatiques, de crises sanitaires et de transformations économiques, la résilience des territoires devient une priorité pour les collectivités territoriales françaises.

L'intelligence artificielle émerge comme un outil puissant pour renforcer cette résilience. L'IA peut aider les collectivités à mieux prévoir et gérer les risques naturels et climatiques. Les modèles prédictifs basés sur l'IA analysent les données historiques et en temps réel pour anticiper les catastrophes naturelles, telles que les inondations, les tempêtes et les vagues de chaleur.

Ces informations permettent aux collectivités de mettre en place des plans de prévention et de réponse plus efficaces.

L'IA joue également un rôle crucial dans la gestion des énergies renouvelables. Les algorithmes de prévision météorologique permettent de mieux anticiper la production d'énergie solaire et éolienne, tandis que les systèmes de stockage intelligent optimisent l'utilisation de ces ressources. Les collectivités peuvent ainsi maximiser l'utilisation des énergies renouvelables et réduire leur dépendance aux énergies fossiles.

L'IA peut en outre, par profilage, aider à déterminer l'accès aux droits sociaux. Ainsi, France Travail<sup>4</sup> « utilise ainsi un système d'IA pour déterminer l'éligibilité des chômeurs à l'allocation chômage d'aide au retour à l'emploi, ainsi que le montant de cette aide et sa durée ».

### Les conditions au développement et à la mise en œuvre des projets d'IA

Le livre blanc de la communauté des acteurs de l'IA dans les territoires<sup>5</sup>, édité par l'ECOLAB du ministère de la Transition écologique, propose un ensemble de préconisations pour la mise en place d'un projet IA.

Le coût financier et écologique des projets IA nécessite une sélection rigoureuse des initiatives les plus utiles et encourage le partage d'expériences entre collectivités.

La mutualisation est devenue une priorité pour répondre aux enjeux communs des territoires. L'IA permet de prédire et de classifier les dynamiques territoriales, optimiser les processus pour réduire la consommation de ressources et améliorer les services publics. Elle constitue ainsi une opportunité pour les acteurs publics de répondre aux enjeux de la transition écologique.

Le pilotage dynamique et optimisé des ressources, la prédiction de séries temporelles, et la résolution de modèles de diffusion complexes sont d'autres applications potentielles de l'IA.

Les collectivités sont proactives dans la mise en œuvre de politiques publiques et programmes dédiés à la transition énergétique et environnementale. Déployée de manière responsable et éthique, l'IA peut soutenir l'action publique en détectant des anomalies, améliorant les contrôles, et facilitant le travail des agents locaux. En automatisant

---

<sup>4</sup> <https://www.vie-publique.fr/parole-dexpert/293547-ia-quel-potentiel-et-quels-risques-dans-les-services-publics>

<sup>5</sup> <https://greentechinnovation.fr/storage/V3-Livre-blanc-1.pdf>

certaines tâches répétitives, l'IA permet de réduire le temps d'expertise technique et de faciliter la prise de décision.

Il convient également de souligner qu'une des conditions du développement de l'IA au sein des collectivités territoriales demeure, comme on le voit dans le secteur bancaire par exemple, le recrutement d'experts en capacité de garantir l'intégrité des données, l'utilisation éthique de celles-ci et de prévenir les risques de défaillance des systèmes.

En dépit des 908 millions d'euros dédiés à la transformation numérique des services publics dans le cadre de France Relance, des ressources devront être affectées à ces tâches essentielles pour nourrir la confiance des usagers et des collectivités.

Le budget de l'État, qui donne aujourd'hui lieu à des coupes déléatoires dans les comptes des collectivités, n'augure pas de conditions propices pour répondre à ce besoin accru de contrôle et de sécurité.

Enfin, une autre condition au développement de l'IA reste l'acculturation.

### **L'IA DANS LA DÉCISION PUBLIQUE : DE L'ACCULTURATION À L'UTILISATION**

L'étude « Quels sont les facteurs qui influencent la perception de l'adoption de l'intelligence artificielle par les managers publics ? »<sup>6</sup> publiée par la Commission européenne en novembre 2024 montre que l'adoption de l'IA dans le secteur public n'est plus une simple perspective.

#### **Le constat d'une utilisation avant tout technique**

Dans ses conclusions, cette étude montre que l'IA est une réalité tangible dans les organisations publiques. Cependant, cette technologie est moins utilisée dans la prise de décision politique que dans les travaux internes.

À titre d'exemple, la plateforme Delibia est conçue pour simplifier et automatiser les tâches administratives. Grâce à son modèle de langage avancé et à sa technologie de *retrieval augmented generation* (RAG), elle permet de réduire considérablement le temps consacré aux tâches répétitives. Les agents peuvent ainsi se concentrer sur des missions à plus forte valeur ajoutée, améliorant ainsi la qualité du service public. En s'appuyant sur une base de données riche et constamment mise à jour, la plateforme offre aux élus et agents des réponses fiables à leurs questions. Cela permet de prendre des décisions éclairées et de mieux répondre aux besoins des citoyens.

Il convient de souligner que disposer d'une expertise interne en intelligence artificielle est l'un des facteurs favorisant le recours à l'IA au sein des organisations, au même titre qu'une stratégie claire concernant les aspects techniques mais aussi légaux, éthiques et de gouvernance. Favoriser l'utilisation de l'IA dans nos organisations requiert un investissement à long terme.

L'intelligence artificielle transforme rapidement tous les secteurs, y compris celui des collectivités territoriales. Mais comment les élus peuvent-ils tirer pleinement parti de cette technologie sans une formation adéquate ? L'IA n'est pas seulement une technologie de pointe. Elle représente une révolution dans la manière dont les données sont traitées et utilisées.

---

<sup>6</sup> <https://interoperable-europe.ec.europa.eu/collection/public-sector-tech-watch/document/what-factors-influence-perceived-artificial-intelligence-adoption-public-managers>

Pour les élus, comprendre les enjeux technologiques de l'IA est essentiel pour prendre des décisions éclairées. Une formation adéquate permet de démystifier l'IA et de comprendre ses applications potentielles dans les collectivités territoriales. C'est en ce sens que le conseil départemental de la Nièvre a souhaité, le 9 avril dernier, organiser un séminaire à destination de ses élus et de son comité directeur sur les opportunités et les risques liés à l'IA. Assurer un niveau de compréhension suffisant et homogène permet de favoriser l'adoption des solutions d'IA.

## **Au-delà des organisations publiques, la place des citoyens !**

Cette adoption passe aussi par l'acculturation et la volonté des citoyens. Si les citoyens sont prêts à utiliser des services à base d'IA, les organisations publiques ont plus de facilité à se saisir des possibilités offertes par cette technologie. L'inclusion des citoyens dans la définition d'une stratégie d'utilisation de l'IA, les échanges réguliers entre techniciens et citoyens et l'évaluation de la disposition de ces derniers à utiliser des services d'IA sont autant d'encouragements à l'intégration de ces solutions dans les organisations publiques.

Mais la place des citoyens, c'est aussi leur protection.

Si le potentiel de l'IA peut paraître évident pour mieux cibler les populations précaires, faciliter l'automatisation des droits sociaux ou lutter contre la fraude fiscale, son développement, qui s'appuie sur la collecte et l'analyse de données à caractère personnel, fait courir, dans le cadre de certaines politiques publiques (solidarité, mais aussi sécurité avec la reconnaissance faciale), le risque d'une atteinte aux libertés fondamentales.

En effet, au-delà du droit au respect de la vie privée, l'utilisation de données personnelles sans que les individus n'en aient toujours pleinement conscience, questionne également la liberté d'aller et venir, la liberté d'expression, de manifestation et de conscience.

# Profils d'appropriation de l'intelligence artificielle générative dans l'éducation

Par Loubna MOURTAJJI et Nathalie CHISS

Membres du LEFMI (Laboratoire d'économie, finance, management et innovation)

L'intelligence artificielle bouleverse les pratiques dans tous les secteurs, y compris l'éducation où des outils comme ChatGPT suscitent de nombreuses attentes et interrogations. Cet article explore, de façon empirique, la perception des enseignants face à cet outil, identifiant les facteurs influençant son adoption ainsi que les réticences manifestées. L'analyse qualitative des entretiens semi-directifs révèle que la perception d'utilité de ChatGPT par les enseignants est très contrastée.

Ces perceptions sont largement influencées par des variables contextuelles, telles que le lien avec la technologie, la discipline enseignée, le statut ou encore l'âge des enseignants. Les résultats obtenus permettent d'identifier quatre profils d'enseignants et d'apporter des recommandations afin d'optimiser l'utilisation de ChatGPT dans un cadre éducatif de manière efficace et responsable, tout en tenant compte des implications éthiques, pratiques et pédagogiques de cette intégration.

## INTRODUCTION

La technologie modifie les pratiques éducatives et redéfinit les rôles traditionnels des enseignants, leur interaction avec les étudiants et avec les outils technologiques. L'arrivée de l'intelligence artificielle, et d'outils d'IA générative comme ChatGPT, questionne la communauté éducative qui oscille entre enthousiasme et scepticisme.

Pour comprendre ces dynamiques, nous mobilisons le concept d'appropriation qui éclaire les processus par lesquels un individu adopte, transforme ou s'oppose à une technologie en fonction de ses besoins, ses valeurs et son contexte (matériel, financier, temporel, organisationnel). En effet, l'appropriation ne se limite pas à une simple utilisation technique, mais engendre une reconfiguration des pratiques et des identités professionnelles (Paquelin, 2009). Appliqué à ChatGPT, ce concept offre un cadre pertinent pour explorer les différents profils d'enseignants utilisateurs, leurs motivations, et les tensions qui émergent face à une IA aux implications pédagogiques, éthiques et sociales complexes (Collin et Marceau, 2022).

Dans un contexte où l'éducation est appelée à devenir de plus en plus numérique, il est important de dépasser l'approche technique et utilitaire de l'outil pour saisir les dimensions humaines de cette appropriation. L'ambition de cet article est de contribuer à une meilleure compréhension des pratiques pédagogiques, augmentées par l'IA, tout en intégrant une réflexion sur les enjeux éthiques et responsables liés à son usage. À travers une analyse typologique des profils d'enseignants utilisateurs, nous explorons les contours de ce nouveau paysage éducatif, entre opportunités, défis et nécessité de régulation. Ce

travail s'inscrit dans une démarche analytique et prescriptive, cherchant à enrichir le débat sur le rôle de l'IA dans l'enseignement et à fournir des recommandations pour une appropriation réfléchie, éthique et durable de ces outils par les enseignants.

La première partie présente les différents modèles applicables lorsque le concept d'appropriation en pédagogie est appliqué. La méthodologie utilisée est ensuite précisée. Enfin la dernière partie s'intéresse aux résultats obtenus et aux recommandations envisagées.

## CONCEPT D'APPROPRIATION APPLIQUÉ À LA PÉDAGOGIE

Le concept d'appropriation, issu de la sociologie des usages, permet d'envisager le rôle actif des acteurs, à travers la créativité et l'innovation dont ils font preuve. Il s'agit d'étudier les représentations que construisent les acteurs à propos de leurs pratiques (professionnelles, de formation, personnelles). L'appropriation fait ainsi le lien entre la conception et l'usage effectif des outils (Bourdet et Leroux, 2009). Selon Trinquecoste et Bidan (2011), travailler sur l'appropriation des technologies de l'information conduit à travailler sur l'utilisateur devenant usager et « inventant sa propre routine », y compris dans une logique de contournement ou de détournement de la technologie afin de gagner en autonomie par exemple.

Au niveau de l'éducation, plusieurs modèles apparaissent. Traitant de l'intégration de la technologie en pédagogie, Koehler et Mishra (2005) proposent un cadre théorique, la Connaissance du contenu pédagogique technologique (CCPT ou *Technological Pedagogical Content Knowledge*, TPCK en anglais), basé notamment sur l'interaction entre les trois principales composantes des environnements d'apprentissage : le contenu, la pédagogie et la technologie. En effet, les enseignants sont amenés à mobiliser trois champs principaux de connaissances lors de l'intégration d'une technologie : les connaissances liées aux contenus à enseigner, celles sur la pédagogie et celles sur la technologie (Mishra et Koehler, 2006). Les auteurs soulignent (Koehler et Mishra, 2014) que l'intégration ne se définit pas par la quantité ou le type de technologie utilisée, mais par la façon dont elle est utilisée et les raisons pour lesquelles elle l'est. Il ne faut pas oublier l'importance de l'improvisation pédagogique et de la collaboration (avec les étudiants) dont peuvent faire preuve les enseignants. Le développement de nouvelles approches pédagogiques, utilisant de nouveaux outils, est davantage un processus additif qu'un processus de remplacement « de l'ancien par le nouveau ». Ainsi, les enseignants qui utilisent et intègrent de la technologie dans leur pédagogie, comme par exemple l'intégration de l'IA, contribuent à développer ce processus et à l'étendre.

Moersch (1995) propose un cadre de référence pour les différents niveaux de mise en œuvre de la technologie par les enseignants dans les salles de classe. Il en propose une version révisée (Moersch, 2010) afin de prendre en compte les nouvelles approches pédagogiques ainsi que les nouveaux outils numériques à disposition. Sept niveaux d'appropriation sont mis en évidence, du niveau 0 qui correspond à la non-utilisation de ces outils, au niveau 6, qui correspond à l'appropriation la plus aboutie des outils, basée sur une collaboration régulière et une approche centrée sur l'apprenant.

Le modèle SAMR (*Substitution, Augmentation, Modification, Redéfinition*) (Puentedura, 2013 ; Hamilton *et al.*, 2016) qui distingue quatre niveaux d'intégration des outils technologiques et de leurs usages, peut également être mobilisé. Utilisé dans l'enseignement primaire et secondaire, ce modèle, qui se présente sous la forme d'une échelle, encourage les enseignants à passer d'un niveau inférieur au niveau supérieur (correspondant à un niveau d'intégration de la technologie plus élevé et à un niveau d'enseignement amélioré). En ce qui concerne le premier niveau, la substitution, les enseignants utilisent de nouveaux outils, mais sans aucun changement fonctionnel (un document papier

est remplacé par un fichier numérique). Lors du deuxième niveau, l'augmentation, de nouveaux outils sont utilisés ce qui entraîne une pédagogie différente (ouverture d'un forum pour augmenter l'interactivité au sein du cours). Ces deux premiers niveaux correspondent à de l'amélioration (selon Puentedura). Le troisième niveau, la modification, concerne l'usage de technologies qui nécessitent de repenser les tâches de manière significative au cours desquelles l'apprenant est amené à avoir un rôle actif (un cours classique est remplacé par la réalisation d'un exposé ou l'utilisation d'un jeu de rôle). Enfin, le niveau de redéfinition est atteint lorsque la technologie est utilisée pour créer de nouvelles tâches qui n'étaient pas concevables antérieurement (imaginer son cours grâce à l'IA et à la *gamification* par exemple). Ces deux derniers niveaux correspondent à une transformation (selon Puentedura).

Dans leur étude traitant de l'intégration du numérique dans les écoles élémentaires, Besneville *et al.* (2019) identifient huit profils d'enseignants, allant de l'enseignant « intégrateur global », le profil 1, qui tire profit du numérique tant de l'équipement que des ressources, à l'enseignant « non-utilisateur », le profil 8, qui n'utilise ni les équipements ni les ressources numériques dans sa pratique d'enseignant.

Enfin, Collet *et al.* (2024) se sont intéressés à l'appropriation de ChatGPT par les enseignants et ont mis en évidence quatre stades d'appropriation (sur le modèle MTT, Prochaska, 1979 ; Prochaska et DiClemente, 1983). Tout d'abord la pré-contemplation correspond aux enseignants qui ne se sentent pas concernés par cet outil ; il n'y a pas de remise en cause des pratiques actuelles, ChatGPT est synonyme de perte de temps et n'est absolument pas intégré. Ensuite la contemplation concerne les enseignants qui ont besoin d'aide pour intégrer cet outil dans leurs pratiques pédagogiques ; ils réclament un accompagnement. Le troisième stade d'appropriation, l'action, concerne des enseignants qui intègrent ChatGPT dans leurs pratiques, qui font des efforts pour prendre en compte ce nouvel outil. Enfin, le maintien correspond aux enseignants qui ont intégré ChatGPT mais qui sont finalement déçus des résultats obtenus et pensent à l'abandonner.

Seule une méthodologie qualitative, basée sur des entretiens, permet de garantir une approche réflexive permettant de mettre à jour ces différentes pratiques.

## MÉTHODOLOGIE

Afin d'appréhender au mieux les facteurs influençant l'appropriation de ChatGPT par les enseignants, une approche qualitative a été privilégiée en menant des entretiens semi-directifs en face-à-face. L'échantillon est composé de 30 enseignants<sup>1</sup> exerçant dans différents niveaux d'enseignement universitaire et disciplines académiques au sein d'établissements publics et privés en France métropolitaine. Les participants ont été sélectionnés au fur et à mesure selon un échantillonnage raisonné visant à maximiser la diversité des profils et à respecter la représentativité qualitative (Paillé et Mucchielli, 2021). Les répondants sont âgés de 25 à 69 ans, et présentent des caractéristiques distinctives en termes de genre, d'ancienneté, de statut, de domaine de compétence ou encore d'appétence envers les technologies (voir le Tableau 1).

L'ensemble des données qualitatives (verbatim retranscrits) a fait l'objet d'une analyse de contenu (Bardin, 2013 ; Paillé et Mucchielli, 2021) en suivant une approche inductive. Les données de l'ensemble du *corpus* ont été balisées et analysées à l'aide du logiciel d'analyse qualitative 3QI et d'une analyse thématique.

---

<sup>1</sup> Est considéré comme enseignant toute personne assurant des cours auprès d'un public étudiant. Il existe différents statuts d'enseignants : les enseignants titulaires, les enseignants vacataires (ayant souvent une autre activité professionnelle), les enseignants-chercheurs (titulaires) et les enseignants contractuels. Les enseignants vacataires apportent à la formation les connaissances professionnelles.

L'examen approfondi des verbatims, par les chercheurs et par le logiciel a permis d'identifier, au regard de leurs perceptions de leurs attitudes, des profils types d'enseignants utilisateurs.

## DES PROFILS CONTRASTÉS

Notre étude a mis en évidence une diversité de profils d'enseignants face à cette innovation technologique, allant des pionniers enthousiastes aux technophobes réticents. Elle a montré une différence notable entre les enseignants et les professionnels qui interviennent dans la formation. En effet, les intervenants professionnels ont tendance à rechercher davantage de productivité et d'efficacité et d'adopter plus facilement tout nouvel outil leur permettant de gagner du temps. Par ailleurs, l'analyse a montré une relation très significative entre l'âge et l'appropriation de l'outil, les enseignants plus jeunes ont tendance à s'approprier et à profiter de l'outil plus facilement. Cette hétérogénéité des attitudes appelle une approche différenciée et nuancée dans les stratégies d'adoption.

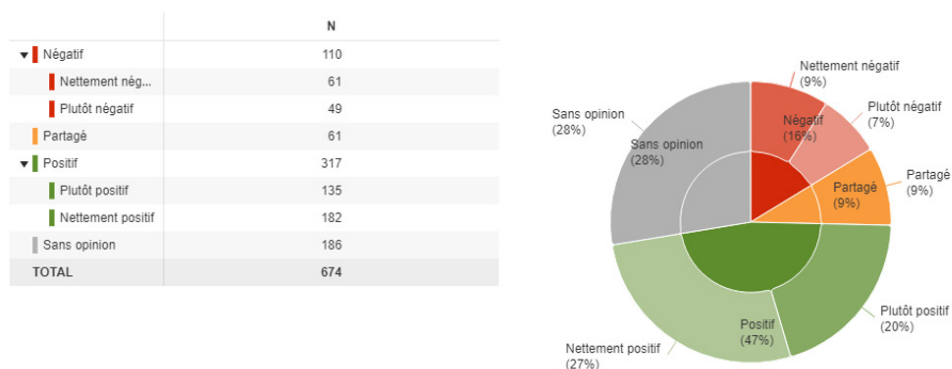


Figure 1 : Les résultats de l'analyse des orientations par le logiciel IQ3 (Source : Sphinx).

## Les pionniers enthousiastes

Les pionniers enthousiastes sont les plus enclins à adopter rapidement et pleinement ChatGPT. Ces enseignants sont généralement très à l'aise avec les technologies et sont souvent les premiers à adopter de nouvelles innovations pédagogiques. Ils expérimentent régulièrement de nouvelles façons d'utiliser l'outil pour optimiser leur enseignement. Ils perçoivent une forte utilité de l'outil et essaient d'explorer son potentiel. Leur attitude est très positive envers ChatGPT dès le départ. Ils utilisent ChatGPT, comme assistant, pour automatiser certaines tâches, comme l'organisation des *e-mails* ou la création de contenu pédagogique. Ce type de profil est à l'aise techniquement et se préoccupe essentiellement de la qualité et de la précision des réponses de ChatGPT ainsi que des questions éthiques d'utilisation de l'IA par les étudiants. Leurs intentions sont de profiter pleinement de l'outil et de l'intégrer de façon à obtenir de meilleurs résultats éducatifs.

## Les pragmatiques prudents

Les pragmatiques prudents adoptent ChatGPT de manière plus prudente, après avoir évalué ses bénéfices concrets. Ils reconnaissent les avantages de ChatGPT mais préfèrent en limiter l'usage aux tâches qu'ils jugent sans risque. Ils l'utilisent surtout pour des véri-



Tableau 1 : Descriptif de l'échantillon (Sources : auteurs).

Numéro	Sexe	Date de l'entretien	Statut	Domaine de compétence	Tranche d'âge
1	M	21/03/2024	Titulaire, EC*	Physique	50-60
2	F	09/04/2024	Vacataire, formatrice	Distribution, <i>merchandising</i>	40-50
3	M	29/03/2024	Vacataire, chef d'entreprise	<i>Marketing</i> digital	40-50
4	F	21/03/2024	Titulaire, EC	Biologie	50-60
5	M	19/03/2024	Titulaire, EC	Économie	50-60
6	M	19/03/2024	Titulaire, professionnel	Négociation, <i>management</i>	40-50
7	F	21/03/2024	Titulaire, EC	Chimie, environnement	50-60
8	F	29/03/2024	Vacataire, chef d'entreprise	Graphiste	40-50
9	M	18/03/2024	Titulaire	<i>Marketing</i>	60-70
10	F	21/03/2024	Titulaire	Expression, communication	50-60
11	M	19/03/2024	Titulaire, professionnel	Droit du travail	40-50
12	M	12/04/2024	Vacataire, chef d'entreprise	<i>Marketing</i> digital	40-50
13	F	12/04/2024	Vacataire, professionnel	Conseil, orientation	50-60
14	F	12/04/2024	Titulaire, EC, externe privé	Logistique	50-60
15	M	05/06/2024	Titulaire, EC, externe public	Physique médicale	50-60

Numéro	Sexe	Date de l'entretien	Statut	Domaine de compétence	Tranche d'âge
16	M	05/04/2024	Titulaire, professionnel	Communication	30-40
17	M	05/04/2024	Vacataire, chef d'entreprise	Communication	50-60
18	M	18/03/2024	Titulaire	Gestion	50-60
19	F	23/04/2024	Titulaire, EC	Mathématiques	30-40
20	M	02/04/2024	Titulaire, EC	Finance	50-60
21	F	25/04/2024	Titulaire, EC externe public	Management, gestion	30-40
22	M	20/03/2024	Vacataire, professionnel	Informatique	60-70
23	M	25/03/2024	Titulaire, EC	Économie	50-60
24	F	23/04/2024	Titulaire, EC	Économie	40-50
25	M	16/04/2024	Titulaire, EC externe privé	Mathématiques numériques	20-30
26	M	16/04/2024	Titulaire, EC externe privé	Mathématiques numériques	20-30
27	F	16/04/2024	Titulaire, EC externe privé	Géographie	20-30
28	M	16/04/2024	Titulaire, EC externe privé	Informatique	30-40
29	M	16/04/2024	Titulaire, EC externe privé	Informatique	20-30
30	M	04/07/2024	Titulaire, EC externe privé	Informatique	30-40

\* EC : enseignant-chercheur.

fications rapides, des recherches d'informations ponctuelles, ou encore pour des révisions et simplifications de contenus. Leur perception de l'utilité dépend des retours d'expérience positifs qu'ils ont reçus. Ils sont conscients des avantages de l'outil mais restent vigilants quant à l'impact sur la réflexion critique des étudiants. Ils ont une approche prudente et sélective de l'outil. Souvent neutres au départ, leur adoption sera progressive, en fonction des normes subjectives dominantes dans leur environnement (pression des collègues, adjonctions administratives, etc.) et au fur et à mesure que l'outil fait ses preuves. Ils soulèvent néanmoins des interrogations sur son éthique et sa maîtrise et la capacité de ChatGPT à remplacer l'interaction humaine de l'enseignant. Leurs intentions sont d'améliorer l'efficacité de leurs enseignements sans prendre de risques majeurs.

### Les sceptiques réservés

Les sceptiques réservés sont souvent des enseignants plus expérimentés. Ils doutent de l'utilité réelle de ChatGPT et craignent que cela ne complique plus qu'autre chose leur travail. Ils sont plus critiques envers ChatGPT. Pour eux, l'outil représente une menace potentielle pour l'autonomie intellectuelle des étudiants et pour la sécurité des informations partagées, mais restent ouverts à l'idée de l'utiliser à condition que des protocoles stricts soient mis en place pour garantir l'éthique et la sécurité. Ces enseignants sont généralement moins à l'aise avec la technologie et leur perception de sa facilité d'utilisation est assez faible. Ils limitent leur utilisation de ChatGPT à des tests exploratoires ou à des activités très spécifiques. Leur attitude initiale est plutôt minimale et expérimentale, nécessitant des preuves solides pour les convaincre. Leur réticence peut être due à un manque de compétences technologiques (non-maîtrise des *prompts*), une résistance au changement (nouvelles méthodes pédagogiques) ou une opposition philosophique (la peur de l'utilisation des données personnelles) ou éthique (plagiat et propriété intellectuelle). Ils ne sont pas opposés par principe mais sont très préoccupés par les risques éthiques et la nécessité d'un cadre de régulation strict. Ils craignent la dépendance et la disparition de l'interaction.

### Les technophobes

Les technophobes (les anti-chatGPT) remettent en cause l'utilité même de ChatGPT qu'ils jugent contraire à une pédagogie humaine. Ils se méfient de l'outil, le considérant comme une menace pour leur profession et pour les étudiants. Ce type d'enseignants préfère ne pas utiliser ChatGPT, ni dans leur vie personnelle, ni dans leur pédagogie, par souci de confidentialité et pour préserver la qualité de l'enseignement. Ils évitent la technologie en général et se contentent des outils minimums. Ils privilégient les pratiques pédagogiques traditionnelles et n'ont aucune confiance en la technologie. Ils craignent la perte de contrôle sur la pédagogie, la disparition du métier et du statut et l'impact négatif sur la qualité de l'apprentissage. Leurs recherches se font toujours à la bibliothèque. Leurs intentions sont de garder le contrôle sur leurs méthodes, qu'elles jugent testées et prouvées. Ces enseignants n'ont pas envie de changer leurs méthodes d'enseignement.

Les témoignages présentés dans l'Encadré situé page suivante illustrent la perception partagée par plusieurs participants concernant l'utilisation de ChatGPT en éducation.

## RECOMMANDATIONS ET CONCLUSION

Cette segmentation en profils distincts n'est pas figée. Certains enseignants peuvent évoluer d'un profil à un autre selon leur formation, leurs expériences avec l'outil et selon le contexte institutionnel. Cela nécessite cependant d'adapter les stratégies de déploiement de ChatGPT, de la communication à la formation, pour faciliter son adoption et appropriation par les différents groupes d'enseignants.

### Verbatims

« Pour moi, ChatGPT est comme un assistant, je peux lui poser des questions ou lui demander de reformuler des explications pour mes étudiants » (E2)

« Pour ma part, j'ai intégré les IA génératives un petit peu dans toutes les sphères, perso pro je l'intègre sans problème » (E19)

« J'utilise ChatGPT pour la correction de mes *e-mails* et la préparation de mes cours, c'est un gain de temps énorme » (E3)

« ChatGPT est utile, mais je m'en sers seulement pour de petites vérifications. Pour mes cours, je ne veux pas que mes étudiants deviennent dépendants » (E2)

« Je permets l'utilisation de ChatGPT seulement dans des contextes bien définis, où les étudiants ne peuvent pas se contenter de copier-coller » (E4)

« ChatGPT m'intéresse, mais il doit y avoir des règles claires pour que son usage soit éthique et sans risque pour les données » (E4)

« On ne dit pas aux étudiants de ne pas l'utiliser, l'étudiant peut l'utiliser, mais pas dans des situations où l'étudiant est censé faire tourner sa propre intelligence intellectuelle » (E29)

« Moi, je suis vieux jeu. Pour moi, ChatGPT n'a pas sa place dans ma pédagogie » (E1)

« Je refuse de l'utiliser, je ne veux pas que mes données ou celles de mes étudiants soient capturées par des serveurs étrangers » (E4)

« Pour moi ça ne remplace pas, ça nous aide » (E16)

« Pour moi, ça n'a d'intelligence que de nom. C'est abusif pour moi c'est restrictif même en plus c'est jamais qu'un logiciel de pillage ou de copiage » (E18)

« Personnellement je vois que les étudiants même pour un exercice en classe pour des questions qui font directement référence aux documents fournis, les étudiants utilisent ChatGPT au lieu de chercher par eux-mêmes » (E21)

« J'utilise par exemple pour donner des cours, j'ai déjà utilisé ChatGPT en direct dans des cours pour trouver des petites réponses, pour montrer des cas d'utilisation aux étudiants » (E26)

« Et pour la pédagogie j'ai pu utiliser ChatGPT ou d'autres outils d'IA pour faire des QCM, par exemple créer des QCM à partir de mes supports de cours et ça j'ai pas trouvé cela très concluant » (E27)

« Si on ne comprend pas son fonctionnement on risque de lui attribuer trop de pouvoir, trop de crédibilité » (E28)

En mettant en place des stratégies diversifiées et des solutions adaptées aux besoins spécifiques de chacun, les enseignants peuvent être efficacement soutenus dans cette transition technologique. Il est d'ailleurs nécessaire d'adapter les modèles d'adoption aux spécificités de l'IA générative (Venkatesh, 2022 ; Mourtajji et Arts-Chiss, 2024). Parallèlement, la sensibilisation des étudiants, non seulement aux opportunités offertes par ChatGPT, mais aussi à ses limites, constitue un levier essentiel pour une intégration réfléchie de cet outil. Toutefois, la création d'un cadre équitable et durable pour l'utilisation de ChatGPT ne repose pas uniquement sur les enseignants. Elle nécessite une approche collaborative impliquant développeurs, éducateurs, décideurs politiques et chercheurs. L'objectif est de positionner ChatGPT comme un allié pédagogique plutôt qu'une source de tensions ou de résistances au sein de la communauté éducative.

Tableau 2 : Les stratégies à adopter pour une intégration éthique (Sources : auteurs).

<b>Stratégies communes</b>	<ul style="list-style-type: none"> <li>• Mettre en place une charte éthique au sein de chaque établissement ;</li> <li>• Mettre en place une plateforme pour le partage d'expériences d'utilisations et de tutoriels ;</li> <li>• Organiser des formations régulières en fonction de l'évolution de l'outil.</li> </ul>
<b>Pionniers</b>	<ul style="list-style-type: none"> <li>• Les encourager en leur offrant des accès prioritaires qui leur permettent de découvrir les nouvelles fonctionnalités ;</li> <li>• Leur donner l'occasion de partager leurs expériences ;</li> <li>• Les responsabiliser pour superviser d'autres nouveaux utilisateurs.</li> </ul>
<b>Prudents</b>	<ul style="list-style-type: none"> <li>• Proposer des formations concrètes en lien avec leurs domaines ;</li> <li>• Les rassurer en leur proposant toutes les informations nécessaires au sujet de la sécurité des données, du plagiat.</li> </ul>
<b>Sceptiques</b>	<ul style="list-style-type: none"> <li>• Proposer des ateliers de discussions pour lever le voile sur les freins de ChatGPT ;</li> <li>• Montrer des études de cas réels sur l'impact positif de chatGPT sur l'apprentissage des étudiants.</li> </ul>
<b>Technophobes</b>	<ul style="list-style-type: none"> <li>• Proposer des formations sur l'usage de la technologie en général ;</li> <li>• Proposer des formations avec un suivi personnalisé ;</li> <li>• Proposer des applications ludiques de ChatGPT dans un premier temps afin de gagner leur confiance.</li> </ul>

## RÉFÉRENCES BIBLIOGRAPHIQUES

BARDIN L. (2013), *L'analyse de contenu* (2<sup>e</sup> éd.), Paris, France, PUF.

BESNEVILLE E., BRILLANT C., CAESAR M., CERISIER J.-F., DEVAUCHELLE B. *et al.*, (2019), « Le numérique éducatif à l'école élémentaire en tension entre politiques nationales, politiques locales et logiques d'appropriation par les enseignants », *Écoles, territoires et numérique : quelles collaborations ? quels apprentissages ?*, Oct., Clermont-Ferrand, France, hal-02314186.

BOURDET J. F. & LEROUX P. (2009), « Dispositifs de formation en ligne : de leur analyse à leur appropriation », *Distances et savoirs*, 7(1), pp. 11-29.

COLLET P., MEINER S., SOMAT A. & SILONE F. (2024), « Élaboration d'une échelle de mesure des stades d'appropriation de ChatGPT par les enseignant-es », MARSOUIN, Mai, 10.1016/j.jenvp.2013.01.002255.hal-04595823

- COLLIN S. & MARCEAU E. (2022), « Enjeux éthiques et critiques de l'intelligence artificielle en enseignement supérieur », *Éthique publique, Revue internationale d'éthique sociétale et gouvernementale*, Vol. 24, n°2.
- HAMILTON E. R., ROSENBERG J. M. & AKCAOGLU M. (2016), "The substitution augmentation modification redefinition (SAMR) model: A critical review and suggestions for its use", *TechTrends*, 60, pp. 433-441.
- KOEHLER M. J. & MISHRA P. (2014), "Introducing tpck", In *Handbook of technological pedagogical content knowledge (TPCK) for educators* (pp. 13-40), Routledge.
- KOEHLER M. J. & MISHRA P. (2005), "What happens when teachers design educational technology? The development of technological pedagogical content knowledge", *Journal of Educational Computing Research*, 32(2), pp. 131-152. <https://doi.org/10.2190/OEW7-01WB-BKHL-QDYV>
- MISHRA P. & KOEHLER M.J. (2006), "Technological pedagogical content knowledge: A framework for teacher knowledge", *Teachers College Rec.*, 108, pp. 1017-1054.
- MOERSCH C. (2010), "LoTi turns up the heat!", *Learning & Leading with Technology*, 37(5), pp. 20-23.
- MOERSCH C. (1995), "Levels of technology implementation (LoTi): A framework for measuring classroom technology use", *Learning & Leading with Technology*, 23, p. 40.
- MOURTAJJI L. & ARTS-CHISS N. (2024), "Unleashing ChatGPT: Redefining technology acceptance and digital transformation in higher education", *Administrative Sciences*, 14(12), 325, <https://doi.org/10.3390/admsci14120325>
- PAILLE P. & MUCCHIELLI A. (2021), *L'analyse qualitative en sciences humaines et sociales* (5<sup>e</sup> éd.), Armand Colin.
- PAQUELIN D. (2009), *L'appropriation des dispositifs numériques de formation : du prescrit aux usages*, L'Harmattan.
- PUENTEDURA R. (2013), "SAMR and TPCK: An introduction", Hipassus, blogue de Ruben R. Puentedura, [http://www.hipassus.com/rrpweblog/archives/2013/03/28/SAMR andTPCK\\_AnIntroduction.pdf](http://www.hipassus.com/rrpweblog/archives/2013/03/28/SAMR_and_TPCK_AnIntroduction.pdf)
- PROCHASKA J.O. & DiCLEMENTE C.C. (1983), "Stages and processes of self-change of smoking: Toward an integrative model of change", *Journal of Consulting and Clinical Psychology*, 51(3), pp. 390-395.
- PROCHASKA J.O. (1979), *Systems of psychotherapy: A transtheoretical analysis*, Dorsey.
- TRINQUECOSTE J. F. & BIDAN M. (2011), « Regards croisés sur le processus d'appropriation des Technologies de l'Information et de la Communication », *Management & Avenir*, 45(5), pp. 175-178.
- VENKATESH V. (2022), "Adoption and use of AI tools: A research agenda grounded in UTAUT", *Annals of Operations Research*, 308(1), pp. 641-652.

# IA et libertés : un défi pour la régulation

Par Marie-Laure DENIS

Présidente de la Commission nationale  
de l'informatique et des libertés (Cnil)

Pour encadrer les systèmes d'intelligence artificielle, l'Union européenne a renforcé sa régulation en introduisant le Règlement sur l'Intelligence Artificielle (RIA), en complément d'autres règlements, tel le Règlement Général sur la Protection des Données (RGPD). Ces textes visent à garantir la transparence et la sécurité des systèmes d'IA, le premier en limitant les risques systémiques, et le second à travers la protection des droits individuels, en protégeant la vie privée.

Ce cadre légal entend aussi prévenir certaines atteintes aux libertés fondamentales, comme l'expansion de la surveillance ou l'utilisation excessive de décisions automatisées. Toutefois, la régulation de l'IA nécessite une coopération étroite entre régulateurs, concepteurs et utilisateurs pour garantir une innovation technologique respectueuse des droits fondamentaux et des valeurs sociales.

## INTRODUCTION : UNE SOCIÉTÉ FAÇONNÉE PAR LA DONNÉE

L'intelligence artificielle (IA) s'est introduite dans de nombreux aspects de la vie quotidienne, transformant les interactions numériques individuelles, et à plus large échelle les pratiques professionnelles et scientifiques. Ses bénéfices sociaux sont déjà notables. Par exemple, dans la recherche génomique, l'IA aide à analyser des mutations génétiques pour mieux comprendre des maladies complexes telles que le cancer. Sur le plan économique, sa diffusion croissante pourrait générer des gains de productivité en réduisant les tâches répétitives, libérant ainsi du temps pour des activités à plus forte valeur ajoutée et créant de nouveaux emplois liés à l'utilisation de ces technologies.

Néanmoins, l'utilisation massive des systèmes automatisés à base d'intelligence artificielle soulève des questions de transparence, fait craindre une perte d'autonomie et d'expertise, et ainsi l'atteinte à certaines libertés, notamment lorsque les décisions humaines sont assistées par des algorithmes susceptibles d'exercer des biais discriminatoires. En outre, l'impact de l'IA sur l'organisation du travail pourrait s'avérer important, avec la possibilité de remplacement de certains emplois par des systèmes automatisés, engendrant des tensions économiques et sociales. À une échelle plus individuelle, la collecte massive de données, en particulier personnelles, soulève de sérieuses inquiétudes en matière de vie privée. Car ces outils d'IA sont d'autant plus efficaces qu'ils « consomment » des données, ce qui pose des défis complexes pour respecter les principes de confidentialité.

Ainsi, le recours à l'IA n'est pas sans risques, et appelle un cadre de régulation, ce que l'Union européenne a entendu en proposant un règlement dédié, en complément des régle-

mentations existantes. L'angle de la protection des données permet, déjà, de capturer un grand nombre des aspects associés aux risques individuels que présente l'IA.

## **LA PROTECTION DES DONNÉES PERSONNELLES, UNE PERSPECTIVE NÉCESSAIRE À LA RÉGULATION DE L'IA**

### **Cadre légal : enjeu d'articulation des textes (RGPD et règlement IA)**

En novembre 2022, Open AI a mis à la disposition du grand public une IA générative, ChatGPT, capable d'interagir avec des usagers en produisant des réponses relativement sophistiquées. Après un bref effet de sidération, un écosystème protéiforme s'est attelé à la création de normes et de guides pour partie déjà amorcées avant l'arrivée de ChatGPT. Les « normes » s'entendent ici au sens large puisque les solutions proposées allaient de la création de référentiels à des mesures plus radicales quoique temporaires comme le blocage du déploiement de ChatGPT sur le sol italien<sup>1</sup>. L'Union européenne, ayant initié un projet de règlement pour l'IA en 2021, y a alors intégré les IA génératives, exigeant des obligations spécifiques, comme la transparence sur l'utilisation des données d'apprentissage<sup>2</sup>, ou le respect de certaines conditions liées à l'interaction avec les individus<sup>3</sup>, qui recourent en partie les obligations du Règlement général de la protection des données (RGPD), rendant impératif l'articulation des textes pour soutenir le développement responsable de l'IA.

### **Du marché à la gouvernance : un texte transverse au service des spécificités sectorielles**

Le RGPD indique les limites et les conditions d'une utilisation et d'une libre circulation des données européennes, ressource essentielle au fonctionnement des systèmes d'IA, tandis que le RIA prévoit l'implémentation des mesures réglementaires, en incluant l'extraction de normes harmonisées pour les industriels<sup>4</sup>, structurant la régulation par marchés<sup>5</sup>. Cette particularité impose une coordination inter-régulatoire de différents secteurs ciblés par le texte : la santé, les infrastructures critiques, l'éducation et l'emploi, l'accès aux services privés et publics essentiels...

La protection des données est, elle aussi, transverse, même si la logique du texte ne suit pas une surveillance par secteurs : à ce titre, chaque autorité nationale de protection des données assure le respect du texte, si besoin dans le cadre d'une coopération avec

---

<sup>1</sup> <https://www.garanteproperty.it/home/docweb/-/docweb-display/docweb/9870847>

<sup>2</sup> Règlement (UE) 2024/1689 du Parlement européen et du Conseil du 13 juin 2024 établissant des règles harmonisées concernant l'intelligence artificielle, article 53, [https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=OJ:L\\_202401689](https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=OJ:L_202401689)

<sup>3</sup> Règlement (UE) 2024/1689, article 50, [https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=OJ:L\\_202401689](https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=OJ:L_202401689)

<sup>4</sup> Règlement (UE) 2024/1689, article 40, [https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=OJ:L\\_202401689](https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=OJ:L_202401689)

<sup>5</sup> Règlement (UE) 2024/1689, Annexe III, [https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=OJ:L\\_202401689](https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=OJ:L_202401689)



ses homologues. Cependant, les données personnelles constituent une denrée économique influençant les dynamiques des marchés numériques. En cela, les décisions des régulateurs ont une influence sur le marché, conditionnant l'accès aux bases de données pour l'entraînement des modèles. Cette situation illustre la tension qui existe entre la protection des données et la compétitivité économique, car l'accès aux données représente un atout pour les entreprises déjà bien établies sur le marché européen.

### Leviers du régulateur et création d'un modèle pérenne

Ainsi, l'adaptation des géants technologiques du secteur de l'IA, majoritairement extra européens, aux standards de l'UE, soulève deux enjeux majeurs : d'une part, la mise en conformité de leurs produits avec la législation européenne, et d'autre part, l'adaptabilité de ces produits aux usagers européens.

Or le rôle du régulateur est de veiller à ce que cette adaptation des modèles aux langues et à la culture européennes n'ait pas lieu en dépit de la protection de la vie privée des usagers européens.

Dans ce contexte, les autorités compétentes disposent de différents leviers pour la mise en conformité : des outils coercitifs comme les sanctions et les amendes (le *hard power*). C'est le cas des actions de la Cnil contre une société proposant un outil de reconnaissance faciale à partir d'une banque de données de milliards d'individus<sup>6</sup>, et plus généralement des sanctions entraînées par les violations de données personnelles. Elles misent aussi sur la responsabilisation des fournisseurs *via* un *soft power*, un rôle d'accompagnement, au moyen de dispositifs de bacs à sable<sup>7</sup>, qui permettent de tester les innovations dans un cadre circonscrit. Le régulateur a aussi pour objectif premier de produire du droit souple. C'est le cas de la production de doctrine, comme les fiches IA de la Cnil<sup>8</sup>, analysant l'application du RGPD lors de la conception de systèmes d'IA.

L'application du RGPD dans le contexte de l'IA fait aussi l'objet de décisions collégiales à l'échelle européenne. Ainsi les autorités européennes de protection des données discutent-elles actuellement des conditions admissibles pour la collecte de données à des fins d'entraînement des modèles<sup>9</sup>.

### LES LIBERTÉS À L'ÉPREUVE DU DÉPLOIEMENT DES MODÈLES D'IA : DÉVELOPPER DES SYSTÈMES AU SERVICE DE L'INTÉRÊT GÉNÉRAL

La transition de l'échelle individuelle à celle du groupe constitue une distinction fondamentale entre l'approche du RGPD et celle du RIA. Or la conception d'un système d'IA au service de l'intérêt général repose sur l'articulation de ces deux perspectives.

---

<sup>6</sup> <https://www.legifrance.gouv.fr/cnil/id/CNILTEXT000046444859>

<sup>7</sup> Règlementaires au sens du RIA, règlement (UE) 2024/1689, article 57, [https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=OJ:L\\_202401689](https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=OJ:L_202401689), ou sans dérogation, (<https://www.cnil.fr/fr/bac-sable-intelligence-artificielle-et-services-publics-la-cnil-accompagne-8-projets-innovants>).

<sup>8</sup> <https://www.cnil.fr/fr/les-fiches-pratiques-ia>

<sup>9</sup> [https://www.edpb.europa.eu/news/news/2024/edpb-stakeholder-event-ai-models\\_en](https://www.edpb.europa.eu/news/news/2024/edpb-stakeholder-event-ai-models_en)

## Surveillance et VSA : un enjeu de liberté publique à concilier avec un enjeu de sécurité

La vidéosurveillance algorithmique (VSA) illustre cette tension entre sécurité – bénéfices potentiels pour une communauté de personnes – et libertés – impact à l'échelle individuelle. En effet, ces systèmes visent à détecter des événements anormaux, parfois basés sur le comportement individuel. La Loi JOP<sup>10</sup> posait trois conditions essentielles à leur usage expérimental lors des Jeux Olympiques et Paralympiques de 2024 à Paris : d'abord, l'interdiction de procéder au traitement de données biométriques, en particulier de la reconnaissance faciale ; ensuite, l'impossibilité de recouper ces données avec d'autres traitements de données personnelles. Ces deux premiers critères avaient pour effet d'empêcher l'utilisation d'images identifiantes des citoyens dans l'espace public, ce qui peut conduire à des travers que le règlement IA a également bien identifié<sup>11</sup>, interdisant les systèmes de crédit social, comme celui déployé en Chine<sup>12</sup>. Enfin, la dernière condition était l'absence de décision individuelle automatisée, empêchant par exemple le déclenchement automatique d'appel aux forces de l'ordre, se cantonnant au « signalement d'attention ». La Cnil, en tant qu'autorité de protection des données personnelles a été chargée de suivre ces expérimentations.

## Interactions humain-machines modifiées à l'échelle sociale : enjeux de transparence et d'autonomie

Cette stricte limitation des capacités des systèmes de VSA demandait en somme que ces systèmes ne se substituent pas au jugement humain. En effet, les systèmes d'IA d'aide à la décision présentent des bénéfices significatifs dans l'allègement des tâches administratives (aide à la rédaction, traitement des factures, tri de candidatures en ressources humaines) et le soutien à l'expertise métier. Le risque est, cependant, la perte d'autonomie des professionnels, une réduction des effectifs ou moyens attribués aux employés, – augmentant les attentes en matière de productivité – et enfin l'opacité de la procédure de décision.

Ainsi, l'arrêt de la Cour de Justice de l'Union européenne en décembre 2023<sup>13</sup> a, en application du RGPD, établi que l'outil de "credit scoring" de la société allemande SCHUFA, bien que validé par un employé de la banque, constituait une décision automatisée. Dès lors que ce genre de décisions présente un encadrement renforcé au sein du RGPD, la CJUE a entendu renforcer la protection des données personnelles dans le cadre de l'utilisation d'algorithmes.

<sup>10</sup> LOI n°2023-380 du 19 mai 2023 relative aux jeux Olympiques et Paralympiques de 2024, [https://www.legifrance.gouv.fr/jorf/article\\_jo/JORFARTI000047561989](https://www.legifrance.gouv.fr/jorf/article_jo/JORFARTI000047561989)

<sup>11</sup> Règlement (UE) 2024/1689, article 5.1(c), [https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=OJ:L\\_202401689](https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=OJ:L_202401689)

<sup>12</sup> [https://www.lemonde.fr/idees/article/2020/01/16/le-credit-social-les-devoirs-avant-les-droits\\_6026047\\_3232.html](https://www.lemonde.fr/idees/article/2020/01/16/le-credit-social-les-devoirs-avant-les-droits_6026047_3232.html)

<sup>13</sup> <https://curia.europa.eu/juris/document/document.jsf?jsessionid=E3D1DC708C777FB310636E7197610C45?text=&docid=280426&pageIndex=0&doclang=fr&mode=lst&dir=&occ=first&part=1&cid=4793519>

En février 2024, une décision de la Cour suprême autrichienne<sup>14</sup> a illustré la manière dont l'IA décisionnelle peut engendrer une menace pour les droits et libertés. La Cour a estimé que le système de classification des demandeurs d'emploi du service public de l'emploi en Autriche, qui créait un classement prédictif des demandeurs d'emploi, associant leur profil à des aides ou des services, tombait dans le périmètre de l'article 22 du RGPD, permettant d'en encadrer les impacts.

De son côté, le RIA impose un contrôle humain pour garantir le bon fonctionnement des systèmes à haut risque<sup>15</sup> et prévenir des erreurs trop coûteuses, en repérant les anomalies.

### À l'échelle des utilisateurs, la personnalisation fine repose sur la collecte massive

Le fait d'inclure quelqu'un dans une catégorie « performative », c'est-à-dire dont la description donne lieu à l'attribution de droits et de services, correspond au terme juridique de « profilage » tel qu'on le retrouve dans le RGPD<sup>16</sup>. Ce traitement consiste à analyser automatiquement des données personnelles pour estimer des caractéristiques comme la productivité. Le profilage est soumis aux mêmes restrictions que les décisions automatisées dont il est un cas particulier.

En effet, ces classifications ou ces estimations du comportement ne sont pas anodines, car les systèmes d'IA, en intégrant des caractéristiques personnelles, risquent aussi de reproduire des biais ou des stéréotypes inférés du profil d'une personne, pouvant mener à des discriminations des usagers, comme l'illustre le cas aux États-Unis de l'algorithme de justice prédictive COMPAS. Cet outil, utilisé par les juges, montrait des biais discriminatoires dans le calcul de score de récidives, en influençant la sévérité de la peine<sup>17</sup>. Dans des domaines moins sensibles, ces biais peuvent néanmoins affecter l'efficacité d'un service en fonction de la représentativité d'une catégorie de population dans les données d'entraînement. Pour prévenir de telles dérives, le RIA prévoit des audits destinés à détecter et prévenir les risques de biais<sup>18</sup>.

Au-delà du terme juridique, qui contient l'idée de prédiction du comportement, l'acception commune du terme de « profilage » inclut des formes variées de personnalisation utilisées par les systèmes de recommandation de contenu.

La personnalisation fine des outils numériques du quotidien permet de cibler les besoins spécifiques des utilisateurs, *via* une collecte de données personnelles (habitudes de consommation, santé, situation géographique et familiale, etc.). Pour atténuer ces risques, le RGPD impose transparence, information préalable des utilisateurs et possibilité de refuser le traitement sous la forme d'*opt-out*<sup>19</sup>.

---

<sup>14</sup> [https://gdprhub.eu/index.php?title=VwGH\\_-Ro\\_2021/04/0010-11&mtc=today](https://gdprhub.eu/index.php?title=VwGH_-Ro_2021/04/0010-11&mtc=today)

<sup>15</sup> Règlement (UE) 2024/1689 du Parlement européen et du Conseil du 13 juin 2024 établissant des règles harmonisées concernant l'intelligence artificielle, article 14, [https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=OJ:L\\_202401689](https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=OJ:L_202401689)

<sup>16</sup> Règlement (UE) 2016/679, article 4, [https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=OJ:L\\_202401689](https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=OJ:L_202401689)

<sup>17</sup> <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>

<sup>18</sup> Règlement (UE) 2024/1689, article 10, [https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=OJ:L\\_202401689](https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=OJ:L_202401689)

<sup>19</sup> Règlement (UE) 2016/679, articles 12 (transparence) et 21 (droit d'opposition).

Le RIA complète ce cadre en interdisant strictement les systèmes susceptibles de manipuler, d'exploiter les failles cognitives ou émotionnelles des utilisateurs<sup>20</sup>. En effet ces personnalisations peuvent avoir l'effet délétère d'influencer les comportements, par exemple en incitant à l'achat ou en créant des bulles de filtre, qui enferment l'utilisateur dans un microcosme fondé sur un *consensus*. Ceci interroge l'un des fondements de la démocratie, dans un contexte où la traçabilité de l'information devient toujours plus complexe.

## CONCLUSION

En somme, la préservation de notre modèle de société, fondé sur un équilibre entre respect des droits fondamentaux et innovation, exige une coordination de tous les acteurs concernés, tant sur le plan de la régulation qu'au niveau des interactions entre la régulation et les acteurs économiques. Les régulateurs ne peuvent à eux seuls garantir l'harmonie entre l'innovation technologique et la protection des usagers. La responsabilité doit être partagée entre régulateurs, concepteurs, déployeurs et utilisateurs, chaque acteur jouant un rôle clef dans la chaîne de valeur. Il en va donc de la responsabilité collective de créer et de déployer des technologies de manière à favoriser l'intérêt général, en cherchant à articuler à la fois le progrès économique et la protection des droits fondamentaux.

---

<sup>20</sup> Règlement (UE) 2024/1689, article 5.1(c) à l'exception de certaines pratiques médicales et des pratiques commerciales en publicité déjà autorisées (Considérant 29).

# Le Droit à l'heure de "l'intelligence artificielle"

Par **Didier GUÉVEL**

Professeur émérite de Droit privé et Sciences criminelles  
de l'Université de Sorbonne Paris Nord,  
Membre de l'IRDA Paris XIII (ER 3970),  
Doyen honoraire de la Faculté de Droit, Sciences politiques et sociales,  
Ancien Directeur du Département de Droit de l'Université du Havre,  
Membre associé du LexFEIM Le Havre (EA 1013),  
Membre du Conseil scientifique de la MSH Paris Nord (USR 3258),  
Membre du Comité multi parties-prenantes  
du Forum annuel de la gouvernance d'Internet,  
Profesor Honorario de l'Université de Mendoza.

L'IA présente un intérêt pour les juristes en tant que complément informatif. Déjà les praticiens et les théoriciens du Droit disposent d'IA spécialisées. Restent deux problèmes : celui du contrôle de l'originalité des travaux des étudiants et surtout celui du choix, en amont, des données utilisées.

*Texte achevé le 22 octobre 2024.*

On ne compte plus les études et rencontres consacrées par les juristes à l'intelligence artificielle ; nous ne tenterons pas de les citer ici de peur d'en oublier et de vexer les auteurs et organisateurs de celles que nous aurions pu omettre.

Les législateurs (*lato sensu*) se sont emparés de la question. On songe, bien sûr, notamment, au règlement européen 2024/1689<sup>1</sup>. Mais ce sont plutôt les usages quotidiens du procédé que nous voudrions évoquer.

Nous avons déjà eu l'occasion à plusieurs reprises de dire<sup>2</sup> que ce que l'on nomme l'intelligence artificielle (ci-après IA) ne mérite probablement pas les fantasmes qu'elle développe : des peurs (car chaque siècle doit avoir les siennes) et des gloires publicitaires<sup>3</sup>. Au demeurant, les économistes semblent partager ce point de vue<sup>4</sup>. Il s'agit seulement, selon nous, d'un instrument précieux qu'il faut savoir manier et maîtriser.

Les juristes ont bien sûr été tentés de tester les IA généralistes en les interrogeant sur des sujets dont ils connaissaient tous les pièges. Les résultats ont été à la hauteur de l'humour qu'ils espéraient exercer à leur égard : pitoyables (erreurs de vocabulaire et incompréhensions).

---

<sup>1</sup> HO-DAC M. (2024), « Premier décryptage du règlement européen sur l'intelligence artificielle (AI Act) : vers un standard mondial de l'IA de confiance ? », Dalloz 2024, p. 1678.

<sup>2</sup> Voir par exemple, GUÉVEL D. (2022), « Propos conclusifs » in colloque « Intelligence artificielle et santé » (dir. Guilhem Julia), Campus Condorcet, 22 février 2022, reporté au 22 novembre 2022, <https://fromsmash.com/1sqMtcLYMo-ct>.

<sup>3</sup> Si l'on en croyait les publicitaires, même les produits à récurer ou les crèmes de beauté seraient « issus de l'intelligence artificielle ».

<sup>4</sup> Voir par exemple, DOCTORONW C. (2024), « L'intelligence artificielle est une bulle », propos recueillis par A. Piquard, in *Le Monde*, 1<sup>er</sup>-2 sept. 2024, *Économie & Entreprise*, p. 13.

sion du fond des problèmes). Ces échecs nous semblent avoir deux causes. La première tient peut-être à la circonstance que les IA utilisées étaient d'origine anglo-américaines et étaient donc inaptées à appréhender le vocabulaire français et surtout les concepts du Droit continental. La seconde, plus grave et plus lourde de conséquences, tenait au fait que les données servant à élaborer les résultats étaient uniquement venues de l'Internet et classées en fonction de leur fréquence et de leur répétition et non en fonction de leur pertinence.

Désormais les grands éditeurs juridiques développent leur propre IA à partir de leurs propres données. Le progrès est évident : on aura moins tendance à confondre *Common Law* et droit romano-germanique et, surtout, les sources seront sérieuses. Toutefois, l'on n'obtiendra rien de pertinent tant que ces différentes IA ne seront pas croisées pour offrir aux praticiens (par une voie officielle ?) une information qui ne privilégiera plus tel ou tel éditeur<sup>5</sup> et qui aura su objectivement hiérarchiser les données (même dans les publications juridiques sérieuses, une affirmation tirée d'un simple éditorial peut ne pas avoir la même valeur qu'une remarque contenue dans une analyse doctrinale fouillée et la phrase simplificatrice issue d'un mémento peut ne pas avoir un intérêt identique à celle détaillée issue d'un grand traité). Bien entendu, nous ne pourrions pas aller jusqu'à procéder à une classification des auteurs eux-mêmes, comme la pratique le fait pourtant.

L'IA, comme l'interrogation de l'Internet que les juristes pratiquent depuis l'origine, peut donc avoir un intérêt pour un chercheur ou un praticien, en passe d'achever un travail, afin de vérifier qu'un point du sujet qu'il traite ne lui a pas échappé : une aide et non une substitution.

Outre la question de leur éventuel propre usage de l'instrument, les enseignants du Droit, se sont évidemment posé celle de son utilisation par les étudiants. Pour l'instant le danger est faible.

On l'a dit, les IA généralistes risquent fort de ne rapporter que des mauvaises notes et les IA spécialisées sont coûteuses. Mais il reste bien sûr le cas des étudiants fortunés et celui de ceux pouvant avoir accès à une IA spécialisée payée par un professionnel.

Dans ces occurrences, et pour combattre les fraudes éventuelles, il est probable, si ce n'est déjà fait, qu'à l'image des logiciels débusquant le plagiat, soient rapidement élaborés des logiciels traquant l'usage de l'IA.

Reste un dernier problème : celui de la démocratisation, de la généralisation du recours à l'IA et de son utilisation éventuelle par les justiciables. Il sera malaisé pour un juge d'expliquer au plaideur que la solution "évidemment exacte" que l'IA généraliste a donnée à son litige n'a aucune valeur juridique, est inexacte et que, contrairement à ce qu'elle a pu lui faire accroire, il va perdre son procès... Mais la question n'est pas nouvelle et elle en soulève une autre : celle du bien-fondé d'un recueil généralisé des données juridictionnelles laissées gratuitement à la disposition d'appétits mercantiles.

---

<sup>5</sup> On peut légitimement craindre qu'un éditeur X n'utilisera pas dans son IA les données pourtant plus pertinentes mais publiées chez un éditeur Y.

# Les défis éthiques de la convergence de l'IA, des neurosciences, de l'informatique et de l'ingénierie

Par le Dr Laure TABOUY

Neuroscientifique et éthicienne, dans l'équipe Éthique et Épistémologie  
du CESP-INSERM U1018, Université de Paris-Saclay  
et au Centre Gilles Gaston Granger (UMR 7304), Université d'Aix-Marseille

Les questionnements, défis et dilemmes éthiques, juridiques, philosophiques et théologiques qu'ouvre la convergence des neurosciences, de l'informatique et de l'ingénierie sont colossaux et vertigineux.

Comprendre les enjeux éthiques de l'IA, du numérique et des neurosciences, afin de les appréhender et de les opérationnaliser sur le terrain, demande de comprendre et de connaître l'histoire, le pourquoi et le comment. Même si ces systèmes d'intelligence artificielle offrent de nouvelles possibilités, il n'en demeure pas moins que ce sont des systèmes informatiques, des outils, des machines et qu'ils ne doivent pas être pris pour ce qu'ils ne sont pas.

Cet article s'attache à partager une vision globale sur les enjeux et les défis de cette convergence en mettant en avant l'articulation des enjeux. C'est une ouverture à la réflexion, afin d'approfondir plus spécifiquement les enjeux éthiques selon des usages spécifiques de cette convergence dans des contextes variés.

Les avancées concernant les neurosciences, l'informatique et l'ingénierie sont nombreuses et vertigineuses [1, 2]. Elles permettent, entre autres, de mieux percer les mystères de notre cerveau, depuis ses aspects fondamentaux jusqu'aux aspects les plus fonctionnels et computationnels. Mais qu'est-ce que cette convergence et d'où vient-elle ? Quels enjeux éthiques porte-t-elle ? La convergence des neurosciences, de l'informatique et de l'ingénierie au sens large pose des enjeux, crée des tensions, des questionnements et des dilemmes sans précédent [1, 3, 4]. Ces avancées vont de pair avec le développement de certaines techniques d'exploration du cerveau, invasives ou non invasives : les neurotechnologies. Ces dernières sont des dispositifs se situant à l'interface entre le cerveau et la machine pour visualiser, décrypter, modifier et moduler le fonctionnement cérébral et les dysfonctionnements pathologiques. Leur essor permet d'observer, d'enregistrer et d'intervenir dans le cerveau sans l'endommager et de pratiquer certaines interventions chirurgicales de grande précision.

Cet article s'attache à partager une vision globale sur les enjeux et les défis de cette convergence en introduisant ce qui caractérise, selon moi, ces défis et cette convergence à travers l'histoire. C'est une ouverture à la réflexion, qui se focalise sur les enjeux éthiques selon des usages spécifiques de cette convergence, avec une démarche pédagogique et vulgarisatrice, qui peut servir de base de connaissances pour une réflexion plus poussée

en fonction des dilemmes éthiques identifiés et des finalités voulues de ces réflexions. La partie finale de cet article inclut le partage d'un point de vue plus personnel sur la manière dont nous pourrions être acteurs dans ces débats afin de permettre de garder un équilibre et d'aiguiser notre esprit critique.

## UNE MISE EN PERSPECTIVE HISTORIQUE DE LA CONVERGENCE DES NEUROSCIENCES, DE L'INFORMATIQUE ET DE L'INGÉNIERIE

Afin de comprendre les dilemmes et les tensions qu'ouvre cette convergence, il nous faut plonger dans l'histoire. C'est le résultat d'une longue évolution alliant progrès scientifique et technique, avec la recherche de la simplification des calculs et de la pensée. L'informatique est étroitement liée à la représentation linguistique des informations, des données et des connaissances, ainsi qu'à l'histoire des automates depuis l'Antiquité et de l'horlogerie, en passant par la Pascaline de Blaise Pascal en 1642, automatisée par Gottfried Wilhelm Leibniz. Parallèlement, la pensée philosophique a lié le langage à la pensée, et donc à l'intelligence, ainsi qu'aux émotions et aux sentiments qu'il exprime.

J'identifie cinq vagues historiques successives.

### 1<sup>e</sup> vague

C'est entre les années 1920 et 1950, avec la création d'IBM et le développement des systèmes d'intelligence artificielle, des neurosciences et de la cybernétique lors de 10 conférences Macy, que sonne le début d'une épopée qui va balayer la deuxième moitié du XX<sup>e</sup> siècle jusqu'à aujourd'hui.

Alan Turing, qui permit avec d'autres chercheurs entre 1936 et 1942 de craquer le code de la machine Enigma, a mis au point une machine dite de Turing, un modèle abstrait du fonctionnement des appareils mécaniques de calcul, l'ancêtre de l'ordinateur d'aujourd'hui. Sa publication "Computing Machinery and Intelligence" [5] de 1950, sur le thème de l'intelligence artificielle (terme officiellement né en 1956 lors de la conférence de Dartmouth) avait pour but de répondre à une question qui anime aujourd'hui tous les débats : « Est-ce que les machines peuvent penser ? » et introduit ce que l'on appelle aujourd'hui le test de Turing.

### 2<sup>e</sup> vague

La 2<sup>e</sup> vague se situe entre les années 1950 et 1960. À cette période, les neurosciences, héritières de l'histoire de la médecine, de la neurologie et des débuts de la cybernétique, naissaient puis devenaient la discipline académique que l'on connaît aujourd'hui.

Cette 2<sup>e</sup> vague est surtout le début d'un intérêt grandissant avec quelques théoriciens des systèmes d'intelligence artificielle et quelques timides succès autour des jeux. Les recherches se concentrent dans les laboratoires de recherche, quelques entreprises, dans les mondes des jeux et des applications militaires.

### 3<sup>e</sup> vague

La 3<sup>e</sup> vague se caractérise par l'arrivée des premières entreprises proposant des systèmes d'experts et des robots industriels. Elle se situe entre les années 1965 et 1990. Durant cette période, on retrouve également l'envoi du premier courriel et l'invention du microprocesseur en 1971. C'est en 1983 que Motorola lance son premier téléphone mobile.

Ces systèmes experts sont des logiciels et servent entre autres comme outils d'aide à la décision. La détermination de l'ensemble des faits et règles qui vont composer la base de



connaissances de ces systèmes experts pose des questions éthiques. Comment décrire le comportement d'un expert face à un problème particulier, sa manière de le résoudre, son expérience et sa connaissance pratique ?

### 4<sup>e</sup> vague

Les années 2000-2020 caractérisent la 4<sup>e</sup> vague. La miniaturisation des technologies donnant lieu aux nanotechnologies et aux biotechnologies, et les avancées des neurosciences, des sciences cognitives et de l'informatique donnent ce que l'on appelle la convergence NBIC dans les années 2000, conduisant à ce que l'on appelle aujourd'hui les neurotechnologies.

Mais cette 4<sup>e</sup> vague est surtout marquée par l'accessibilité et la concentration massive de grandes bases de données à partir de 2008-2010, par la puissance informatique et les énormes moyens financiers de certaines entreprises. Le *deep learning* est à son tour mis au-devant de la scène. 1997 est l'année où Deep Blue, la machine d'IBM, bat aux jeux d'échecs le champion en titre Kasparov, suivie de 2002 avec l'introduction du premier robot ménager, iRobot. Les réseaux sociaux se développent entre 2004 et 2017. En 2014, Facebook crée DeepFace un système de reconnaissance faciale capable de reconnaître les visages humains, et 2015 sonne la victoire de DeepMind d'IBM face au champion en titre de l'AlphaGo, Lee Sedol.

Ce premier quart du XXI<sup>e</sup> siècle est réellement marqué par une accélération des innovations et des technologies accessibles au grand public, par la création des réseaux sociaux et leur capacité à capter l'attention des utilisateurs pour en faire un *business*.

### 5<sup>e</sup> vague

La 5<sup>e</sup> vague, initiée en 2020 avec l'arrivée de GPT-3 d'OpenAI, est un tournant bien particulier qui ne ressemble pas aux autres. Elle repose sur toutes les vagues antérieures et est héritière de toute leur histoire.

Les convergences des avancées importantes dans les domaines de l'informatique, des neurosciences, de la biologie, des mathématiques, de la physique, ajoutées à la croissance exponentielle des puissances de calcul et des capacités technologiques, à l'accessibilité massive des données, et à l'internationalisation des échanges entre les individus, donnent à cette vague un visage particulier. Comme toute vague d'innovations majeures, elle enthousiasme par les possibilités qu'elle offre et inquiète par les risques auxquels elle expose.

## DES DÉFIS ?

Les systèmes d'IA, le *machine learning* et le *deep learning* sont historiquement des ouvertures à la façon dont le cerveau fonctionne. Les uns sont imbriqués dans les autres. Ils visent à réaliser des machines capables de simuler, d'imiter l'intelligence humaine et la façon dont l'information circule entre les neurones d'un cerveau humain, et de permettre à des ordinateurs, des machines, des robots de paraître « penser » et « d'agir » comme des êtres humains.

Même si ces systèmes d'intelligence artificielle offrent de nouvelles possibilités, il n'en demeure pas moins que ce sont des systèmes informatiques, des outils, des machines, et qu'ils ne doivent pas être pris pour ce qu'ils ne sont pas. Le *machine learning*, ou apprentissage automatique, repose sur des modèles mathématiques et statistiques, et vise à donner aux machines la capacité d'apprendre à partir de données. Alors que le *deep learning*, ou apprentissage profond, repose sur les neurosciences, la biologie. Ces systèmes imitent la façon dont les êtres humains acquièrent des connaissances et des

savoirs. L'idée est de s'inspirer de la manière dont fonctionne le cerveau humain, avec des réseaux de neurones, pour pousser les analyses et apprendre à extraire des données de façon autonome. Les IA génératives ou LLM sont du *deep learning* et sont des systèmes d'intelligence artificielle capables de générer du texte, des images ou d'autres médias en réponse à des invites, ou *prompts* en anglais. Les invites sont des interfaces en ligne de commande, donc des interfaces homme-machine dans lesquelles la communication entre l'utilisateur et l'ordinateur s'effectue en mode texte.

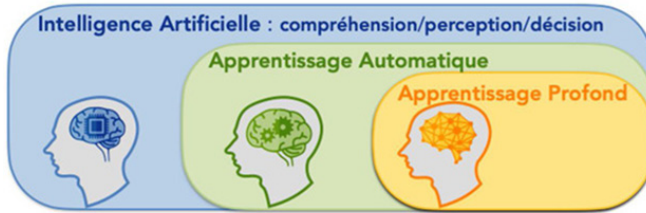


Figure 1 : Intelligence Artificielle, Machine Learning et Deep Learning : Quels liens et quelles différences ?

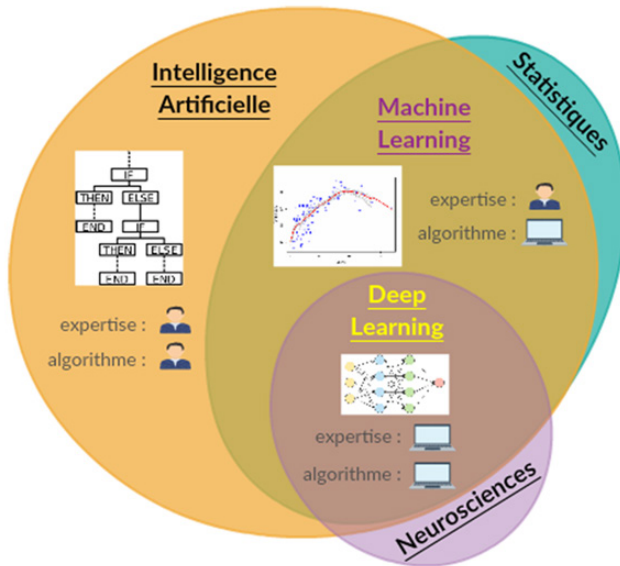


Figure 2 : Interactions entre IA, ML et DL (© Pensée Artificielle, <https://www.geekmaispasque.com/2019/10/intelligence-artificielle-machine-learning-deep-learning-liens-differences/?cn-reloaded=1>).

Mais la crainte que des IA soient capables de manipuler les humains, d'accéder à des systèmes d'armes, d'exploiter des failles de cybersécurité, voire peut-être d'acquérir une forme de conscience et/ou de devenir incontrôlables au point de menacer l'existence de l'humanité est une réalité. Les progrès du numérique, des systèmes d'IA et des neurosciences ouvrent de nouveaux dilemmes pour les droits de l'homme : le droit de garder ses

pensées privées, le droit à la liberté de pensée et remettent en cause le concept juridique de libre arbitre, et donc les fondements de la responsabilité juridique. Ce qui est en jeu, c'est l'accessibilité probable au fonctionnement de notre cerveau et donc à notre intime, à ce que nous sommes au plus profond de nous, à ce qu'il y a de plus intime et de plus secret et ce qui fait notre humanité, notre vulnérabilité, notre dignité, notre fragilité humaine. Cette convergence nous invite à nous questionner sur notre propre humanité, sur ce que cela veut dire être un être humain dans ce monde dans lequel nous vivons.

Tous ces progrès depuis les années 1950 laisseraient entrevoir la possibilité qu'un jour nous puissions décrypter les rêves, l'esprit, la conscience, « télécharger » l'esprit sur un support numérique... de lire la conscience, les intuitions, les intentions, les pensées [7, 8], et animent le fantasme d'une « IA consciente » [9, 10, 11]. La mise en lumière de mécanismes cérébraux de plus en plus précis et spécifiques entrouvre un autre rapport au cerveau, au monde et rouvre la question de la découverte de la signature neuronale de la conscience [12] qui échappe pourtant encore aux neuroscientifiques.

Au gré de leur invention, élaboration et de leur histoire, les machines, la technologie informatique, les ordinateurs, envisagés comme de véritables « organismes » artificiels, ont fourni aux êtres humains des modèles heuristiques et souvent inédits d'intelligibilité du vivant. La conscience, les pensées, l'intelligence humaine ont longtemps été liées à des zones cérébrales spécifiques, comme le cortex préfrontal par exemple. Mais de nouvelles recherches remettent en question ce point de vue, offrant une perspective plus large [13, 14] : l'intelligence et la conscience ne sont pas confinées à des régions cérébrales localisées et précises, ce sont des propriétés globales de l'ensemble du cerveau. Elles proviennent de connexions distribuées à l'échelle du cerveau dans son ensemble, plutôt que de régions spécifiques localisées. Le cerveau humain est l'organe le plus mystérieux du corps humain, le centre de contrôle de notre corps, c'est notre boîte noire, le circuit électrique le plus puissant au monde. Pas moins de 200 milliards de cellules (neurones et cellules gliales) et  $10^{15}$  synapses communiquent, échangent, régulent les informations qui circulent dans notre cerveau. Il traite les informations sensorielles et nous permet, entre autres, d'agir, de former des pensées, de prendre des décisions et d'emmagasiner des connaissances. Compte tenu des capacités de notre cerveau, il semble presque paradoxal que nous en sachions encore si peu, pourtant, il aurait une capacité de stockage des données cérébrales de l'ordre du pétaoctet [15, 16], avec une optimisation de l'énergie inégalée. Ce qui fait que notre cerveau reste aujourd'hui la « machine » la plus puissante et la plus efficace au monde. La plasticité synaptique et neuronale et son adaptabilité en temps réel aux signaux et aux nombres de transmissions d'informations qu'il reçoit sont des propriétés singulières qui permettent au cerveau d'être très économe et très puissant. Ces recherches repoussent les limites des théories traditionnelles sur la conscience et l'intelligence, les interrogent, les bousculent, et montrent que cette convergence souligne leur complexité, faisant allusion à des mécanismes neuronaux inexplorés derrière les capacités cognitives.

Plus nous nous approchons des rouages de cette « boîte noire » qu'est notre cerveau, qui fait de nous des êtres humains singuliers, dotés de conscience, plus la frontière entre la réalité des recherches en neurosciences, des possibles envisagés et les fantasmes d'un homme augmenté et amélioré est de plus en plus mince. Ces systèmes viennent interroger le sens de ce que cela veut dire être un être humain et nous mettre en face de notre humanité. Car la trame de fond de ces innovations en fin 2024, c'est la réalité de l'homme augmenté (qui existe déjà à bien des égards dans le domaine médical, la frontière entre l'homme « réparé » et l'homme « augmenté » étant très mince), et le fantasme d'une société transhumaniste par la mise en commun des différentes technologies et des données collectées.

Tout cela légitime la nécessité de considérer les craintes au regard des évolutions de cette convergence et de s'interroger sur la signification de ces concepts et sur la portée des recherches sur nos sociétés. Mettre en garde contre une façon de raisonner qui exprime

son attente des avancées de ces innovations qui pourraient avoir des effets nuisibles au regard de ces possibilités fantasmées sur l'humanité est plus que nécessaire. Si l'on pousse à l'extrême et bien au-delà de leurs réalisations concrètes les possibilités qu'on leur fantasma, il y a d'excellentes raisons de réfléchir aux enjeux éthiques, juridiques et politiques de l'IA et des neurosciences.

## L'ÉTHIQUE PEUT-ELLE ÊTRE LA PIERRE ANGULAIRE POUR UN FUTUR SOUHAITABLE ?

Nous naviguons entre espoirs et craintes, tandis que ces innovations continuent de progresser et d'envahir notre quotidien. Tous les domaines de la société sont en train de connaître de profondes mutations. Il y a du bon dans ces innovations et dans ces recherches. Ce sont de puissants outils mis au service de la médecine humaine et du citoyen, favorisant des projets et des innovations. Ils sont porteurs de promesses, de rêves, d'espoirs fantastiques et révolutionnaires avec des avantages et des potentiels incroyables. Et en même temps, ils peuvent représenter un risque et même un danger pour l'humanité et l'être humain, suscitant aussi des inquiétudes, car il est impossible de savoir comment les humains réagiront face à un monde en profond bouleversement.

La question de l'éthique nous entraîne donc dans un exercice périlleux, en nous invitant à prendre de la hauteur sur ces enjeux qu'ouvrent les neurotechnologies : pouvons-nous raisonner, penser méthodiquement et consciemment dans un monde d'incertitude ?

Les réflexions éthiques et philosophiques liées à ces innovations s'expriment en général par le biais de questionnements, d'intuitions... Cela nécessite une mise en débat des parties prenantes, de tenir compte des conséquences des décisions, des finalités des actions et des décisions, des contextes dans lesquels les décisions sont prises, des principes, des valeurs, de l'histoire et des cultures, des pays, des individus qui les composent. De telles réflexions, essentielles à l'innovation responsable, permettent alors de poser factuellement les choses et de pouvoir entrer dans un discernement, quant aux usages, ajusté de ces recherches en neurosciences, en informatique et en ingénierie dans chaque situation.

L'éthique et la philosophie permettent de poser les questions, les dilemmes qu'ouvre cette convergence devant les évidences, de creuser les argumentations devant les rapidités de jugement, d'exercer une raison et un esprit critique, de questionner et d'analyser les valeurs, de formuler des chartes, des codes de conduite. C'est une école d'interrogation, de prise de recul, de questionnement et d'argumentation face à l'emprise des passions, des intérêts et des subjectivités. Une prise de hauteur nécessaire pour prendre conscience et ajuster l'agir en vue d'un bien commun, comme le souligne Paul Ricœur : « L'éthique est le mouvement même de la liberté qui cherche une vie bonne, dans la sollicitude envers autrui et dans un juste usage des institutions sociales », l'éthique n'est pas un ensemble de règles à suivre dans le présent pour être bon et parfait, c'est le devoir de chacun de considérer l'impact de ses actions présentes dans le futur, sur les autres et le bien commun.

Il est là pour moi le sens et le rôle de ce pivot réflexif qu'est l'éthique dans ce contexte. Il est d'aider à éclairer les débats, d'inviter à être acteurs en prenant de la hauteur, d'être là où il y a des tensions et des enjeux afin qu'au sein de ces tensions, les questions importantes, les dilemmes éthiques puissent être identifiés et donc être discutés. Cette mise en dialogue des différents points de vue est alors possible afin d'avancer et de prendre des décisions. L'éthique va être la partie prenante qui accompagne depuis l'élaboration d'un projet, d'une innovation jusqu'à sa finalité prévue. Cette éthique par conception permet d'introduire dès le début une réflexion approfondie qui va être une boussole pour une innovation responsable.

L'accélération de ces innovations rend indispensables les réflexions sur les enjeux sociétaux, éthiques et légaux qu'elles soulèvent, ainsi que la conception interdisciplinaire de garde-fous, de systèmes d'évaluation et de surveillance et de cadres de gouvernance adaptés aux valeurs sociologiques, éthiques et juridiques de la France et de l'Europe. En somme, il s'agit d'appeler à des pratiques neuroscientifiques et scientifiques au sens large, plus réfléchies, confortées et mises en débat par les sciences sociales, pour examiner le potentiel et les limites des questions et des méthodologies des neurosciences, de l'IA et de leurs impacts sur la société. Et cela appelle aussi à ce que la philosophie se laisse bousculer par les recherches et les avancées en neurosciences et informatique et qu'elle se laisse interpellé pour qu'elle puisse éclairer en apportant sa réflexion empirique ancrée dans le terrain et la réalité des enjeux d'aujourd'hui pour un futur souhaitable. Dès lors, ni le neuroscientifique, ni l'éthicien, ni le juriste, ni même le philosophe ne peuvent faire cavalier seul. C'est donc avant tout un défi pour tous les chercheurs, que de s'ouvrir au dialogue interdisciplinaire, que d'ouvrir leurs recherches, leurs rêves de chercheurs et leurs envies d'un futur plus extraordinaire à la société. Il en va de notre responsabilité de chercheurs vis-à-vis de la société.

Il faut dévoiler les mécanismes qui sont derrière ces technologies et ne pas se laisser trop impressionner. Pour discerner quelque chose de bon ou de mauvais dans les usages, il faut connaître et comprendre ces technologies et leurs mécanismes.

La neuroéthique [22, 23], qui se place à l'intersection entre les sciences humaines et les neurosciences et qui invite à discerner entre ce qui – dans les transformations engendrées par les neurotechnologies – est souhaitable et ce qui l'est moins, incarne l'approche pluridisciplinaire indispensable à toute innovation responsable. De plus, nous remarquons une mise en commun progressive des réflexions neuroéthiques et de l'éthique de l'IA [24], ce qui donne un véritable élan pour comprendre comment les connaissances et les recherches en neurosciences, en informatique et en ingénierie peuvent affecter le futur de la société. Des protections appropriées des données cérébrales, des espaces privés et de l'identité individuelle doivent être intégrées dans notre compréhension des droits de l'Homme. Il ne faut donc pas être ignorant du contexte international, qui est en ébullition concernant les recherches émergentes sur les neurosciences et les applications des systèmes d'IA. Depuis 2019, l'OCDE [17, 18, 19] et l'Unesco [20, 21] agissent et réfléchissent activement pour mettre en place des systèmes de protection, des lieux de réflexions, des recommandations et des normes juridiques pour border, border les systèmes d'IA, le numérique et les neurotechnologies<sup>1</sup>. L'IA Act, le DATA Act et le RGPD que l'Union européenne a mis en place sont des textes en vigueur qui se penchent très sérieusement sur la problématique des données.

Comprendre d'où viennent l'intelligence, la conscience<sup>2,3</sup>, le génie humain, la pensée, les rêves, les intuitions et les intentions va au-delà des recherches scientifiques de « paillasses », c'est une ouverture à un monde spirituel, de l'esprit, qui nous fait passer au-delà de la conscience, de l'intelligence, de la pensée physique et mesurable, quantifiable et visible *via* les neurotechnologies et l'IA. À côté de réflexions sur les sciences dites « exactes », il est réellement souhaitable d'ajouter d'autres réflexions du côté des sciences morales et politiques, voire de la théologie. De fait, le Vatican, par exemple, mène une réflexion sur l'IA. Le message de la journée pour la paix du 1<sup>er</sup> janvier 2024 était sur

---

<sup>1</sup> <https://www.unesco.org/en/ethics-neurotech?hub=8329>

<sup>2</sup> <https://theconversation.com/comment-les-neurosciences-expliquent-elles-la-conscience-232312>

<sup>3</sup> <https://www.radiofrance.fr/franceculture/podcasts/la-science-cqfd/theorie-de-l-esprit-l-examen-de-conscience-9088912>

l'intelligence artificielle<sup>4</sup> en soulignant que les applications technologiques et les innovations les plus avancées ne doivent pas être utilisées pour faciliter la résolution violente des conflits, mais « pour paver les voies de la paix : les progrès dans le développement de formes d'intelligence artificielle servent [...] la cause de la fraternité humaine et de la paix » [6]. De plus, le Vatican mène une réflexion en coopération avec l'Unesco<sup>5</sup> par le biais du Centre Catholique International de Coopération (CCIC), qui a pour objectif, en co-création avec la Commission nationale française de l'Unesco, d'élaborer un aide-mémoire qui devrait aider les jeunes, les familles, les éducateurs de toute nature et la société à appréhender les enjeux éthiques inhérents aux technologies de l'IA, en parallèle à un cycle de tables rondes. J'ai d'ailleurs eu l'honneur d'intervenir dans ce cadre, en tant que neuroscientifique et éthicienne, le 12 novembre 2024 à la table ronde du CCIC : « Éthique et IA, de quoi parle-t-on ? »<sup>6</sup>.

Seul l'humain humanise. L'être humain devient une personne, s'humanise, par les relations qu'il engage avec d'autres personnes. Seule une personne qui n'est pas moi, mais qui est face à moi, personnalise la personne. Les relations interpersonnelles restent et resteront essentielles à l'humanisation de l'être humain, qui n'est pas calculable. Ces recherches nous montrent que ce matérialisme et ce réductionnisme dans lesquels les neurosciences et l'informatique nous amènent ne doivent pas étouffer une autre manière de voir, d'envisager qu'il y a au cœur de ce dualisme corps-âme, un autre lieu, un lieu plus intime, plus secret, peut-être pour certains un lieu de rencontre, avec un autre, incalculable, non palpable, non matériel mais qui nous rencontre dans cette humanité que nous incarnons. Finalement, c'est peut-être bien un trio, corps-âme-esprit, qu'il nous faut envisager dans cette convergence et cette quête de ce que sont l'intelligence, la pensée et la conscience, un trio qui fonctionne ensemble pour révéler notre humanité, ce qui fait que nous sommes des êtres humains et non des machines. Alors à votre avis, sommes-nous notre cerveau ? Allons-nous pouvoir vraiment télécharger notre conscience, notre esprit dans une machine ? Y- aura-il des IA conscientes ? Allons-nous pouvoir réellement lire dans les pensées, la conscience et les intentions et intuitions des personnes ?... Que de fantasmes et de rêves... Et surtout, est-ce souhaitable ?

## BIBLIOGRAPHIE

- [1] AMUNTS K. *et al.* (2024), "The coming decade of digital brain research: A vision for neuroscience at the intersection of technology and computing", *Imaging Neuroscience*, 2024, 2, pp. 1-35.
- [2] RAYHAN S. (2023), "The intersection of AI and neuroscience: Exploring cognitive enhancements and ethical dilemmas", doi:10.13140/RG.2.2.36679.06568
- [3] WEXLER A. & SPECKER SULLIVAN L. (2023), "Translational neuroethics: A vision for a more integrated, inclusive, and impactful field", *American Journal of Bioethics Neuroscience*, 14, pp. 388-399.
- [4] SURIANARAYANAN C., LAWRENCE J. J., CHELLIAH P. R., PRAKASH E. & HEWAGE C. (2023), "Convergence of artificial intelligence and neuroscience towards the diagnosis of neurological disorders - A scoping review", *Sensors*, 23(6), pp. 1-29.
- [5] TURING A. (1950), "Computing machinery and intelligence 1950", *Mind*, 59(236), pp. 433-460.

<sup>4</sup> <https://www.vatican.va/content/francesco/fr/messages/peace/documents/20231208-messaggio-57-giornatamondiale-pace2024.html>

<sup>5</sup> <https://ccic-unesco.org/intervention-du-ccic-au-debat-de-politique-generale/>

<sup>6</sup> <https://ccic-unesco.org/replay-ethique-et-ia-de-quoi-parle-t-on/>

- [6] FRANÇOIS P. & VATICAN (2024), « Message de Sa Sainteté pour la 57<sup>e</sup> journée mondiale de la paix [en ligne sur le site du Vatican] », pp. 1-10.
- [7] GILBERT F. & RUSSO I. (2024), “Mind-reading in AI and neurotechnology: evaluating claims, hype, and ethical implications for neurorights”, *AI Ethics*, 4, pp. 855-872.
- [8] BROWN C. M. L. (2024), “Neurorights, mental privacy, and mind reading”, *Neuroethics*, 17, pp. 1-19.
- [9] FINKEL E. (2023), “If AI becomes conscious, how will we know?”, doi: 10.1126/science.adk4451
- [10] CHALMERS D., BUTLIN P. & BIRCH J. (2024), “Taking AI welfare seriously”.
- [11] FINKEL E. (2023), “Researchers propose test for AI sentience”, *Science*, (80)381, pp. 822-823.
- [12] YURCHENKO S. B. (2023), “A systematic approach to brain dynamics: cognitive evolution theory of consciousness”, *Cogn. Neurodyn.*, 17, pp. 575-603.
- [13] THIELE J. A. & FASKOWITZ J. (2024), “Choosing explanation over performance: Insights from machine learning-based prediction of human intelligence from brain connectivity”, *PNAS Nexus*, 3, pp. 1-15.
- [14] LUPPI A. I. *et al.* (2024), “Unravelling consciousness and brain function through the lens of time, space, and information”, *Trends Neurosci.*, 47, pp. 551-568.
- [15] ID W. L. N. (2024), “On human nanoscale synaptome: Morphology modeling and storage estimation”, *PLoS One*, 19(9), pp. 1-18 doi:10.1371/journal.pone.0310156
- [16] NOWINSKI W. L. (2024), “Storage estimation in morphology modeling of the human whole brain at the nanoscale”, *J. Comput. Sci.*, 81, 102346.
- [17] OCDE (2019), « Recommandation du Conseil sur l’intelligence artificielle ».
- [18] OECD (2023), “Emerging privacy-enhancing technologies: Current regulatory and policy approaches”, OECD Digital Economy Papers, No. 351.
- [19] OECD (2019), “Recommendation on responsible innovation in neurotechnology”, *Oecd/Legal/0457* 10.
- [20] UNESCO (2021), « Recommandation sur l’éthique de l’intelligence artificielle », *Unesco* 43.
- [21] UNESCO (2022), “Ethical issues of neurotechnologies”, comité bioéthique, doi:10.1016/S0196-0644(05)81763-9.
- [22] TABOUY L. (2021), « Ce que les neurotechnologies soulèvent comme enjeux éthiques et légaux pour la recherche, les neuroscientifiques, les entreprises et la société », *Annales des Mines - Réalités Industrielles*, août 2021, pp. 65-73.
- [23] GOERING S. *et al.* (2021), “Recommendations for responsible development and application of neurotechnologies”, *Neuroethics*, 14, pp. 365-386.
- [24] SALLES A. & FARISCO M. (2024), “Neuroethics and AI ethics: a proposal for collaboration”, *BMC Neurosci.*, 25, <https://doi.org/10.1186/s12868-024-00888-7>

# Towards responsible and ethical AI

## INTRODUCTION

### 05 Towards human intelligence *Nicolas CHAGNY*

This issue of *Enjeux numériques* focuses on the importance of individuals and the planet in the context of artificial intelligence (AI). The contributors, chosen for their diversity of experience and points of view, encourage in-depth reflection on our human interactions, both personal and professional.

The issue points out that, although AI technology has been around since the 1950s, what has really evolved recently is user experience (UX). This evolution has enabled a wider and more democratized use of AI, similar to previous technological revolutions such as Web 2.0, the iPhone, voice assistants and 5G.

The issue also explores Responsible AI and its impact on democratic access to technology, environmental, social and economic implications, as well as privacy and public life issues.

### 08 Challenges and prospects for ethical and sustainable AI *Guillaume BOURGEOIS,* *Luciana GONDIM DE ALMEIDA GUIMARÃES* *& Vincent COURBOULAY*

Artificial intelligence (AI) has emerged as a driving force across industries, but it also raises significant ethical, social, and environmental challenges. This article examines the concept of responsible AI, focusing on key principles such as ethics, transparency, security, and social inclusion. It further assesses AI's environmental impact, with particular attention to data centers' energy consumption and the risk of exacerbating social inequalities. By proposing technical and policy-based solutions, including energy-efficient models and international regulations, this paper aims to promote sustainable AI development, maximizing benefits while minimizing risks.

## IA FOR ALL, ALL FOR IA

### 15 Democratising our relationship with technology, one AI café at a time *Jean CATTAN*

Artificial intelligence is creeping into every corner of our lives, often without us being fully aware of it or equipped to properly define our uses. Between fascination and concerns, the issues related to its use concern us all, whether it is understanding its applications in our professions, in the education of our children or in public services.

It is to respond to these challenges that the National Digital Council initiated Café IA: to support and encourage the creation of spaces for exchange, learning and



dialogue open to all, where everyone can express themselves, ask questions, and imagine their own conditions for using or not using AI. Like the café, an informal place conducive to sharing knowledge, these meetings aim to democratize the understanding of technological issues and to strengthen a local technical democracy.

**19 Towards an “ILearning” society**

***Frédéric BARDEAU***

Artificial intelligence is at the heart of the transformations of our societies. While it promises immense opportunities, it also raises ethical, economic, ecological and social challenges. How can we prevent this technological revolution from leaving part of the population behind? Training the general public in AI is not only a necessity, but an emergency.

**22 The augmented trainer: between artificial intelligence and emotional intelligence**

***Brice GAILLARD***

Artificial intelligence (AI) is shaking up the world of education, profoundly transforming traditional teaching practices. With 78% of university students already using AI as part of their curriculum, trainers face a major challenge: integrating these new tools without losing their human essence. Far from replacing them, AI presents itself as a powerful tool for enhancing them, offering unprecedented possibilities for personalizing and making learning more accessible.

Between technological adaptation and preservation of emotional intelligence, tomorrow’s instructor becomes a learning architect, capable of navigating between digital data and human support. This revolution raises major questions about ethics, accessibility and the central role of the human being in the transmission of knowledge.

**26 For an education in the plurality of alterities and digital attachments in the age of generative artificial intelligence**

***Jean-François LUCAS***

This article supports the need for artificial intelligence education that allows children to understand and use these sociotechnical devices in an informed manner, that is, to make responsible choices. This must take place at school, and outside of school. However, at school, and despite good intentions, the development of the knowledge and skills necessary to understand and seize the opportunities of artificial intelligence is hampered by a lack of means and resources. The development of “outside school” projects and initiatives is interesting in this respect.

The article concludes by emphasizing the importance of developing the emotional skills of children, and of citizens more broadly, so that they can navigate and seize the opportunities of the otherness and digital links that they will be confronted with, and that they will develop, particularly in light of the development of generative artificial intelligence systems.

**33 Artistic creation put to the test by Artificial Intelligence**

***Alain ASSOULINE***

In our present and future relationship with AI, the question of artistic creation probably concentrates most of the problems we should be asking ourselves.

Dance, the visual arts, literature, music and all the other arts have existed since the human race first appeared. They are the different expressions of human emotions and relationships. They are the most profoundly human part of a culture.

What will AI do with this? What should it be allowed to do? Do we still have a choice?

AI at the service of humanity, of women and men, can enhance and perhaps even magnify the arts. An AI at the service of algorithmic logics of efficiency, profitability or power, would reduce art to a communication tool that would undoubtedly be very efficient, but would annihilate diversity, originality, in short, human genius.

There is an urgent need to put in place tools to regulate the use of AI in all artistic fields.

## ENVIRONMENTAL ISSUES

### 37 **The impact of AI on the environment**

*Frédéric GARCIA & Sophie SCHBATH*

With the rapid development of artificial intelligence (AI), particularly generative AI, the link between AI and the environment is raising many questions.

On the one hand, AI is emerging as a key technology in the service of the ecological transition, boosting our predictive and decision-making capacities tenfold to ensure more sustainable use of natural resources and a more detailed understanding of socio-technical and environmental systems.

On the other hand, AI requires a great deal of IT equipment that consumes a lot of energy, water and rare metals, which has a devastating impact on the environment itself, in terms of resource depletion, loss of biodiversity and greenhouse gas emissions.

Given the current explosion in the use of AI by the general public, we are entitled to wonder whether the unanimous calls for the development of “frugal” AI will be enough to ensure that the environmental benefits of AI outweigh its environmental damage.

### 43 **Sustainable AI does not exist**

*Frédéric MARCHAND*

Can sustainable AI exist? This article addresses the environmental implications of artificial intelligence (AI), often ignored in the face of its benefits. AI, while promising advances, generates negative externalities, particularly environmental ones. With the democratization of generative AI, its energy consumption is increasing significantly, illustrated by the exponential increase in demand for computer computing. A measured useful AI and not an excessively futile AI is the path recommended by the author.

### 50 **Developing AI for the common good**

*Laure de La RAUDIÈRE*

The emergence and rapid diffusion of generative artificial intelligence to the general public since the end of 2022 are as much a vector of opportunities as they are of economic and societal challenges. As the architect and guardian of exchange networks as common goods, Arcep warns of several potential impacts of generative

AI. By becoming the new gateways to our access to the internet, the major players in generative AI could call into question the principle of the openness of the internet. In addition, the development of generative AI raises questions about its environmental sustainability. On these issues, several levers must be mobilized so that the development of AI is for the common good.

**58 Integrating AI into an eco-designed service:  
oxymoron or reality?**

***Mathieu DELEMME, Christophe CLOUZEAU,  
Vincent COURBOULAY, Jean-Luc MARINI,  
Emmanuel NURIT, Romuald RIBAUT  
& Claire VERDIER***

Integrating artificial intelligence (AI) into an eco-designed service raises some major paradoxes. On the one hand, eco-design seeks to reduce the environmental impact of digital services by advocating sobriety and optimisation. On the other hand, AI, and in particular generative AI, requires huge quantities of energy, water and computing resources that use critical metals and rare earths.

Faced with AI-related electricity consumption that could soon rival that of certain countries, the very idea of associating AI and eco-design seems contradictory. However, some avenues are emerging for designing more responsible AI. But the main challenge remains to determine whether and under what conditions their use can really be integrated into a sustainable approach.

## **SOCIAL ISSUES**

**65 Artificial intelligence and human rights:  
the inadequacies of the European framework**  
Thomas DUMORTIER

The use of artificial intelligence raises as many concerns as it does hopes. After the era of “AI Ethics”, in the 2010s, the European Union and the Council of Europe wished to establish a firmer framework which recently materialized, respectively in a regulation and a framework convention. Although these texts show a desire to protect human rights, they arouse a number of concerns in this area.

These concerns relate, on the one hand, to the insufficient delimitation of unacceptable AI systems because of their damage to human rights and, on the other hand, to the uncertainty about the scope of the guarantees established at all stages of life of an AI system to preserve these rights, starting with the human rights impact assessment.

**70 Generative AI and misinformation**  
***Nicolas CURIEN***

Misinformation means information which is either unwillingly wrong or deliberately falsified, then called disinformation, “infox” or fake news. Although these information diseases have existed for long, the growth of digital technologies drastically increased their impact. Today they spread as a virus all over the Internet, because fake content is more attractive than genuine facts and thus benefits from the business model of large platforms, based on the capture of users’ attention.

In a report published in December 2024, the French Academy of technologies signals the potential danger of an amplification of misinformation due to generative AI. This report underlines the progress of a curative AI used to counter misinformation, as well as the elaboration of a pioneering European regulatory framework. It also makes several targeted proposals in order to reinforce society's resilience in the face of fake news and to promote the sincerity of information.

**77 AI-xiety: between myth and reality,  
is artificial superintelligence already here?  
*Aurélie GIARD-JACQUET & Stéphanie POTTECHER***

In March 2023, a call for caution signed by 30,000 people reignited the debate: should we put the brakes on the development of artificial intelligence? This article delves into the contemporary anxieties surrounding AI, between dystopian fantasies and real challenges. Drawing on literary and cinematic references, we ask: is AI still a tool at our service, or an existential threat? By addressing the issues of alignment, the concentration of power, and the costs induced by AI, we explore what, tomorrow, could turn the destiny of humanity upside down.

**85 Mental health at work and artificial intelligence:  
between psychological support and the risk of dependency  
*Christian MAKAYA & George KASSAR***

This article presents an exploratory study on the use of conversational artificial intelligence (AI) tools by employees with mental health disorders. Preliminary findings reveal a positive perception of AI tools as a support for productivity and well-being but highlight the risk of heavy dependence on AI, which could deteriorate mental health. Future research directions include conducting a quantitative study with a large and diverse panel and co-designing workplace psychological support systems leveraging AI.

**91 European legislation and social dialogue  
on artificial intelligence  
*Franck GAMBELLI***

The new European regulation on artificial intelligence will apply on top of existing or future legislation that also applies to generative AI systems. This article takes a closer look at European social legislation and legislations concerning the free circulation of goods. At this stage, the discussions between the European social partners provide a clearer picture of the main issues for employers and workers.

**100 AI and the Commons: combining  
technological power and terrestrial habitability  
*Emmanuelle ROUX***

In the face of climate, social, and democratic crises, artificial intelligence can either amplify inequalities or, on the contrary, become a catalyst for collective transformation. To achieve this, it is necessary to go beyond its role as a tool for the commons and view it as a common in itself. This implies creating models governed collectively, as envisioned by platforms like Hugging Face or open-source initiatives like BLOOM. By integrating AI into a logic of digital commons, we could not only democratize its access but also reduce its ecological impacts, while strengthening citizen engagement and transparency.

**106 Towards “gender by design” artificial intelligence?**

***Peggy VICOMTE & Camille SALINESI***

Artificial intelligence (AI) is reshaping our personal and professional lives, establishing itself as an omnipresent technology. Pervasive in fields such as healthcare, urban planning, agriculture, education, and culture, it offers new opportunities but also raises significant societal and ethical challenges.

While AI holds promise, it also amplifies pre-existing biases found in its training data, reinforcing discrimination and stereotypes, as seen in documented cases in recruitment and healthcare. Furthermore, the lack of diversity in digital professions exacerbates these issues, highlighting the urgent need for more inclusive AI.

Interdisciplinary initiatives combining social sciences and humanities are paving the way for AI to become a tool for equity and collective progress. A “gender by design” approach to AI development may provide the framework to ensure justice and inclusivity in its creation and use.

## CHALLENGES FOR PROSPERITY

**111 AI and the future of civilisation:  
transhumanist dystopia or creative metamorphosis?**

***Boris SIRBEY & Hervé BÉRAUD***

Humanity is experiencing exponential acceleration marked by technological, social, and ecological transformations, pushing the limits of a civilizational paradigm rooted in patriarchal domination. The historical turning point of the “Point 0”, highlighted by Boris Sirbey, leads to a global collapse, a transhumanist dystopia, or a creative metamorphosis, depending on the combined evolution of our collective consciousness and technological mastery.

AI, as a reflection of our knowledge and a mirror of our ethical choices, holds immense potential: catalyzing enhanced awareness and systemic solutions. However, it remains confined by private interests, limiting its ability to become a shared asset of wisdom.

For a positive transition, we must reconcile technology with humanity. “Artists of Amicable Intelligence”, inspired by Alain Damasio, could create ethical and collaborative interfaces and embody this creative metamorphosis, transforming AI into a tool for collective emancipation.

The future will depend on our ability to harmonize innovation and wisdom, making AI a lever to transcend our limits and build an inclusive and balanced civilization.

**117 AI and the transformation of professions:  
creation or destruction?**

***Guy MAMOU-MANI & Axel MAMOU-MANI***

Artificial intelligence, as a catalyst for transformation across all sectors, is redefining professional skills and roles. While some jobs are disappearing or undergoing profound changes, new opportunities are emerging. However, AI also raises ethical, social, and economic challenges, including the digital divide, growing inequalities, and loss of control. By rehumanizing work and fostering an inclusive and equitable transition, this technological revolution can become a powerful lever for human and social innovation. This article explores these disruptions and outlines the conditions for a responsible AI that serves everyone.

**122 The AI-compatible organisation or the art of knowing how to recruit technology**  
***Romain RABIER***

Understand, decide and act. This is the winning combination for organisations that are able to integrate artificial intelligence and make it a lever for overall performance! Infobesity and the age-old biases of financial capitalism still lead too many decision-makers to take an imprecise approach to the subject. That's why understanding the major issues that AI poses for the world of work will shed light on decisions and actions that are as much about economic gains as about the psychological security of employees, as much about maintaining social links as about the quest for societal impact that we have every right to expect from organisations.

**129 Building an ambitious and peaceful framework for using AI in education and training**  
***Oriane LEDROIT***

At a time when professions are changing, some are disappearing, others are emerging, enabling everyone to train and renew their skills is one of the sine qua non conditions for: building our technological and ecological resilience; nurturing our power of creation and innovation; reindustrialising the region; giving each of us a place in society and with that place (re)empowering us to act; and finally, in a way, holding our own.

These major challenges demand that our teaching environments and learning methods – from the earliest age and throughout life – evolve, adjust and take advantage – in an ambitious and pragmatic way – of the opportunities offered by technology and in particular the power of artificial intelligence. This means thinking calmly about our relationship with AI in education and training. What are the issues? What are the challenges? What benefits? What framework can we use to ensure that AI serves a learning society?

**135 Artificial intelligence and territories**  
***Fabien BAZIN***

Artificial intelligence has increased its penetration in public policies, unevenly, but its potential is undeniable, both for progress on reliable shared databases and as a response to access to social rights and the resilience of territories. Nevertheless, the risks of AI for fundamental rights and freedoms cannot be underestimated.

**140 Appropriation profiles of generative artificial intelligence in education**  
***Loubna MOURTAJJI & Nathalie CHISS***

Artificial intelligence is revolutionizing practices in all sectors, including education, where tools such as ChatGPT are raising many expectations and questions. This article empirically explores teachers' perceptions of this tool, identifying the factors influencing its adoption as well as the reluctance expressed. Qualitative analysis of semi-directive interviews reveals that teachers' perceptions of ChatGPT's usefulness are highly contrasted.

These perceptions are largely influenced by contextual variables, such as the link with technology, the subject taught, the status or age of the teachers. The results obtained enable us to identify four teacher profiles and make recommendations for

optimizing the teachers' use of ChatGPT in an effective and responsible way, while considering the ethical, practical and pedagogical implications of this integration.

## ISSUES FOR PRIVATE AND PUBLIC LIFE

### 150 AI and freedoms: a challenge for regulation

*Marie-Laure DENIS*

To provide a framework for artificial intelligence systems, the European Union has stepped up its regulation by introducing the Artificial Intelligence Regulation (AIR), in addition to other regulations such as the General Data Protection Regulation (GDPR). These texts aim to guarantee the transparency and security of AI systems, the former by limiting systemic risks, and the latter by protecting individual rights and privacy.

This legal framework is also intended to prevent certain infringements of fundamental freedoms, such as the expansion of surveillance or the excessive use of automated decisions. However, AI regulation requires close cooperation between regulators, developers and users to ensure that technological innovation respects fundamental rights and social values.

### 156 Law in the age of artificial intelligence

*Didier GUÉVEL*

AI is of interest to lawyers as an informative complement. Legal practitioners and theorists already have specialized AI. Two problems remain: that of controlling the originality of the students' work and especially that of the choice, upstream, of the data used.

### 158 The ethical challenges of the convergence of AI, neuroscience, computer science and engineering

*Laure TABOUY*

The ethical, legal, philosophical and theological questions, challenges and dilemmas raised by the convergence of neuroscience, computer science and engineering are colossal and dizzying.

Understanding the ethical challenges of AI, digital technology and neuroscience, in order to grasp them and put them into practice in the field, requires an understanding of the history, the whys and the wherefores. Even if these artificial intelligence systems offer new possibilities, the fact remains that they are computer systems, tools and machines, and should not be taken for what they are not.

This article sets out to share a global vision of the issues and challenges involved in this convergence, by highlighting the interconnectedness of the issues. It provides a basis for reflection, with a view to examining more specifically the ethical issues involved in the specific uses of this convergence in a variety of contexts.

*Issue editor*  
*Nicolas CHAGNY*

## Ont contribué à ce numéro

**Alain ASSOULINE**, aujourd'hui peintre, a été le fondateur et le président du réseau des écoles WebForce3 délivrant des formations aux métiers du numérique pour un accès immédiat à un emploi. Implanté dans 50 territoires, 10 000 personnes y ont été formées entre 2014 et 2022.

Son credo : réinventer la formation professionnelle initiale et continue à l'ère digitale en accompagnant les apprenants sur un marché du travail en constante mutation pour leur permettre d'être toujours en situation de s'adapter aux changements des métiers ou d'occuper ceux qui émergent.

Chevalier de la Légion d'Honneur, il a été administrateur de Numeum, le syndicat des entreprises du numérique, après avoir été président de Cinov-Numérique, le syndicat des TPE/PME du numérique, et président de la commission numérique de la CPME.

Alain Assouline est un des pionniers du numérique en France.

En 1994, il a fondé Les Argonautes, une des premières agences digitales de France qui s'est distinguée en créant, entre autres, le site Apec.fr, la plateforme Peugeot de déploiement de ses sites internet dans 90 pays et, pour l'Éducation nationale, le Réseau Professionnel des Enseignants ViaÉduc.

Dès 1990, il était expert Apple pour aider les journaux, agences publicitaires et grands studios graphiques à passer au numérique par la formation des salariés et le conseil aux dirigeants.

Il a débuté sa carrière professionnelle en 1976 comme directeur artistique chez RSCG.

→ *La création artistique à l'épreuve de l'intelligence artificielle*

**Frédéric BARDEAU**, passionné de cyberculture, d'*empowerment* et de technologie, travaille depuis 25 ans à faire converger numérique et impact. Il a co-fondé en 2008 l'Agence Limite, entièrement dédiée à la communication responsable des ONG et des fondations, puis en 2013, Simplon.co, entreprise sociale qui a formé aux compétences et aux métiers techniques du numérique plus de 35 000 demandeurs d'emploi (dont 40 % de femmes, 12 % de personnes en situation de handicap, 6 % de personnes réfugiées et 45 % d'apprenants de niveau bac ou inférieur), et accompagne les mobilités professionnelles des salariés au travers d'un réseau de 125 écoles dans 25 pays (Europe, Moyen-Orient, Afrique, Colombie, Cambodge). Depuis 2018, Simplon a formé plus de 3 000 personnes sur les métiers et les compétences liés à l'IA. Au travers de la Simplon Foundation, des programmes de sensibilisation à l'IA et des *hackathons* centrés sur les IA génératives permettent de toucher des publics éloignés de l'emploi, des associations et de faire connaître des cas d'usages à impact positif.

Frédéric Bardeau est également expert certifié par Bpifrance dans le cadre du dispositif Booster Data IA prise en charge au titre de France 2030. Fellow Ashoka et EPIC Foundation depuis 2015, entrepreneur social de l'année 2017 du BCG, Personnalité IT de l'année 2022 pour *Le Monde Informatique*, Prix de l'Audace HEC en 2023, il est également l'auteur, avec Nicolas Danet, de *Anonymous, peuvent-ils changer le monde ?* (2011) et de *Lire, écrire, compter, coder* (2014) publiés chez FYP Éditions.

→ *Vers une société « IAApprenante »*

**Fabien BAZIN** est président du conseil départemental de la Nièvre depuis le 1<sup>er</sup> juillet 2021. Ce professionnel du droit public a été élu pour la première fois au conseil général, puis départemental, en 2004, sur le canton de Lormes puis de Corbigny. De 2011 à 2014, il était vice-président en charge des TIC et des territoires, puis à partir de 2014 en charge de l'autonomie, de la solidarité, des nouvelles sociabilités et des citoyennetés ainsi que du développement numérique.



Il a été nommé vice-président de la commission « Politiques territoriales et ruralité » au sein de l'association des Départements de France.

Fabien Bazin a été maire de Lormes de 2001 à juin 2021 et, à ce titre, vice-président du Pays Nivernais Morvan en charge des innovations et des Villages du futur ainsi que de la Communauté de communes Morvan, sommets et grands lacs, en charge du développement économique.

Entre 1991 et 2015, il a été attaché parlementaire du député de la Nièvre Bernard Bardin (1991-1993), du sénateur de la Nièvre René-Pierre Signé (1995-1997), du député de la Nièvre Christian Paul (1997-2013) et de la sénatrice de la Nièvre Anne-Emerly-Dumas (2013-2015), après avoir été collaborateur au cabinet du président du conseil général de la Nièvre (1993-1994) et chargé de mission au Comité départemental du tourisme de la Nièvre (1994-1995).

Concepteur du Bouclier rural en 2009, auteur de la réflexion sur le Village du futur en 2014, Fabien Bazin est convaincu que l'avenir appartient aux territoires ruraux volontaristes. À Lormes, petite ville de moins de 1 500 habitants, il a permis de tester des idées innovantes pour façonner un modèle de développement rural original, souvent médiatisé.

→ ***Intelligence artificielle et territoires***

**Hervé BÉRAUD**, bâtisseur d'utopies concrètes et explorateur de futurs désirables, ouvre des chemins là où d'autres voient des impasses. Chef d'entreprise et pionnier de l'innovation sociétale, il donne vie au « secteur quaternaire » : un modèle économique inédit où le bien commun de connaissances partagées devient le moteur d'un monde durable.

Cofondateur de CollectivZ, qui conçoit l'entreprise comme un lieu de civilisation, et de ze.game, qui prépare au futur du travail, il accompagne les organisations et les personnes dans leur transition vers des modèles collaboratifs et résilients. Émerveillé par le vivant sous toutes ses formes, il mobilise la puissance du collectif, les *soft skills* et l'intelligence artificielle pour convertir chaque défi stratégique en opportunité de régénération et d'innovation.

Diplômé de l'ESSEC et de KEDGE BS, *coach* exécutif certifié, il s'appuie sur une riche expérience en transformation culturelle et *design* pédagogique. Conjuguant poésie, pragmatisme et alliance des intelligences, il propose des approches systémiques innovantes pour :

- préparer les entreprises à naviguer dans l'incertain ;
- favoriser la cohésion d'équipe et l'apprenance ;
- cultiver les *soft skills* et la créativité collective ;
- ancrer l'action du management dans une dynamique de sens et de durabilité.

Sa mission ? Relier les élans de vie aux nécessités économiques pour participer à l'émergence d'une civilisation où les écosystèmes, organiques et régénérateurs, deviennent les alliés d'un bien commun au service d'un futur profondément humain et durable.

→ ***IA et futur de la civilisation : dystopie transhumaniste ou métamorphose créatrice ?***

**Guillaume BOURGEOIS** est titulaire d'un doctorat en informatique de l'Université de La Rochelle en France. Il est également titulaire d'un master en informatique, spécialisé dans les systèmes informatiques pour l'ingénierie, obtenu à l'université de Pau et de la région de l'Adour en France. Il poursuit actuellement un stage postdoctoral en informatique, supervisé conjointement par l'Université de La Rochelle en France et l'École de technologie supérieure (ETS) au Canada.

→ ***Enjeux et perspectives pour une IA éthique et durable***

**Jean CATTAN** est secrétaire général du Conseil national du Numérique, docteur en droit public et chargé d'enseignements en droit et régulation du numérique à Sciences Po Paris et à l'Université Panthéon-Assas. Il a notamment été conseiller du président de l'Arcep de 2018 à 2020. Il est l'auteur d'une cinquantaine d'articles sur les enjeux du

numérique et co-auteur avec Serge Abiteboul de *Nous sommes les réseaux sociaux* (Odile Jacob, septembre 2022).

→ *Démocratiser notre rapport à la technologie, un Café IA à la fois*

**Nicolas CHAGNY** est depuis plus de 25 ans au cœur du numérique, comme chef d'entreprise, engagé pour un numérique ouvert et responsable, pour l'inclusion et l'égalité des chances.

Il pilote des projets numériques, d'abord comme développeur, puis comme chef de projet, directeur de projet et directeur général. Il a eu l'occasion de diriger d'importants projets digitaux pour, notamment, l'Apec, l'Aftral, l'Afnic, l'UIMM, Deloitte, Orange, le Crédit agricole, Peugeot, Radio France, le Réseau Canopé et Teréga. Il a dirigé l'agence *web* Les Argonautes et l'agence de contenu Médiasactu, puis a rejoint en tant que directeur général adjoint digital et *data* le groupe de communication Makheia.

Associé et membre du comité stratégique du réseau d'écoles du numérique WebForce3 depuis 2014, il en devient le directeur général de septembre 2019 à novembre 2023, après la cession du réseau au groupe Collège de Paris. Début 2024, il crée NS Pulse, cabinet de conseil en stratégie, développement et relations publiques pour les projets et entreprises à impact, qui ont pour mission d'améliorer notre société. Il participe également aux activités de CollectivZ / Ze.Game.

Nicolas Chagny est engagé dans la défense des utilisateurs d'internet, pour un internet libre et ouvert *via* l'ONG Internet Society France, qu'il préside depuis 2015, et il est membre de la Commission nationale consultative des droits de l'homme (CNCDH).

→ *Pour une intelligence humaine*

**Nathalie CHISS**, docteur ès Sciences de Gestion, Maîtresse de Conférences, membre du LEFMI (Laboratoire d'Économie, Finance, Management et d'Innovation - UR4286) de l'Université Picardie Jules Verne (UPJV) et membre du groupe MAIA (Maîtrise des Applications en Intelligence artificielle). Elle s'intéresse notamment au processus de prise de décision et aux nouveaux comportements du consommateur, et à l'impact de l'IA dans différents secteurs d'activité (éducation, banque, entrepreneuriat).

→ *Profils d'appropriation de l'intelligence artificielle générative dans l'éducation*

**Christophe CLOUZEAU** est un expert GreenUX, responsable du pôle Écoconception numérique au sein de la société Temesis. Il s'intéresse depuis 2010 à l'impact environnemental du numérique. Il contribue à différents groupes de travail au sein de l'INR (Institut du Numérique Responsable), du collectif GreenIT, des Designers Éthiques, de Boavizta, de Flupa et de l'Afnor et du W3C. Il est co-rédacteur du RGENS, le Référentiel général de l'écoconception des services numériques.

→ *Intégrer l'IA dans un service éco-conçu : oxymore ou réalité ?*

**Vincent COURBOULAY** est ingénieur et maître de conférences en informatique à l'Université de La Rochelle. Depuis une dizaine d'années, il s'est spécialisé dans les pratiques numériques responsables. En 2018, il a cofondé l'Institut du numérique responsable, dont il assure la direction scientifique. Il se concentre actuellement sur le concept d'intelligence artificielle responsable.

→ *Enjeux et perspectives pour une IA éthique et durable*

→ *Intégrer l'IA dans un service éco-conçu : oxymore ou réalité ?*

**Nicolas CURIEN** est diplômé de l'École polytechnique (1973) et de Télécom Paris (1975), docteur en mathématiques appliquées de l'Université Paris 6 (1976), ingénieur des Mines. Il est depuis 2000 membre fondateur de l'Académie des technologies. Professeur émérite en économie, il a été titulaire de la chaire « Économie et politique des télécommunica-

tions » du Conservatoire national des arts et métiers (CNAM) de 1992 à 2005, où il a aussi dirigé le laboratoire d'économétrie.

En alternance avec son cursus d'enseignant-chercheur, il a été responsable d'études économiques au ministère des PTT (1976-1984) puis au ministère de la Défense (1986-1988), avant d'être nommé directeur-adjoint de l'ENSAE, École nationale de la statistique et de l'administration économique (1989-1991). Il a deux fois été désigné par le président du Sénat comme membre du collège d'une autorité de régulation : d'abord à l'Arcep, régulateur des communications électroniques et des postes (2005-2011) puis à l'Arcom, régulateur des services audiovisuels et de la communication en ligne (2015-2021). Il est l'auteur de nombreux articles scientifiques et de plusieurs livres dans les domaines de l'économie industrielle, l'économie des réseaux et l'économie numérique.

→ **IA générative et mésinformation**

**Laure de LA RAUDIÈRE** est la présidente de l'Autorité de régulation des communications électroniques, des postes et de la distribution de la presse (Arcep) depuis le 28 janvier 2021.

Diplômée de l'École normale supérieure, ingénieur des télécommunications, elle a passé toute sa carrière dans le secteur des télécommunications et du numérique. Elle a été aussi députée d'Eure-et-Loir de 2007 à 2021, particulièrement investie sur les enjeux économiques et sociétaux relatifs aux enjeux du numérique. Elle est élue locale à Saint-Denis-des-Puits depuis 2001 et conseillère départementale du canton d'Illiers-Combray depuis 2015.

→ **Pour un développement de l'IA au service du bien commun**

**Mathieu DELEMME** est un entrepreneur social et environnemental dans le secteur du numérique. Il préside Ctrl-a, premier groupe numérique pour le secteur de l'intérêt général. Ctrl-a rassemble les agences :

- ecedi, reconnue depuis 2001 comme partenaire privilégié des projets *web* d'intérêt général et d'utilité publique ;
- numerik-ea, entreprise adaptée dédiée aux métiers du *web* ;
- Temesis, agence de référence sur le numérique responsable ;
- Smart Agence, spécialisée dans l'engagement digital.

Depuis plus de 20 ans, il œuvre au service d'un numérique juste et utile, accompagnant des organisations publiques, associatives et privées dans la définition et la mise en œuvre de leurs stratégies de communication numérique et dans le déploiement de leurs dispositifs digitaux.

Engagé de longue date pour porter une communication numérique qualitative, porteuse de sens et au service de l'intérêt général, il a contribué dès 2005 à la mise en œuvre du RGAA (Référentiel Général d'Amélioration de l'Accessibilité). Il a lancé en 2018 et dirige depuis le groupe de travail « Éco-conception des services numériques » au sein de l'Institut du Numérique Responsable (INR). Les travaux de ce groupe ont servi de fondation pour l'élaboration du RGENS (Référentiel général de l'écoconception des services numériques) et sont utilisés par le W3C pour établir les standards internationaux d'écoconception.

→ **Intégrer l'IA dans un service éco-conçu : oxymore ou réalité ?**

**Marie-Laure DENIS**, diplômée de l'Institut d'études politiques de Paris en 1988 et ancienne élève de l'École Nationale d'Administration (promotion « Condorcet »), a consacré 20 ans de son parcours professionnel à la régulation au sein d'autorités administratives indépendantes.

Membre du Conseil d'État, elle débute sa carrière dans l'administration en 1990 et exerce des fonctions comme membre ou directrice de cabinets de collectivités locales ou ministériels. En 2004, Marie-Laure Denis devient membre du Conseil supérieur de l'Audiovisuel (devenu l'Arcom) jusqu'en 2011, date à laquelle elle rejoint l'Autorité de régulation des communications électroniques et des postes (Arcep), également comme membre du collège,

jusqu'en 2017. Elle a également été membre de la commission des sanctions (CORDIS) de la Commission de régulation de l'Énergie (CRE) de 2017 à 2019.

Depuis le 2 février 2019, elle est présidente de la CNIL. Son mandat a été renouvelé par décret du président de la République du 30 janvier 2024 pour une durée de 5 ans.

→ ***IA et libertés : un défi pour la régulation***

**Thomas DUMORTIER** est titulaire d'un doctorat en droit public. Ses travaux de recherche portent sur l'encadrement des libertés publiques, la bioéthique et les discriminations. Après sa thèse, il a également collaboré en tant qu'expert aux rapports de l'Agence des droits fondamentaux de l'Union européenne. Il est conseiller juridique au sein de la Commission nationale consultative des droits de l'homme depuis 2016.

→ ***L'intelligence artificielle et les droits humains : les insuffisances du cadre européen***

**Brice GAILLARD** est diplômé d'EPITECH en 2008, suivi d'un Executive MBA en 2014. Il s'est consacré depuis près de 20 ans à l'innovation dans les technologies éducatives. Dès l'âge de 21 ans, il a fondé sa première entreprise, affirmant un goût marqué pour l'entrepreneuriat et la transformation digitale.

Il est le directeur général d'Apolearn, une entreprise qu'il a créé en 2015 avec Sami Labidi, et qui s'est imposée comme une référence incontournable dans le domaine des Edtech.

Avec Apolearn, Brice Gaillard contribue à accompagner les professionnels de la formation et les entreprises dans leur transition numérique. La plateforme de digital *learning* Apolearn se distingue par sa capacité à conjuguer technologie de pointe et simplicité d'utilisation. Aujourd'hui, Apolearn s'est imposée comme un outil apprécié internationalement, avec plus de 2 millions d'apprenants formés à travers 50 pays. Adaptée à tous les formats pédagogiques, elle propose une expérience intuitive pour les formateurs et les apprenants, intégrant des outils d'intelligence artificielle, des fonctionnalités multi-modales et des solutions personnalisées pour accompagner la transformation digitale des organismes de formation.

À travers Apolearn, il poursuit une mission claire : démocratiser l'accès à des outils de formation performants et transformer la manière dont les savoirs sont transmis dans un monde en constante mutation.

→ ***Le formateur augmenté : entre intelligence artificielle et intelligence émotionnelle***

**Franck GAMBELLI** est actuellement conseiller de l'Union des industries et métiers de la métallurgie (UIMM), de formation juridique à la Sorbonne, il a été directeur juridique de la fédération française des industries mécaniques jusqu'en 2013 et directeur Santé, Sécurité, Environnement et RSE à l'UIMM de 2003 à 2023. Il est l'auteur de plusieurs publications juridico-technique notamment d'études sur le droit des législations techniques, de la normalisation et des relations intra-industriels.

Il a été administrateur de l'INRS, l'Ineris, l'Anses, et l'Anact, président de la branche accidents du travail de la Sécurité sociale de 2005 à 2014. Actuellement il est membre du Comité consultatif pour la Santé au travail de la Commission européenne, administrateur de l'Agence de Bilbao (EU OSHA) et du Cofrac.

→ ***Législations et dialogue social européens autour de l'intelligence artificielle***

**Frédéric GARCIA** est directeur de recherche Inrae au sein de l'unité MIAT. Chercheur en intelligence artificielle, ses travaux visent à développer des méthodes de modélisation, de simulation et d'optimisation des comportements intelligents. Il a dirigé le département de Mathématiques et Informatique appliquées de l'Inrae (2012-2016), a été le premier délégué au numérique de l'Inrae (2015-2016) et co-dirigé l'Institut Convergences #DigitAg sur l'agriculture numérique (2016-2020). Il est aujourd'hui missionné sur l'Intelligence

artificielle par son département Mathématique et Sciences du Numérique. Ses recherches actuelles, dans le domaine de la cognition végétale, portent sur l'étude et la modélisation du comportement biophysique de défense des plantes.

Fortement investi depuis 2012 dans la formation des doctorants et des encadrants au sein du dispositif Eden, il accompagne l'autonomisation intellectuelle du jeune chercheur dans un contexte d'une recherche interdisciplinaire aux enjeux sociétaux forts. Il encourage l'inscription des pratiques de recherche dans une démarche réflexive et ouverte sur les interactions, et défend en particulier l'idée d'un rapprochement entre art et science, quant à leurs liens à l'expérimentation, la créativité et l'innovation.

→ ***Les impacts de l'IA sur l'environnement***

**Auréli GIARD-JACQUET** est une spécialiste de l'intelligence artificielle générative et une défenseure engagée des talents féminins dans le secteur technologique. Forte d'une carrière en *marketing* opérationnel dans des groupes internationaux tels que EDS, Pitney Bowes et Allianz Trade, elle a développé une expertise à la croisée de la stratégie, de l'innovation et des nouvelles technologies. En 2018, elle a enrichi ses compétences en *blockchain* et intelligence artificielle grâce au programme exécutif conjoint de HEC Paris et des Mines ParisTech.

Depuis 2020, elle produit des contenus pour démocratiser les technologies émergentes et encourager les femmes à se tourner vers les métiers d'avenir. Avec son *podcast* Techlipstick, elle met en lumière des expertes et entrepreneuses du domaine. En 2023, elle a fondé Business Crush, une entreprise dédiée à des formations pratiques et adaptées aux besoins des entreprises, s'appuyant sur sa connaissance approfondie des processus organisationnels.

→ ***AI-xiety : entre mythes et réalité,  
la superintelligence artificielle est-elle déjà là ?***

**Luciana GONDIM DE ALMEIDA GUIMARÃES** est titulaire d'un doctorat en génie du bois et des matériaux biosourcés de l'Université Laval, à Québec, Canada. Elle possède également une maîtrise en logistique et recherche opérationnelle obtenue à l'Université Fédérale du Ceará, ainsi qu'un second doctorat en administration de l'Université Fédérale de Pernambuco. Elle est actuellement professeure à l'Université Potiguar (UNP) et chercheuse à l'École de Technologie supérieure (ÉTS).

→ ***Enjeux et perspectives pour une IA éthique et durable***

**Didier GUÉVEL** est professeur émérite de Droit privé et Sciences criminelles de l'Université de Sorbonne Paris Nord, membre de l'IRDA Paris XIII (ER 3970), doyen honoraire de la Faculté de Droit, Sciences politiques et sociales. Il est également ancien directeur du département de Droit de l'Université du Havre ; membre associé du LexFEIM Le Havre (EA 1013) ; membre du conseil scientifique de la MSH Paris Nord (USR 3258) ; membre du comité multi parties-prenantes du Forum annuel de la gouvernance d'internet ; Profesor Honorario de l'Université de Mendoza, Docteur d'État (1990 ; sujet de la thèse : Essai sur une tendance contemporaine : la dispersion de la fonction juridictionnelle). Il est enfin auteur d'ouvrages juridiques et de très nombreux articles privilégiant l'interdisciplinarité.

→ ***Le Droit à l'heure de "l'intelligence artificielle"***

**George KASSAR** est enseignant permanent, membre du corps professoral et de l'équipe de recherche à Ascencia Business School. Il détient un Doctor of Business Administration (DBA), un Postgraduate Certificate in Business Research Methods (PgCert), un Master of Business Administration et un Bachelor of Science - Management Information Systems. Après avoir consacré plus de 16 ans au secteur bancaire, principalement dans la gestion de la performance et l'analyse commerciale, il s'est orienté vers le monde académique, ayant enseigné dans le supérieur depuis 2009. Il s'investit depuis dans l'enseignement des disciplines fondamentales du commerce, notamment le management, le *marketing*

et la gestion des systèmes d'information, avec l'objectif de former des professionnels compétents et visionnaires.

Ses intérêts de recherches portent sur les stratégies de gestion de la performance, tant au niveau individuel qu'organisationnel. Il s'intéresse particulièrement à l'utilisation des outils technologiques et des analyses commerciales pour fournir aux gestionnaires des perspectives approfondies et les aider à prendre des décisions éclairées. Par ailleurs, il explore les pratiques de gestion des connaissances qui peuvent stimuler l'apprentissage et l'innovation, contribuant ainsi à l'amélioration continue des performances des organisations et à leur adaptation aux évolutions du marché.

→ ***Santé mentale au travail et intelligence artificielle :  
entre soutien psychologique et risque de dépendance***

**Orianne LEDROIT** travaille depuis plus de 15 ans dans le numérique ; développant une expertise des impacts sociaux, sociétaux et environnementaux des technologies.

Entrée dans le numérique en mobilisant des logiciels d'amélioration de la relation d'une ville à ses habitants, elle s'engage progressivement dans des collectifs œuvrant pour l'ouverture des données et la participation citoyenne dans des projets urbains mais aussi pour une meilleure compréhension des transitions à l'œuvre en initiant des démarches de « Voyages Apprenants ».

À partir de 2016, elle pilote la politique gouvernementale de lutte contre « l'illectronisme » au ministère de l'Économie et des Finances. En 2021, elle coordonne les opérations d'Openclassrooms, *scale-up* qui tire parti des technologies pour favoriser l'acquisition de compétences et l'insertion dans l'emploi.

Orianne Ledroit dirige aujourd'hui l'association EdTech France qui rassemble plus de 420 entreprises françaises mobilisant les technologies au service de l'éducation et de la formation.

Après avoir co-animé durant 4 ans à Sciences Po Lille un cours de M2 intitulé « Numérique, démocratie et action publique », elle a pris en 2021 la responsabilité de plusieurs modules au sein de l'Executive Master « Humanités numériques » de Sciences Po Paris.

Parce que le numérique porte à la fois des opportunités et des risques, elle cherche à agir pour massifier les premières tout en limitant les seconds.

→ ***Construisons un cadre ambitieux et apaisé pour mettre les IA  
au service de l'éducation et de la formation***

**Jean-François LUCAS**, docteur en sociologie, est délégué général de Renaissance Numérique. Par le passé, il a été directeur conseil chez Auxilia et Chronos, accompagnant des collectivités territoriales et des entreprises dans le développement de territoires dits intelligents, au service d'une meilleure inclusion et du développement durable. Il a également travaillé de nombreuses années dans le champ de la conception et du management de projets de recherche innovants (École polytechnique fédérale de Lausanne ; Télécom ParisTech) regroupant des acteurs académiques (en sciences humaines et sociales et en ingénierie), industriels, publics, ainsi que des artistes.

→ ***Pour une éducation à la pluralité des altérités et des attachements  
numériques à l'heure de l'intelligence artificielle générative***

**Christian MAKAYA** est enseignant-chercheur à Ascencia Business School, à Paris, et chercheur associé à l'université Paris-Nanterre (laboratoire CEROS). Ses travaux de recherche, publiés dans des revues comme le *Journal of Business Research*, la *Revue française de gestion*, la *Revue Internationale de Psychologie et Comportements Organisationnels* ou *Entreprendre et Innover*, portent sur un renouvellement critique de la pratique entrepreneuriale dans le contexte de crise environnementale et climatique. Il travaille notamment à inscrire l'entrepreneuriat dans une réconciliation entre les formes du vivant et s'intéresse aux communautés marginalisées, notamment du fait de leur diversité cognitive et psychologique. Il est responsable du développement et de la valo-

risation de la *Revue française de gestion* et membre du *board* éditorial de l'*International Journal of Technoentrepreneurship*.

→ ***Santé mentale au travail et intelligence artificielle : entre soutien psychologique et risque de dépendance***

**Axel MAMOU-MANI** est le fondateur de la Rainer School, une école innovante dédiée à l'IA générative et à l'automatisation, qui forme et accompagne vers l'emploi des jeunes en difficulté, dont les métiers sont particulièrement menacés par l'automatisation et les évolutions technologiques. Diplômé de l'ESCP, il a débuté sa carrière comme consultant en stratégie digitale chez Accenture, avant de rejoindre Louis Vuitton en tant que chef de projet en innovation digitale, où il a piloté des projets stratégiques fondés sur la *blockchain*. Engagé dans l'insertion professionnelle à travers l'éducation et l'entrepreneuriat, il aspire à transformer les défis de la révolution technologique en opportunités concrètes pour les jeunes, en leur offrant les compétences et les perspectives nécessaires pour réussir dans le monde de demain.

→ ***IA et transformations des métiers : création ou destruction ?***

**Guy MAMOU-MANI** est entrepreneur, *business angel*, enseignant en école de commerce et à Sciences Po, et conférencier. Ancien co-président du Groupe Open, il a débuté sa carrière chez CSC-GO International avant de fonder en 1995 la filiale française de Manugistics, leader européen du *supply chain management*. Il a présidé Syntec Numérique (2010-2016) et MiddleNext (2006-2010) et a été vice-président du Conseil national du Numérique (2016-2017). Engagé pour l'égalité, il est cofondateur du mouvement #JamaisSansElles. Auteur de *L'Apocalypse numérique n'aura pas lieu* et *Pour un numérique humain*, il est Chevalier de la Légion d'Honneur.

→ ***IA et transformations des métiers : création ou destruction ?***

**Frédéric MARCHAND**, fort de 25 ans d'expérience dans la transformation digitale des organisations, milite pour un numérique plus inclusif plus respectueux de la planète. Co-fondateur et CEO de Digital4Better, acteur de l'ESS dont la raison d'être est d'agir pour un numérique plus responsable. Digital4better est l'éditeur de la solution fruggr, cette plateforme labellisée GreenTech Innovation par le ministère et lauréat du concours d'innovation iNov permet aux grandes organisations de mesurer, comparer et améliorer leur empreinte numérique.

Il a été élu dans les 100 personnalités en France qui donnent un sens au numérique. Son livre, *40 mots pour un numérique responsable*, propose un index didactique d'un numérique à impacts positifs. Il a également participé avec des personnalités de la RSE à l'écriture d'un livre collectif, *Les 101 mots de la RSE*.

→ ***L'IA durable n'existe pas***

**Jean-Luc MARINI**, avec plus de 30 ans d'expérience en IA, rejoint OpenStudio en tant que directeur du Laboratoire d'Intelligence Artificielle (Lab IA) et de l'agence de Lyon. Il est multi-diplômé avec un doctorat en sciences de l'information, une maîtrise en Mathématiques appliquées, et un diplôme en informatique de gestion de l'IAE de Lyon, où il a également enseigné pendant 25 ans.

Il a un esprit entrepreneurial marqué par l'innovation et a co-fondé plusieurs entreprises, dont Servidoc et Search'XPR. Son arrivée chez OpenStudio apporte un *leadership* fort, une créativité ainsi qu'une expertise supplémentaire en IA où il dirige le pôle de recherche et développement.

→ ***Intégrer l'IA dans un service éco-conçu : oxymore ou réalité ?***

**Loubna MOURTAJJI**, docteur ès Sciences de Gestion, Maîtresse de Conférences, membre du LEFMI (Laboratoire d'Économie, Finance, Management et d'Innovation - UR4286), référente d'innovation pédagogique à l'UPJV et membre du groupe MAIA

(Maîtrise des Applications en Intelligence Artificielle). Ses travaux de recherche portent sur l'analyse de discours, l'impact de l'intelligence artificielle sur différents domaines (ESG, *marketing*) et l'innovation pédagogique.

→ ***Profils d'appropriation de l'intelligence artificielle générative dans l'éducation***

**Emmanuel NURIT** a d'abord été *freelance* dans le numérique avant de fonder en 2006 Openstudio, entreprise de services digitaux dont il est président, spécialisée dans les projets de développement sur mesure. Il défend et promeut depuis l'origine un numérique responsable et ouvert. C'est pourquoi OpenStudio a toujours contribué à des projets *open source*, et notamment le CMS e-commerce thelia dont l'entreprise est éditrice. OpenStudio a également pris très tôt la mesure de l'empreinte environnementale de son activité et a engagé sa transition écologique et énergétique, qui s'est notamment traduite la labellisation « Numérique Responsable » en 2021.

→ ***Intégrer l'IA dans un service éco-conçu : oxymore ou réalité ?***

**Stéphanie POTTECHER**, diplômée de Telecom Paris, est une ingénieure et entrepreneure passionnée par l'intelligence artificielle et la transmission des savoirs. Ancienne enseignante en mathématiques et informatique au lycée et dans le supérieur, elle a également occupé des postes stratégiques en *data science* et direction technique dans des entreprises comme la Société générale, Orange et Biogen. En 2022, elle crée Beauvoir, une école inclusive dédiée aux arts numériques, avec pour mission de préparer les talents de demain dans des domaines tels que le *web*, l'intelligence artificielle, la *data*, l'automatisation et le *marketing* digital. Elle incarne une vision innovante et inclusive de la formation, inspirant de nouvelles vocations dans un secteur en pleine évolution.

→ ***AI-xiety : entre mythes et réalité, la superintelligence artificielle est-elle déjà là ?***

**Romain RABIER**, expert en transformation complexe des organisations depuis 20 ans, d'abord sur les enjeux de transformation digitale et d'innovation, puis sur l'évolution des métiers et des compétences, a fondé en 2020 le cabinet Smart Leaders, après 15 ans passés dans divers cabinets de conseil.

Le rôle de Smart Leaders est d'accompagner les dirigeants, les DRH et les managers à garder un haut niveau de compétence à l'ère de l'IA et à donner à leur organisation un coup d'avance grâce à une gestion des compétences prospective et stratégique. Doté d'une plateforme innovante dédiée au futur du travail et fortement axé sur le travail entre pairs, Smart Leaders a publié en 2024 l'étude « IA en entreprise, du buzz à l'action » qui a impliqué une quarantaine de professionnels.

Auteur du livre éponyme *Smart Leaders*, le guide pratique des compétences du futur, édité aux éditions Ellipses, Romain Rabier est également membre du Centre des Jeunes Dirigeants (CJD) au sein duquel il pilote un groupe de travail sur l'IA. Il est enfin très impliqué dans la transmission de savoir en tant que mentor d'étudiants, professeur vacataire à ESCP Business School et référent jeunesse au sein du CJD92.

→ ***L'organisation IA-compatible ou l'art de savoir recruter la technologie***

**Romuald RIBAUT** est le fondateur de la marque Neonumeric dans les années 2000, devenant l'un des premiers importateurs de clés USB et de baladeurs MP3 en France. Concerné par l'impact environnemental des produits électroniques, il s'est ensuite impliqué dans la création de la filière nationale agréée de recyclage des DEEE en France en 2005. Il rejoint l'éco-organisme Ecologic en 2009 au poste de directeur *marketing* jusqu'à ce jour. Il a rejoint l'Alliance Green IT (AGIT) en 2011, devenant le secrétaire général en 2013, le trésorier en 2017 et le vice-président en 2020.

Il est également impliqué depuis plus de 15 ans dans des organisations engagées dans les transitions écologiques, la RSE et l'économie circulaire, notamment aux côtés de Corinne



Lepage et Myriam Maestroni. Il est également investi au sein de la commission nationale « Nouvelles responsabilités entrepreneuriales » du Medef depuis 2008 et intervient comme expert Green IT au sein de l'Institut national de l'Économie circulaire.

→ *Intégrer l'IA dans un service éco-conçu : oxymore ou réalité ?*

**Emmanuelle ROUX** est portée par la volonté de relever les grands défis de notre époque, qu'il s'agisse de l'Anthropocène, des enjeux démocratiques ou de l'évolution technologique – et notamment l'impact de l'intelligence artificielle sur nos modes de vie et nos organisations. Depuis plus de 15 ans, elle accompagne dirigeants et équipes dans leurs transformations stratégiques, en veillant à concilier innovation, performance et responsabilité. Son parcours est marqué par plusieurs étapes clés. En 2009, elle découvre la puissance de l'*open source* et s'implique dans le mouvement « Maker », puis crée, dès 2012, ses premiers tiers lieux dédiés au partage et à la collaboration. En 2016, elle fonde leChaudron.io pour encourager une appropriation inclusive de la *data* et de l'IA ; cette initiative donnera lieu, entre autres, à la création d'une *data school* pour le groupe Crédit du Nord. En 2018, elle lance fabriquerdemain.io, prolongement naturel de son engagement à concevoir de nouveaux modèles de société, d'économie et de gouvernance en mobilisant l'intelligence collective.

Persuadée que la transformation de nos modes de vie, pour préserver notre habitabilité planétaire, doit se faire dans une dynamique de coopération, elle facilite la mise en réseau d'acteurs variés et s'attache à rendre accessibles des concepts complexes, de l'innovation à l'intelligence artificielle. Son objectif est clair : stimuler la curiosité, la capacité d'action et la co-construction au service des projets de demain.

En 2023, Emmanuelle Roux rejoint l'Institut national de l'information géographique et forestière (IGN) en tant que conseillère du directeur général, devenant ainsi « cartographe de l'Anthropocène ». À ce titre, elle contribue à la stratégie de l'opérateur public de la donnée et s'engage à promouvoir un usage responsable et ouvert des informations géographiques, au bénéfice de l'ensemble des territoires et de leurs habitants.

→ *IA et Communs : conjuguer puissance technologique et habitabilité terrestre*

**Camille SALINESI** est Professeur des Universités en informatique à l'Université Paris 1 Panthéon-Sorbonne, où il co-dirige l'Observatoire de l'Intelligence artificielle. Spécialiste reconnu en conception de logiciels et en ingénierie des exigences, ses travaux explorent notamment les enjeux sociétaux des technologies à base de logiciels – dont en particulier l'IA, tels que l'éthique, la confiance, la transparence, l'*empowerment* et les biais, mais aussi les enjeux démocratiques de l'IA, comme par exemple le rôle de l'IA dans les phénomènes de désinformation, de radicalisation, de haine en ligne, et de discrimination.

Il est également engagé dans des initiatives interdisciplinaires visant à promouvoir l'inclusion et l'équité dans le développement des technologies numériques. Il contribue activement à des projets de recherche alliant sciences humaines et sociales et technologies, et milite pour une IA responsable et inclusive, qui au-delà de l'innovation, vise le progrès.

→ *Vers une intelligence artificielle “gender by design” ?*

**Sophie SCHBATH** est directrice de recherche INRAE dans l'unité MaIAGE qu'elle a dirigée de 2012 à 2022. Ses travaux de recherche, à la croisée des statistiques et de la bioinformatique, ont principalement porté sur l'analyse statistique des séquences génomiques et leur composition en motifs inattendus. Elle s'est fortement impliquée dans l'animation de la communauté bioinformatique française en co-crédant et co-dirigeant le GDR de Bioinformatique Moléculaire (2006-2013) puis en présidant la société savante SFBI (2010-2016). Depuis 2016, elle est la responsable scientifique de la plateforme de bioinformatique Migale. En 2022, elle intègre le comité exécutif du Grand Défi « Ferments du Futur » de la stratégie France 2030.

Depuis 2020, elle a infléchi ses activités scientifiques et d'animation vers des questions en lien avec les impacts environnementaux de la recherche. Elle fait ainsi partie du Comité d'Orientation et de Suivi de la politique Responsabilité Sociétale et Environnemental d'Inrae ; elle est missionnée par son département Mathématique et Sciences du Numérique pour animer la thématique « Frugalité Numérique » ; et, en tant que relai développement durable de son unité, elle anime la démarche de réduction de l'empreinte carbone du laboratoire.

→ *Les impacts de l'IA sur l'environnement*

**Boris SIRBEY**, docteur en philosophie et entrepreneur visionnaire, se consacre à la transition vers une civilisation plus consciente et éveillée. Issu d'une famille artistique et marquée par les conflits en ex-Yougoslavie, il explore depuis toujours la relation entre l'individu et la société.

Entre 2012 et 2024, il co-fonde MyJobCompany, CollectivZ, ze.game, le Lab RH, France Apprenante et Tomorrow Theory, organisations qui cherchent toutes à mettre l'innovation au service de l'émergence d'une nouvelle civilisation.

Parmi ses contributions majeures figure *Le Point Zéro*, qui est à la fois un cycle d'ouvrages et une boîte à outils (fresque, diagnostic, parcours de transformation...) pour naviguer dans le monde actuel et anticiper celui de demain. *Le Point Zéro* explore cinq transitions clés – économique, organisationnelle, civilisationnelle, anthropologique et existentielle – et propose 25 sous-scénarios, permettant aux organisations de choisir leur avenir désiré. Boris Sirbey intervient régulièrement en tant que conférencier, partageant sa vision d'une société où l'intelligence collective et la conscience individuelle sont au cœur de la transformation civilisationnelle. Son approche unique, alliant philosophie, innovation RH et engagement sociétal, inspire de nombreux acteurs à repenser les paradigmes actuels et à œuvrer pour une métamorphose durable de nos structures sociales et organisationnelles.

→ *IA et futur de la civilisation : dystopie transhumaniste ou métamorphose créatrice ?*

**Laure TABOUY** est à la fois neuroscientifique et éthicienne, enseignante-chercheuse en éthique des neurosciences et du numérique. Guidée par la passion pour la recherche médicale et scientifique en constante évolution, elle a construit et personnalisé son parcours au fur et à mesure des rencontres, ce qui lui a permis d'expérimenter différents milieux de travail, publics et privés, de partir à l'étranger et de travailler avec des personnes venant de tous horizons.

Elle est diplômée d'un doctorat en neurosciences de l'Université Paris-Cité, d'un master en éthique de l'Université Paris-Saclay, et d'une certification de Digital Ethics Officer (2023) de l'EDHEC BS-ALLL. Depuis 2021, elle s'est engagée dans la réalisation d'un deuxième doctorat philosophie spécialité éthique / neuroéthique au sein du Centre Gilles Gaston Granger (UMR 7304) de l'Université l'Aix-Marseille

Par son travail, elle réfléchit sur la manière de définir et de mettre en œuvre une approche *neuroethics by design* pour les neurotechnologies connectées à l'IA, et sur les enjeux de la convergence des neurosciences avec le numérique. Elle fait des liens entre le monde de la recherche académique et le monde institutionnel, au service de la société, de l'être humain et du bien commun.

→ *Les défis éthiques de la convergence de l'IA, des neurosciences, de l'informatique et de l'ingénierie*

**Claire VERDIER** est doctorante en mathématiques appliquées chez OpenStudio. Elle se spécialise dans les applications de l'intelligence artificielle (IA) et travaille sur des solutions visant à rendre l'IA plus durable et frugale. Dans le cadre de ses recherches, elle collabore sur des projets liés à l'optimisation des ressources technologiques et à la réduction de l'impact environnemental de l'IA.

Elle est également impliquée dans des initiatives liées à la création d'une norme européenne pour l'IA frugale, en collaboration avec l'Afnor. Cela fait partie d'un mouvement vers un numérique plus responsable, dans lequel les technologies *open source* jouent un rôle clé pour développer des solutions efficaces tout en minimisant les coûts environnementaux. En tant que *data scientist*, elle travaille sur des projets tels que le traitement automatique des documents (IDP).

→ *Intégrer l'IA dans un service éco-conçu : oxymore ou réalité ?*

**Peggy VICOMTE**, immergée depuis une vingtaine d'années dans l'écosystème de l'innovation et du numérique, est aujourd'hui déléguée générale de Femmes@Numérique, une fondation regroupant plus de 40 grandes entreprises et 50 associations œuvrant pour une meilleure représentation des femmes dans les filières du numérique.

Issue d'une double formation administration d'entreprises et stratégie de communication, elle a travaillé dès les années 2000 avec une diversité d'acteurs issus de la formation, de la recherche, de l'industrie, de l'entrepreneuriat ou de l'investissement pour développer des projets dans une logique de partenariat entre public et privé. Au fil des années, elle occupe ainsi des fonctions variées telles que responsable communication ou responsable développement. Avant de rejoindre Femmes@Numérique en 2021, elle évolue vers des fonctions de direction d'abord dans le cadre d'un programme d'accélération de *start-ups* du numérique, puis au sein d'un institut de recherche sur le véhicule électrique et autonome. Peggy Vicomte porte aujourd'hui un intérêt marqué pour les politiques publiques et les synergies possibles avec la société civile et le monde économique. En sa qualité d'auditrice du cycle national des auditeurs de l'Institut National des Hautes Études de l'Éducation et de la Formation (IH2EF) et de ses fonctions d'élue au bureau du Conseil Économique, Social et Environnemental d'Île-de-France (Ceser IDF) elle s'intéresse tout particulièrement aux politiques en matière d'éducation, de développement territorial et de prospective.

→ *Vers une intelligence artificielle "gender by design" ?*