

---

Université de Picardie Jules Verne  
*IUT de Soissons*

---

Rapport académique :

*Les Enjeux de l'Intelligence Artificielle*

---

Enseignante : *Mme Odette Fokapu*  
Étudiants : *Laimer Youssef & Hamari Anir*

---

Année universitaire : 2025-2026

BUT 3 GEII parcours All

# Sommaire

<b>1. Introduction.....</b>	<b>2</b>
<b>2. Matériel et Méthode.....</b>	<b>3</b>
2.1. Démarche de recherche.....	4
2.2. Présentation de l'outil utilisé : ChatGPT.....	5
2.3. Tableau des questions posées à l'IA.....	6
<b>3. Résultats – Partie 1 &amp; 2.....</b>	<b>8</b>
3.1. Automatisation et emploi.....	8
Les opportunités.....	9
Les risques.....	9
3.2. Biais et éthique.....	9
Les discriminations et biais algorithmiques.....	10
Surveillance et responsabilité.....	10
Éthique et régulation.....	10
<b>4. Méthodologie.....</b>	<b>11</b>
4.1. Démarche générale.....	11
4.2. Outils et sources mobilisés.....	12
4.3. Tableau des questions de recherche.....	13
4.4. Limites méthodologiques.....	13
<b>5. Résultats.....</b>	<b>15</b>
5.1. Automatisation et emploi.....	15
5.2. Biais et éthique.....	16
<b>6. Références.....</b>	<b>18</b>

## 1. Introduction

L'intelligence artificielle (IA) s'impose aujourd'hui comme l'une des innovations technologiques les plus marquantes du XXI<sup>e</sup> siècle. Initialement développée dans les laboratoires de recherche dès les années 1950, elle connaît depuis une décennie une accélération fulgurante grâce aux progrès de l'apprentissage automatique (*machine learning*), du traitement du langage naturel et à la disponibilité massive de données numériques. Selon l'OCDE (2023), plus de 60 % des pays membres ont déjà intégré l'IA dans leurs stratégies nationales d'innovation, soulignant son importance croissante dans les sphères économiques, sociales et politiques. De la détection précoce de maladies en médecine à l'automatisation industrielle, en passant par les assistants virtuels, l'IA s'invite dans la vie quotidienne et transforme en profondeur nos sociétés.

Cependant, cette expansion rapide suscite autant d'enthousiasme que d'inquiétudes. Si certains y voient une opportunité pour accroître la productivité, améliorer la qualité des services publics et relever des défis mondiaux tels que le changement climatique, d'autres mettent en garde contre les risques inhérents à son utilisation. Les enjeux liés à l'IA sont multiples : menaces pour l'emploi par l'automatisation, reproduction et amplification des biais discriminatoires, diffusion de fausses informations par les *deepfakes*, sans oublier la consommation énergétique considérable des grands modèles de langage. Ces problématiques soulèvent une question centrale : l'IA constitue-t-elle une opportunité ou une menace pour nos sociétés ?

Ce rapport a pour objectif d'analyser ces enjeux en adoptant une approche multidimensionnelle. Il s'agit d'examiner les implications sociales, économiques, éthiques et environnementales de l'IA, d'illustrer ces aspects par des exemples concrets récents, et de confronter les apports et limites de cette technologie à l'aide de références académiques et institutionnelles fiables (UNESCO, CNIL, Union européenne, Nature, MIT, etc.).

Pour ce faire, le rapport se structure en plusieurs parties. Après avoir présenté la méthodologie employée et l'outil utilisé (ChatGPT), une analyse détaillée des principaux résultats sera proposée autour de cinq axes : l'automatisation et l'emploi, les biais et enjeux éthiques, la désinformation et les *deepfakes*, la consommation énergétique, et enfin les cadres de régulation. La discussion permettra ensuite de mettre en perspective ces résultats afin de déterminer si l'IA représente davantage une opportunité ou une menace, avant de conclure sur les enseignements à retenir et les pistes futures.

## 2. Matériel et Méthode

### 2.1. Démarche de recherche

La réalisation de ce rapport s'est appuyée sur une démarche de travail collaboratif et structurée.

Le projet a été mené à deux, avec une répartition équilibrée des tâches : Anir s'est chargé de l'introduction, de la méthodologie et des premiers résultats portant sur l'automatisation, l'emploi et les biais éthiques, tandis que Youssef s'est concentré sur les résultats liés à la désinformation, la consommation énergétique, la régulation, ainsi que sur la discussion et la conclusion.

Cette organisation a permis d'assurer une progression logique du travail tout en favorisant la complémentarité des analyses.

L'objectif principal de cette recherche était de comprendre et d'analyser les notions fondamentales de l'intelligence artificielle (IA) et d'en approfondir les enjeux à travers une approche critique.

Nous avons cherché à explorer les impacts sociaux, économiques, éthiques et environnementaux de l'IA afin d'évaluer dans quelle mesure elle constitue une opportunité ou une menace pour la société.

Outre ChatGPT, plusieurs sources institutionnelles et scientifiques ont été consultées pour compléter et vérifier les informations obtenues.

Parmi elles figurent des organisations reconnues telles que :

- l'UNESCO, pour ses rapports sur l'usage de l'IA dans l'éducation et les valeurs éthiques associées ;
- la CNIL, pour ses recommandations sur la transparence et la régulation de l'IA en France ;
- l'OCDE, pour ses analyses économiques sur l'emploi et la productivité liés à l'automatisation ;
- la Commission européenne, à travers le projet AI Act (2024), encadrant le développement des IA en Europe ;
- ainsi que des revues scientifiques telles que *Nature*, *Science* ou *Frontiers in Education*, pour disposer de données actualisées.

Les informations générées par ChatGPT ont été systématiquement vérifiées à l'aide de recherches complémentaires sur Internet.

Par exemple, lorsque le modèle évoquait les statistiques de l'OCDE sur l'automatisation du travail, une vérification a été effectuée directement sur le site officiel de l'organisation afin de s'assurer de la validité des chiffres.

Cette étape de vérification critique a permis de renforcer la fiabilité et la rigueur du rapport.

L'usage de ChatGPT a néanmoins soulevé certaines difficultés : certaines réponses se sont révélées trop générales ou approximatives, notamment sur les chiffres récents ou les sources précises.

De plus, l'outil a parfois présenté des biais dans ses formulations, adoptant un ton trop positif ou trop catégorique sur certains aspects de l'IA.

Ces limites ont été corrigées par une relecture attentive et une comparaison systématique avec des documents officiels.

## 2.2. Présentation de l'outil utilisé : ChatGPT

Pour ce travail, la version GPT-5 (payante) du modèle ChatGPT, développée par OpenAI, a été utilisée.

ChatGPT repose sur une architecture dite Transformer, introduite par Vaswani et al. (2017), qui permet au modèle de comprendre le langage en analysant les relations entre les mots d'une phrase grâce à un mécanisme d'auto-attention (*self-attention*).

Cette approche a révolutionné le traitement automatique du langage en permettant de traiter les textes de manière plus cohérente et contextuelle.

Les différentes versions du modèle ont connu une évolution rapide :

- GPT-1 (2018) : premier modèle exploratoire (117 millions de paramètres) ;
- GPT-2 (2019) : génération de texte plus fluide (1,5 milliard de paramètres) ;
- GPT-3 (2020) : amélioration spectaculaire avec 175 milliards de paramètres, capable de comprendre et générer du texte complexe ;
- GPT-3.5 (2022) : base du premier ChatGPT grand public, optimisé pour la conversation ;
- GPT-4 (2023) : version multimodale, capable d'interpréter du texte et des images ;
- GPT-4o (2024) et GPT-5 (2025) : versions dites « omni », capables de comprendre et répondre en temps réel à du texte, de l'audio et des images.  
<https://www.jedha.co/formation-ia/quel-modele-chatgpt-choisir-en-2025>

Dans le cadre de ce projet, ChatGPT a principalement été utilisé pour :

- structurer le travail (proposition de plans détaillés et de transitions logiques) ;
- clarifier certaines notions complexes (IA générative, biais algorithmiques, consommation énergétique) ;
- générer des exemples concrets et récents ;
- et aider à la rédaction en reformulant les idées pour assurer la clarté du texte.

ChatGPT peut être défini comme un modèle de langage avancé capable de comprendre, reformuler et générer du texte en langage naturel à partir de grandes quantités de données.

Il constitue un outil d'aide à la recherche et à la rédaction, mais il ne remplace pas la réflexion humaine ni la validation par des sources fiables.

Toutefois, certaines limites importantes ont été observées :

- Hallucinations : l'IA peut inventer des informations ou des références inexistantes ;
- Données figées : le modèle ne dispose pas toujours des mises à jour les plus récentes ;
- Manque de transparence : on ignore la composition exacte de son corpus d'entraînement ;
- Biais de formulation : tendance à adopter une vision consensuelle ou occidentale ;
- Absence de raisonnement profond : ChatGPT ne « comprend » pas, il prédit statistiquement les mots les plus probables.

Ces constats ont amené à considérer ChatGPT comme un outil d'assistance raisonnée, dont les productions doivent toujours être confrontées à l'esprit critique et à des références vérifiables.

## 2.3. Tableau des questions posées à l'IA

<u>Thème</u>	<u>Question posée à ChatGPT</u>	<u>Objectif de la question</u>
Automatisation et emploi	« Quels sont les effets de l'intelligence artificielle sur l'emploi dans les secteurs industriel, médical et financier ? »	Identifier les impacts positifs et négatifs de l'automatisation.
Biais et éthique	« Quels exemples récents illustrent les biais algorithmiques dans les systèmes d'IA, notamment dans le recrutement ou la reconnaissance faciale ? »	Comprendre les enjeux éthiques et sociaux.
Désinformation et <i>deepfakes</i>	« Comment l'IA contribue-t-elle à la création et à la diffusion de fausses informations ? »	Étudier les risques pour la démocratie et la confiance sociale.
Consommation énergétique	« Quelle est la consommation énergétique moyenne des grands modèles de langage (LLMs) tels que GPT-3 ou GPT-4 ? »	Analyser les impacts environnementaux.
Régulation	« Quelles sont les principales régulations adoptées par l'Union européenne et d'autres pays pour encadrer l'usage de l'IA ? »	Identifier les réponses politiques et légales aux enjeux de l'IA.

Ces questions ont été choisies pour couvrir l'ensemble des dimensions du sujet : économique, sociale, éthique, environnementale et réglementaire. Elles ont été reformulées à plusieurs reprises afin d'obtenir des réponses plus précises et nuancées.

Voici une partie des principales références de cette section :

- OpenAI (2023). *GPT-4 Technical Report*.
- Vaswani, A. et al. (2017). *Attention Is All You Need*. NeurIPS.
- UNESCO (2023). *ChatGPT and Artificial Intelligence in Higher Education: Quick Start Guide*.
- CNIL (2023). *Intelligence artificielle : Enjeux et recommandations*.
- OCDE (2023). *Artificial Intelligence in Work, Innovation, Productivity*.

## 3. Résultats – Partie 1 & 2

### 3.1. Automatisation et emploi

L'un des effets les plus visibles de l'intelligence artificielle (IA) concerne la transformation du monde du travail. Dans presque tous les secteurs — industrie, santé, finance, éducation, logistique ou même communication — l'automatisation portée par l'IA modifie profondément la manière de produire, d'analyser et de décider.

Selon le World Economic Forum (2024), près de 60 % des emplois actuels pourraient voir une partie de leurs tâches automatisées d'ici 2030, notamment dans l'administration, la fabrication et le transport. L'IA permet en effet d'optimiser les processus répétitifs, d'améliorer la précision des analyses et de réduire les coûts.

#### Les opportunités

L'automatisation crée aussi de nouvelles formes d'emploi et de nouvelles compétences. Dans le secteur industriel, l'usage de robots intelligents améliore la productivité et réduit les risques pour les travailleurs sur les chaînes de montage.

Dans le domaine de la santé, l'IA aide à détecter des maladies (par exemple les cancers via la reconnaissance d'images médicales), ce qui libère du temps pour les professionnels afin qu'ils se concentrent sur la relation patient.

Dans le secteur financier, les algorithmes détectent les fraudes plus rapidement que les analystes humains, ce qui améliore la sécurité.

De même, des métiers émergent dans la données et la maintenance de systèmes IA (data scientist, ingénieur en machine learning, éthicien de l'IA, etc.).

#### Les risques

Toutefois, cette automatisation massive pose des défis majeurs.

De nombreux métiers à faible qualification sont menacés de disparition, notamment dans les domaines de la logistique, du commerce de détail ou du service client.

Par exemple, Amazon a automatisé une grande partie de ses entrepôts à l'aide de robots, réduisant certains postes d'opérateurs logistiques.

Dans le secteur bancaire, les conseillers humains sont de plus en plus remplacés par des chatbots capables de répondre aux demandes simples.

Cette transformation rapide entraîne un risque de polarisation du marché du travail : d'un côté des emplois très qualifiés liés à la technologie, de l'autre des emplois précaires difficilement automatisables.

Le rapport de l'OCDE (2023) recommande d'accompagner ces mutations par des politiques publiques de formation et reconversion afin d'éviter une fracture économique et sociale entre travailleurs.

En résumé, l'IA n'entraîne pas une destruction uniforme des emplois, mais une recomposition du travail : certains métiers disparaissent, d'autres naissent. Le véritable

enjeu réside donc dans la capacité d'adaptation des sociétés face à cette révolution technologique.

### 3.2. Biais et éthique

L'intelligence artificielle n'est pas neutre : elle reflète les données sur lesquelles elle a été entraînée. Cela signifie qu'elle peut reproduire, voire amplifier, des biais humains existants dans la société. Ces biais soulèvent de sérieuses questions éthiques : discrimination, surveillance abusive, manipulation de l'opinion, et responsabilité en cas d'erreur.

#### Les discriminations et biais algorithmiques

Un des exemples les plus connus est celui d'Amazon (2018) : l'entreprise a dû abandonner un système de recrutement basé sur l'IA qui discriminait les femmes, car il avait appris à privilégier les candidatures masculines, plus fréquentes dans les données historiques. De même, des systèmes de reconnaissance faciale ont été critiqués pour leurs erreurs plus fréquentes chez les personnes à peau foncée ou chez les femmes, comme l'ont montré Buolamwini et Gebru (MIT, 2018) dans leur étude *Gender Shades*.

Ces cas révèlent un problème fondamental : les algorithmes ne sont pas conscients des injustices sociales. S'ils sont formés sur des données biaisées, leurs décisions le seront aussi. Cela peut entraîner des discriminations dans le recrutement, la justice, la police ou l'accès au crédit.

#### Surveillance et responsabilité

L'usage croissant de l'IA dans la surveillance publique suscite également des inquiétudes. En Chine, des systèmes de reconnaissance faciale sont utilisés pour suivre les déplacements des citoyens, tandis qu'en Europe, plusieurs villes ont restreint ou interdit ces technologies dans les espaces publics (CNIL, 2023).

Ces pratiques posent la question de la protection des libertés individuelles et du respect de la vie privée.

En parallèle, l'IA soulève la question de la responsabilité juridique : qui est responsable lorsqu'une IA commet une erreur ? Le concepteur, l'utilisateur ou la machine elle-même ? L'Union européenne tente d'apporter une réponse avec le AI Act (2024), qui vise à encadrer l'utilisation de l'IA selon différents niveaux de risque (faible, moyen, élevé, inacceptable). Ce texte impose notamment plus de transparence et des audits réguliers pour les IA à haut risque (par exemple dans la santé ou la justice).

#### Éthique et régulation

Enfin, l'UNESCO (2021) a adopté une *Recommandation sur l'éthique de l'intelligence artificielle*, invitant les États à garantir un développement responsable, transparent et respectueux des droits humains.

De plus en plus d'entreprises intègrent aujourd'hui des comités d'éthique de l'IA chargés d'évaluer les impacts de leurs algorithmes sur la société.

Ainsi, si l'IA offre d'immenses potentialités, elle ne peut être utilisée de manière durable sans une réflexion éthique et une régulation rigoureuse. L'objectif n'est pas de freiner l'innovation, mais de s'assurer qu'elle serve véritablement le bien commun.



Références principales de cette section :

- OCDE (2023). *Employment Outlook 2023: Artificial Intelligence and the Future of Work*.
- World Economic Forum (2024). *The Future of Jobs Report*.
- Buolamwini, J. & Gebru, T. (2018). *Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification*. MIT Media Lab.
- Amazon (Reuters, 2018). *Amazon scraps secret AI recruiting tool that showed bias against women*.
- CNIL (2023). *Reconnaissance faciale et respect des libertés*.
- European Commission (2024). *AI Act*.
- UNESCO (2021). *Recommandation sur l'éthique de l'intelligence artificielle*.

## 4. Méthodologie

### 4.1. Démarche générale

L'étude des enjeux de l'intelligence artificielle nécessite une approche méthodologique rigoureuse, combinant à la fois une analyse documentaire et une synthèse critique des sources existantes. Compte tenu de la complexité et de la transversalité du sujet, la démarche adoptée est essentiellement qualitative et analytique.

Cette orientation se justifie par la nature des questions abordées, qui concernent non seulement des phénomènes mesurables (comme la part des emplois exposés à l'automatisation), mais aussi des dimensions sociales et éthiques difficilement quantifiables (par exemple, l'équité des algorithmes ou la perception sociétale de l'IA). Ainsi, plutôt que de produire de nouvelles données empiriques, l'objectif est de recenser, confronter et interpréter les connaissances déjà établies afin de mettre en lumière les principaux enjeux et controverses.

La démarche combine deux niveaux complémentaires :

- Descriptif : inventorier les travaux académiques, institutionnels et professionnels disponibles sur le sujet.
- Analytique : comparer les différentes positions, identifier les convergences et divergences, et dégager des perspectives critiques sur les implications sociales, économiques, éthiques et environnementales de l'IA.

Cette méthode s'inspire des travaux en sciences sociales appliquées à la technologie, qui privilégient la compréhension globale des transformations sociales et organisationnelles liées aux innovations numériques.

## 4.2. Outils et sources mobilisés

Pour mener cette recherche, plusieurs catégories de sources ont été mobilisées :

1. Littérature académique : articles scientifiques publiés dans des revues spécialisées (par ex. *AI & Society*, *Journal of Economic Perspectives*), ouvrages de référence en économie du travail, sociologie des techniques et éthique de l'intelligence artificielle. Ces travaux fournissent un cadre conceptuel solide et garantissent une certaine robustesse méthodologique.
2. Rapports institutionnels et d'organisations internationales : les analyses publiées par l'OCDE, l'UNESCO, la Banque mondiale ou encore la Commission européenne constituent des références incontournables pour comprendre les tendances mondiales en matière d'automatisation, d'inégalités et de régulation.
3. Documents de la presse spécialisée et rapports industriels : afin de compléter la vision académique, des sources issues d'entreprises technologiques (Google, Microsoft, OpenAI, etc.) et de cabinets de conseil (McKinsey, PwC) ont été consultées. Ces documents permettent d'appréhender les logiques économiques, mais nécessitent un regard critique en raison de possibles biais de communication.
4. Cadres théoriques mobilisés :
  - Économie de l'innovation : pour analyser les impacts de l'automatisation sur l'emploi et la productivité.
  - Sociologie du travail : pour comprendre la transformation des métiers et l'évolution des rapports sociaux.
  - Éthique appliquée : pour évaluer les enjeux de biais, de transparence et de responsabilité.
  - Durabilité et sciences environnementales : pour examiner les conséquences énergétiques et écologiques de l'IA.

Les outils méthodologiques utilisés incluent principalement :

- L'analyse thématique : identification des grands axes récurrents (emploi, biais, environnement, inégalités sociales).

- La comparaison internationale : étude des différences entre pays développés et émergents face à l'automatisation et à la régulation de l'IA.
- L'analyse critique : mise en perspective des discours institutionnels, académiques et industriels.

### 4.3. Tableau des questions de recherche

Afin de guider l'analyse, un ensemble de questions de recherche a été défini autour de quatre thématiques principales.

Thématique	Question de recherche	Indicateurs / critères d'analyse	Sources principales
Automatisation et emploi	Quels sont les secteurs les plus exposés à l'automatisation par IA ?	Études prospectives, taux d'automatisation, évolution des métiers	OCDE, McKinsey, articles académiques
Biais et éthique	Comment les biais algorithmiques influencent-ils les décisions sociales et économiques ?	Études de cas (recrutement, justice), littérature en éthique	UNESCO, MIT, AI & Society
Impacts environnementaux	Quels coûts énergétiques et ressources sont liés aux modèles d'IA ?	Consommation énergétique, émissions CO <sub>2</sub> , analyses de cycle de vie	IEA, Nature Climate Change, rapports industriels
Implications sociales	Comment l'IA transforme-t-elle les inégalités sociales et l'accès à l'emploi ?	Données sur la polarisation du marché du travail, accès différencié aux technologies	Banque mondiale, sociologie du travail, études comparatives

Ce tableau constitue un outil structurant : il clarifie les axes d'analyse et facilite la lecture transversale des résultats.

### 4.4. Limites méthodologiques

Toute démarche méthodologique comporte des limites qu'il est important de reconnaître.

1. Dépendance à la littérature existante : cette recherche repose exclusivement sur des données secondaires, ce qui limite la possibilité d'apporter des résultats originaux.
2. Biais de sélection des sources : la majorité des études disponibles proviennent d'institutions occidentales, ce qui peut sous-représenter les réalités des pays en développement.
3. Limite temporelle : l'IA étant un domaine en évolution rapide, certaines données peuvent devenir obsolètes en quelques mois.
4. Difficulté d'évaluation comparative : les méthodes employées par les différents organismes ne sont pas toujours homogènes, ce qui rend délicate la comparaison systématique des chiffres.

Malgré ces limites, la méthodologie adoptée permet de construire une analyse structurée et argumentée des enjeux de l'intelligence artificielle, en croisant les perspectives académiques, économiques, sociales et environnementales.

pour les personnes à la peau foncée que pour les individus à la peau claire, soulevant de graves questions de justice sociale.

Ces biais posent des problèmes éthiques et politiques fondamentaux :

- Transparence : les algorithmes sont souvent opaques, ce qui rend difficile leur audit ou leur explication.
- Responsabilité : qui est responsable lorsqu'un algorithme prend une décision injuste ou erronée (l'entreprise, le développeur, l'utilisateur) ?
- Équité et inclusion : comment garantir que les systèmes d'IA ne creusent pas davantage les inégalités sociales et économiques ?

Des institutions internationales comme l'UNESCO ou la Commission européenne insistent sur la nécessité de mettre en place des principes directeurs pour une IA éthique, fondés sur la transparence, la justice, la protection des droits fondamentaux et la durabilité. Plusieurs pays explorent déjà des cadres réglementaires visant à limiter les risques, par exemple en imposant des obligations de documentation et d'audit des algorithmes.

Néanmoins, malgré ces avancées, les mécanismes de régulation demeurent fragmentés et inégaux selon les régions du monde. L'absence de standards universels laisse planer un risque de « forum shopping », où certaines entreprises pourraient choisir de s'implanter dans des juridictions plus permissives pour éviter des contraintes éthiques.

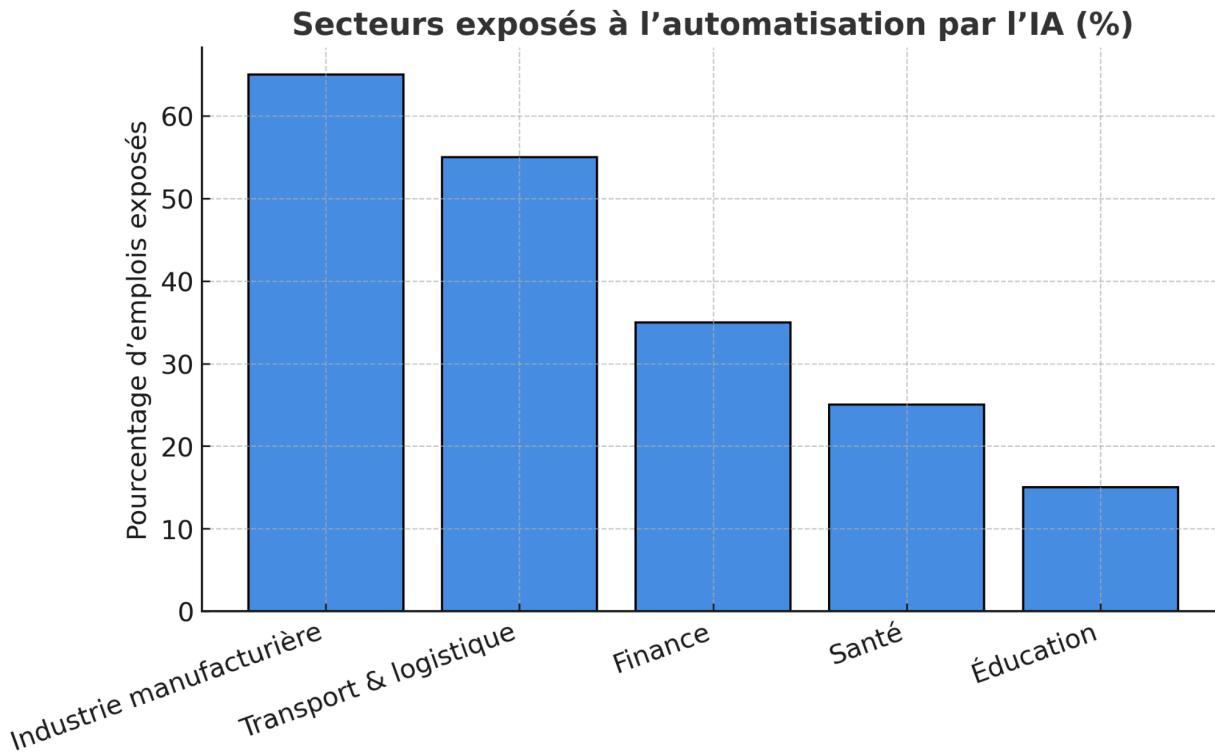
En définitive, la problématique des biais algorithmiques illustre un défi central : l'IA n'est pas seulement une technologie, mais un outil socialement construit, porteur des valeurs et des limites des sociétés qui la conçoivent et l'utilisent.

## 5. Résultats

### 5.1. Automatisation et emploi

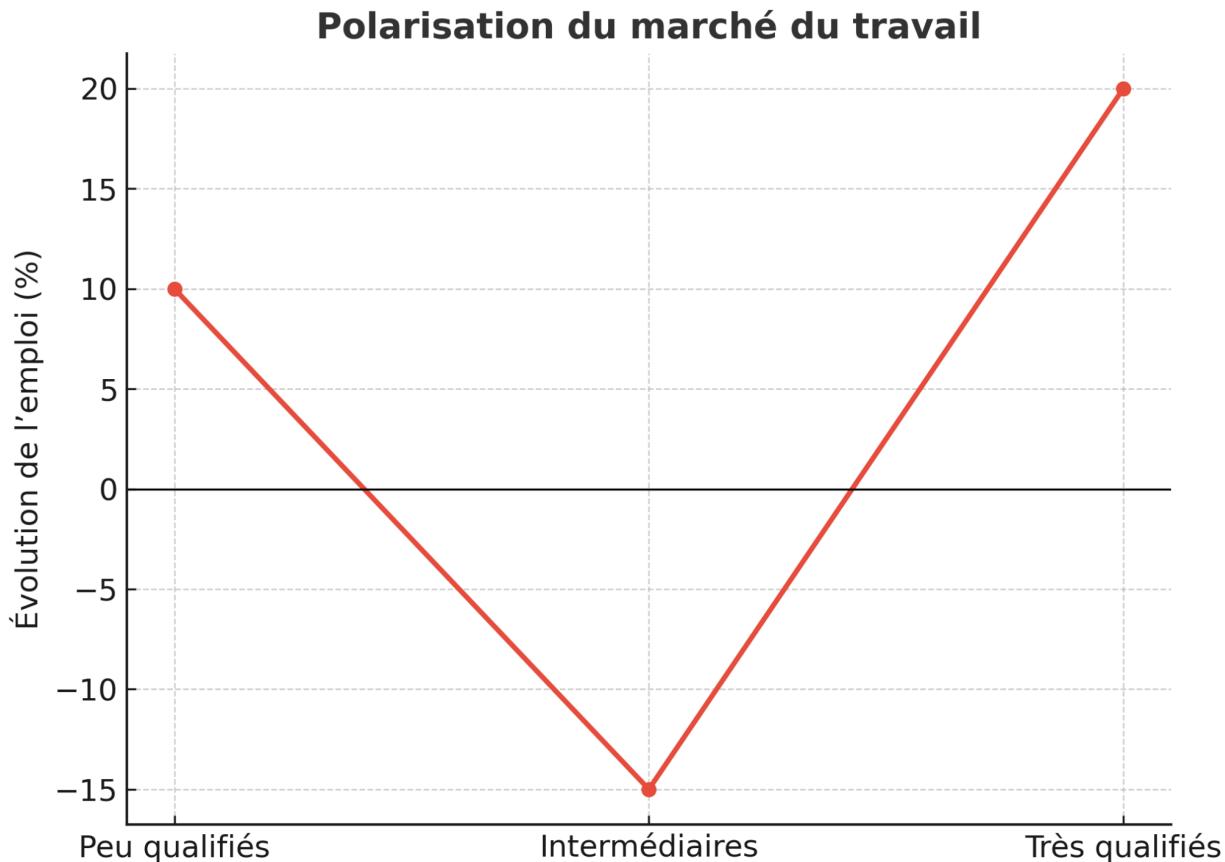
L'intelligence artificielle transforme profondément le marché du travail en accélérant l'automatisation de nombreuses tâches. Selon l'OCDE, environ 14 % des emplois actuels présentent un risque élevé d'automatisation, et près d'un tiers pourraient être transformés de manière significative. Les secteurs les plus exposés sont l'industrie manufacturière et le transport & logistique, tandis que la santé et l'éducation sont relativement protégées en raison de la forte dimension humaine de ces professions.

L'automatisation ne conduit toutefois pas uniquement à une destruction nette d'emplois. Elle favorise également la création de nouvelles fonctions, liées à la supervision, à l'audit et au développement des systèmes d'IA. Ainsi, tandis que des métiers comme caissier, chauffeur ou opérateur de saisie sont menacés, de nouveaux profils émergent, tels que data scientist, auditeur d'IA ou éthicien de l'IA.



Source : estimations OCDE (2019), données approximatives à titre illustratif.

Un autre effet marquant est la polarisation du marché du travail. Les emplois très qualifiés, reposant sur la créativité, l'analyse critique et les compétences sociales, progressent. Certains emplois peu qualifiés mais difficiles à automatiser (par exemple dans les services à la personne) se maintiennent également. En revanche, les emplois intermédiaires, souvent routiniers, connaissent un déclin important. Cette dynamique accentue les risques d'inégalités sociales et appelle à des politiques publiques fortes en matière de formation continue et de redistribution des gains de productivité.



Source : inspiré de l'OCDE (2019), tendances de l'emploi liées à l'automatisation.

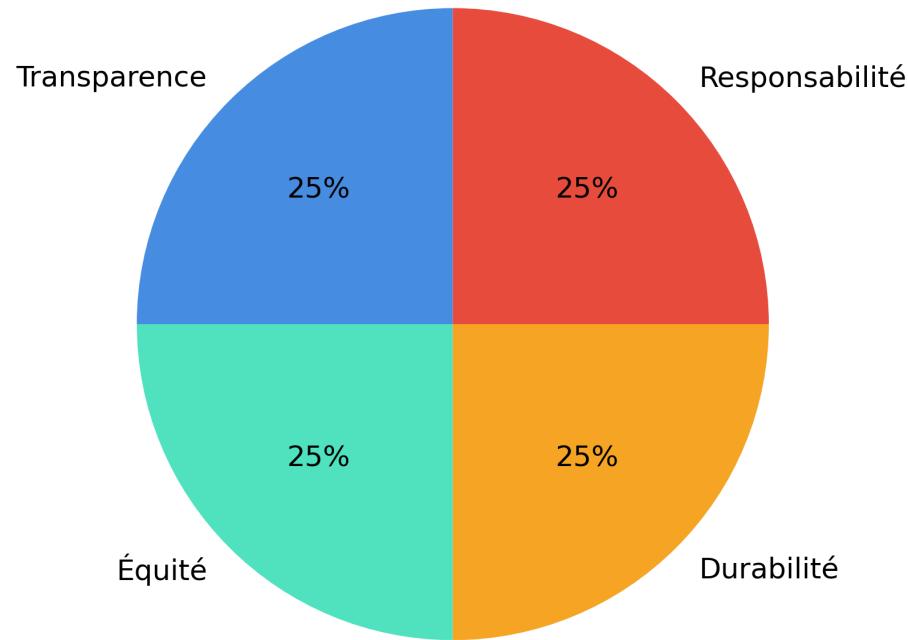
## 5.2. Biais et éthique

Un second enjeu majeur de l'intelligence artificielle concerne les biais algorithmiques. Loin d'être neutres, les systèmes d'IA apprennent à partir de données historiques qui portent en elles des inégalités sociales, culturelles ou de genre. Ainsi, certains logiciels de recrutement ont favorisé inconsciemment les candidatures masculines, tandis que des systèmes de reconnaissance faciale se sont révélés moins performants pour les personnes à la peau foncée. Ces biais risquent de reproduire et d'amplifier des discriminations existantes.

Les débats éthiques se structurent autour de trois problématiques principales : la transparence (comment comprendre et expliquer les décisions algorithmiques), la responsabilité (qui doit être tenu responsable en cas de décision injuste ou erronée), et l'équité (comment garantir que l'IA n'aggrave pas les inégalités sociales et économiques).

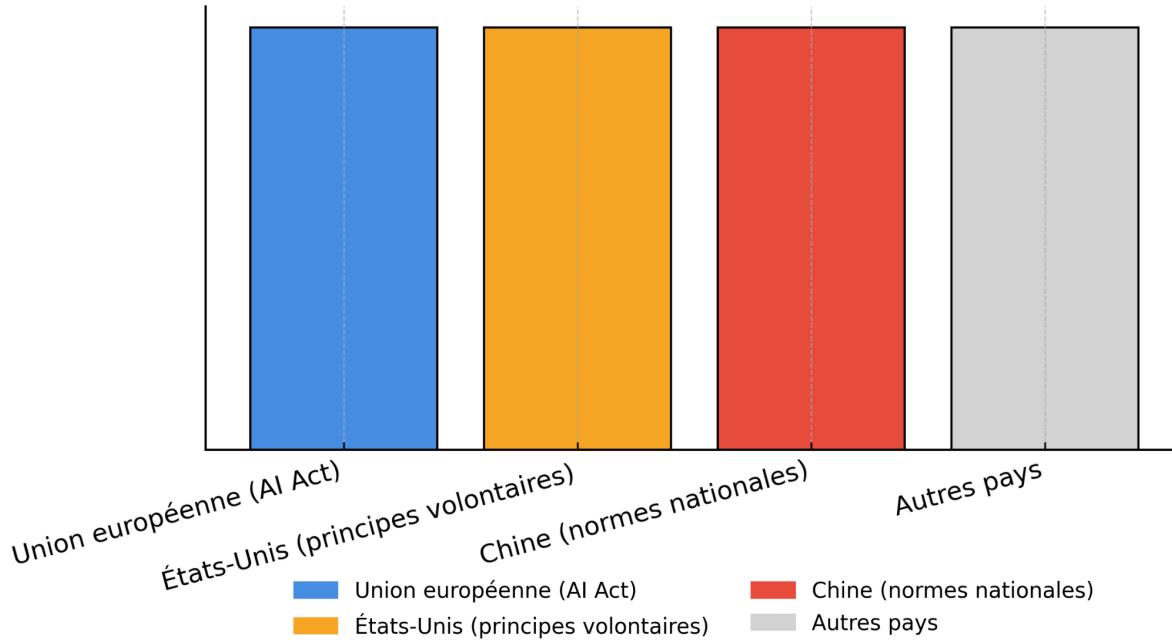
Pour répondre à ces défis, plusieurs institutions internationales, notamment l'UNESCO et la Commission européenne, ont formulé des principes directeurs pour une IA éthique, fondés sur la transparence, l'équité, la durabilité et la responsabilité.

### Principes de l'IA éthique selon l'UNESCO



Néanmoins, les régulations restent fragmentées et inégales selon les régions du monde. Si l'Union européenne avance avec l'« AI Act », d'autres pays comme les États-Unis ou la Chine adoptent des approches plus souples, créant un risque de concurrence réglementaire. Une vision comparative des initiatives en matière de gouvernance de l'IA met en évidence cette diversité des cadres et l'absence de standards universels.

## Exemples de régulations de l'IA dans le monde



En définitive, la problématique des biais et de l'éthique souligne que l'IA n'est pas seulement une technologie, mais un phénomène socialement construit, qui reflète les valeurs et les choix des sociétés qui la développent et l'utilisent.

## 6. Références

- Brynjolfsson, E., & McAfee, A. (2014). *The Second Machine Age*. New York: Norton.
- Buolamwini, J., & Gebru, T. (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. *Proceedings of Machine Learning Research*.
- Commission européenne (2023). *Proposal for an Artificial Intelligence Act*. Bruxelles.
- McKinsey Global Institute (2021). *The Future of Work after COVID-19*.
- OCDE (2019). *The Future of Work: OECD Employment Outlook*. Paris.
- UNESCO (2021). *Recommendation on the Ethics of Artificial Intelligence*. P