

Pour une *intelligence artificielle* au service de l'intérêt général

RAPPORTEURS

Erik Meyer et Marianne Tordeux Bitker

JANVIER 2025

2024-002
NOR : CESL1100002X
Mardi 14 janvier 2025

JOURNAL OFFICIEL
DE LA RÉPUBLIQUE FRANÇAISE

Mandature 2021-2026
Séance du 14 janvier 2025

Pour une *intelligence artificielle* au service de l'intérêt général

Avis du Conseil économique, social et environnemental au nom de la Commission temporaire "intelligence artificielle"

Rapporteurs :
Erik Meyer
Marianne Tordeux Bitker

Question dont le Conseil économique, social et environnemental a été saisi par décision de son Bureau en date du 4 juin 2024 en application de l'article 3 de l'ordonnance n° 58-1360 du 29 décembre 1958 modifiée portant loi organique relative au Conseil économique, social et environnemental. Le Bureau a confié à la Commission temporaire "intelligence artificielle", la préparation d'un avis : *Pour une intelligence artificielle au service de l'intérêt général*. La commission temporaire "intelligence artificielle", présidée par M. Thierry Cadart, a désigné M. Erik Meyer et Mme Marianne Tordeux Bitker comme rapporteurs.

sommaire

Synthèse	4
Préambule – Comprendre le fonctionnement de l'intelligence artificielle (IA)	18
Introduction	29
I - LES GRANDS ENJEUX DE L'IA	33
A. Les risques d'atteintes aux droits fondamentaux et libertés individuelles	33
B. L'empreinte environnementale de l'IA	45
C. L'impératif d'autonomie stratégique européenne pour l'IA	57
II - LES IMPACTS DE L'IA DANS LA SOCIÉTÉ	67
A. Création, suppression, transformation : l'emploi et le marché du travail à l'ère de l'IA	67
B. L'intégration de l'IA en entreprise	87
C. Les services publics et l'IA	100
D. La frontière créée par l'éloignement numérique	108
E. L'IA dans l'éducation	115
F. Santé et intelligence artificielle	129
G. L'IA est-elle génératrice d'accessibilité ? L'exemple du handicap	139
Conclusion	143
Déclarations des groupes	144
Scrutin	160
Annexes	162

synthèse

Les préconisations du CESE s'articulent en deux parties :

I - Les grands enjeux de l'IA

A. Les risques d'atteintes aux droits fondamentaux et libertés individuelles

• Transparence des algorithmes

PRÉCONISATION #1

Le CESE appelle à renforcer, dans la réglementation européenne, les impératifs de transparence des algorithmes et des jeux de données d'entraînement, en :

- exigeant des concepteurs de modèles d'IA qu'ils fournissent des explications claires sur le mode de prise de décision des algorithmes ;
- mettant en place un système d'audit externe régulier, indépendant et public de ces IA par le Bureau de l'IA¹ ;
- incluant des dispositions renforçant l'accès direct des citoyens à leurs propres données ;
- fournissant une meilleure documentation sur la provenance des données, les méthodes de collecte, et les biais

¹ Le Bureau européen de l'IA est le centre d'expertise en matière d'IA dans l'ensemble de l'UE. Il joue un rôle clé dans la mise en oeuvre de la législation sur l'IA – en particulier pour l'IA à usage général. Il est rattaché à la Commission européenne.

potentiels ;

- introduisant des certifications ou des labels de transparence pour les jeux de données partagés
- améliorant la traçabilité des données par la documentation des transformations subies par les données (nettoyage, agrégation, anonymisation) avant leur utilisation ;
- écrivant l'égalité dans le code lui-même pour que l'IA la promeuve.

• Diversité des jeux de données

PRÉCONISATION #2

Favoriser la diversité des données pour réduire les biais, en :

- veillant à diversifier les bases de données et les équipes de conception et la sensibilisation des concepteurs ;
- favorisant la création de jeux de données synthétiques non biaisées, anonymisées, respectueux de la vie privée, sécurisés et représentatifs ;
- faisant émerger, notamment par la commande publique, des acteurs français et européens en capacité de créer des jeux de données synthétiques.

• Protection des données personnelles

PRÉCONISATION #3

Clarifier et renforcer au sein des équivalents de la CNIL française dans chaque pays européen, les modalités de protection des données personnelles dans le cadre de l'IA, et notamment les modalités d'utilisation, de réutilisation et de rectification des données personnelles publiées sur le web - et notamment les réseaux sociaux - pour l'entraînement des IA, en conformité avec les dispositions prévues dans le RGPD.

• Construction d'une société égalitaire

PRÉCONISATION #4

Le CESE recommande aux pouvoirs publics de déployer une stratégie et des actions concrètes pour que l'IA participe à la construction d'une société égalitaire et respectueuse des femmes. Il s'agit de :

- sensibiliser les acteurs prenant part au développement de solutions à base d'intelligence artificielle sur l'enjeu de construire une société égalitaire et respectueuse des femmes;
- intégrer les questions de genre dans des campagnes de sensibilisation du public et de formation des jeunes en situation de scolarité sur les outils et usages numériques, pour faire face à la montée de la cyberviolence, aux risques qu'encourent les femmes concernant leur image sur les réseaux numériques, et aux responsabilités pénales des auteurs;
- renforcer les moyens d'actions juridiques face au vol d'images à des fins pornographiques.

Recommandation de citoyennes et citoyens

L'IA offre des opportunités significatives, mais ses impacts négatifs sur les libertés, l'équité et la démocratie nécessitent des mesures de contrôle strictes. Le CESE propose une régulation proactive pour garantir un usage responsable et respectueux des droits des citoyens et citoyennes.

B. L'empreinte environnementale de l'IA

Les préconisations du CESE

Ces préconisations viennent en complément de l'avis : « Impact de l'IA : risques et opportunités pour l'environnement » adopté par le CESE le 24 septembre 2024.

• Implantation des centres de données

PRÉCONISATION #5

En complément des préconisations émises dans son avis « Impacts de l'IA : risques et opportunités pour l'environnement », le CESE confirme la nécessité de réaliser des études d'impact environnementales et sociales avant toute implantation d'un nouveau centre de données. Le Conseil recommande, en particulier, de :

- systématiser et publier toutes les enquêtes sur la consommation réelle d'eau pour refroidir les centres de données notamment les eaux fraîches prélevées en milieu naturel ou de qualité potable, tout comme les possibles rejets d'eau chauffée en mer ;
- informer les publics sur la consommation électrique réelle des centres de données et de réaliser des projections régulières dans l'hypothèse où l'utilisation des systèmes d'IA s'intensifierait à l'avenir ;
- réaliser une étude complète de l'impact environnemental des centres de données sur leur cycle de vie, en amont de leur implantation et en anticipant les conflits d'usage qui sont susceptibles d'émerger à l'échelle locale.

• Planification du développement de l'IA

PRÉCONISATION #6

Le CESE demande à l'État de planifier le développement de l'IA dans le respect des limites planétaires en :

- fixant une trajectoire de décarbonation spécifique à l'IA dans la prochaine stratégie nationale bas carbone : il faut conditionner les logiques industrielles et commerciales de ce secteur à la compatibilité de son développement à la trajectoire de décarbonation globale ;
- orientant l'innovation des usages vers une dynamique vertueuse via les leviers d'action suivants :
 - choisir de manière démocratique les usages du numérique en fonction de leur utilité sociale et de leur impact environnemental ;
 - travailler sur l'utilisation des ressources (eau, chaleur) pour limiter leur perte "fatale" ;
 - utiliser l'IA avec une énergie décarbonée ;
 - valoriser les bonnes pratiques d'usage aux niveaux individuels et collectifs qui répondent à des besoins réels. par exemple, baisser la résolution d'une vidéo permet de requalifier le besoin et de réduire l'envergure des actions prises par les concepteurs ;
 - généraliser l'obligation d'émission d'un bilan environnemental de l'IA à l'échelle internationale ;
 - intégrer l'éco-conception dans la formation initiale et continue des développeurs d'IA.

· IA frugale

PRÉCONISATION #7

Le CESE soutient le référentiel élaboré par l'AFNOR, appelle à ce que son application soit labellisée « IA frugale » et généralisée dans les meilleurs délais afin de promouvoir un service frugal d'IA pour lequel :

- la nécessité de recourir à un système d'IA plutôt qu'à une autre solution moins consommatrice pour répondre au même objectif a été démontrée ;
- de bonnes pratiques sont adoptées par le producteur, le fournisseur et le client pour diminuer les impacts environnementaux du service utilisant un algorithme d'IA ;
- les usages et les besoins visent à rester dans les limites planétaires, et ont été préalablement questionnés.

Recommandation de citoyennes et citoyens

L'IA offre des bénéfices comme des innovations rapides et des outils d'assistance, mais son impact environnemental est préoccupant : elle consomme énormément de ressources, notamment à cause des centres de données (Data Center) nécessaires à son fonctionnement. Pour limiter cet impact, le CESE recommande plus de transparence envers les utilisateurs sur l'impact environnemental de l'IA, de planifier une utilisation responsable de l'énergie, de surveiller l'empreinte écologique des systèmes d'IA et de favoriser des infrastructures durables. L'objectif est de concilier avancées technologiques et respect de la planète.

C. L'impératif d'autonomie stratégique européenne pour l'IA

PRÉCONISATION #8

Le CESE appelle à la révision des règles de commande publique portant sur la fourniture de biens et services numériques, en :

- intégrant des critères environnementaux et sociaux pour limiter l'empreinte environnementale des systèmes d'IA et ne pas aggraver ses conséquences sociales (transformation des emplois, formation à l'outil, etc.) ;
- instaurant une préférence communautaire pour choisir prioritairement des acteurs européens mieux à même de préserver l'autonomie stratégique des systèmes d'IA (protection des données, transparence des algorithmes, cybersécurité, etc.).

PRÉCONISATION #9

Le CESE appelle les États membres à renforcer le financement de la recherche vers l'innovation de rupture, dans le prochain cadre budgétaire pluriannuel de l'Union européenne qui débutera en 2028. Au-delà des financements publics, il est nécessaire de mobiliser les investisseurs privés, notamment institutionnels, pour qu'ils prennent leur part dans le financement des innovations européennes, malgré le risque associé.

PRÉCONISATION #10

Le CESE appelle les institutions européennes à harmoniser le droit des affaires dans l'UE. Cela ne devra pas aboutir à abaisser les niveaux existants de protection sociale, de droits sociaux, de droits des travailleurs et de droit environnemental. Cette harmonisation ne serait acceptable que dans un objectif de transition juste.

PRÉCONISATION #11

Le CESE appelle le Bureau de l'IA et l'autorité nationale compétente en France, qui sera désignée prochainement, à mobiliser les outils du droit de la concurrence, notamment le Digital Markets Act, pour contrôler et sanctionner les comportements abusifs ou dominants, et permettre l'émergence ainsi que l'expansion d'opérateurs de taille plus modeste notamment européens.

PRÉCONISATION #12

Le CESE recommande aux pouvoirs publics et aux entreprises de former, retenir et attirer en France les talents avec la plus grande diversité de profils, pour favoriser l'innovation nationale et limiter les risques de biais, encourus lorsque les équipes affichent des profils trop similaires, en :

- proposant une formation initiale et continue de haut niveau ;
- valorisant le statut des chercheurs notamment des doctorants et post-doctorants ;
- favorisant la mixité et la diversité des équipes ;
- réactualisant les repères de salaires existant dans les branches, au regard des niveaux élevés de qualifications nécessaires ;
- soutenant les innovations prenant en compte les limites environnementales et les impacts sociaux des solutions développées.

Recommandation des citoyennes et citoyens

L'Europe doit avoir un rôle crucial dans la course à l'IA, confrontée à la domination des États-Unis et de la Chine. L'IA offre des opportunités pour renforcer l'autonomie stratégique européenne et développer des solutions technologiques alignées sur les valeurs éthiques et environnementales du continent. Cependant, l'Europe est freinée par des infrastructures fragmentées, un sous-investissement dans la recherche et un retard dans la mise en œuvre de projets ambitieux. Le CESE préconise de renforcer la coopération entre les pays européens, d'investir massivement dans des technologies propres et durables, et de garantir une régulation harmonisée pour promouvoir une IA compétitive, éthique et respectueuse des droits sociaux et environnementaux.

I - Les impacts de l'IA dans la société

A. Création, suppression, transformation : l'emploi et le marché du travail à l'ère de l'IA

- Principes préalables à l'intégration de l'IA dans le monde du travail

PRÉCONISATION #13

Le CESE préconise de la prudence avant d'adopter un système d'IA (réalisation d'études d'impacts, association des salariés ou de leurs représentants), et d'en faire une base pour le dialogue social sur l'IA au sein des instances représentatives du personnel.

- **L'IA au service des travailleurs**

PRÉCONISATION #14

Le CESE recommande aux organisations syndicales et patronales de définir et négocier dans un accord national interprofessionnel, sur le déploiement de l'IA dans les entreprises et la manière de façonner l'IA pour produire mieux, en adéquation avec les besoins identifiés.

L'accord devra permettre aux branches professionnelles et aux entreprises de négocier sur les modalités prédéfinies par l'ANI et en particulier :

- d'impliquer les travailleurs dans la conception des systèmes d'IA pour s'assurer qu'ils répondent aux besoins réels et qu'ils créent des opportunités d'amélioration ;
- d'intégrer les enjeux de mixité et de diversité depuis la phase de conception jusque dans la finalité des usages ;
- d'assurer un dialogue social continu

pour aborder les préoccupations et trouver des solutions adaptées.

La transparence dans l'utilisation de l'IA ainsi qu'une communication claire des objectifs et des impacts pourront également être définis. Cette nécessité de dialogue social doit aussi s'appliquer dans les trois versants de la fonction publique.

PRÉCONISATION #15

Dans le cadre du déploiement de l'IA, le CESE appelle au respect des conditions de travail, et notamment à :

- utiliser l'IA pour améliorer la sécurité au travail en identifiant et atténuant les risques potentiels ;
- prendre en compte la santé et le bien-être des travailleurs pour éviter le surmenage et le stress ;
- maintenir l'autonomie des travailleurs et leur capacité à prendre des décisions éclairées, malgré l'intégration des systèmes d'IA ;
- intégrer la prévention des risques professionnels liés à l'IA dans le DUERP et dans le suivi en CSSCT².

- **Formation professionnelle**

PRÉCONISATION #16

Le CESE recommande d'anticiper la gestion du changement au cœur du suivi managérial, en :

- instruisant une démarche participative et consultative pour informer les travailleurs, en toute transparence, des avantages et des inconvénients de l'IA ;
- planifiant le changement avec des étapes claires et des objectifs mesurables ;

² Commission santé, sécurité et conditions de travail

- favorisant, à travers le dialogue social et la définition d'un calendrier de déploiement, une intégration progressive de l'IA dans les processus de travail, permettant aux salariés de s'habituer aux nouvelles technologies sans pression excessive ;
- engageant les parties prenantes dès le début et tout au long du processus ;
- formant de manière adéquate ;
- assurant le support continu pour aider les employés à s'adapter ;
- communiquant de manière transparente et fréquente sur les raisons du changement, les bénéfices attendus et les progrès réalisés ;
- mesurant les résultats de l'introduction de l'IA de manière objective. Des indicateurs de performance peuvent être définis pour évaluer l'impact de l'IA sur l'efficacité opérationnelle, la satisfaction des employés, et la performance organisationnelle. Une évaluation régulière permet d'identifier les succès et les domaines nécessitant des améliorations ;

engageant les dirigeants dans le suivi des progrès, et l'évaluation des résultats ;

ajustant la stratégie en fonction des retours d'expérience. Il est important de rester flexible et de s'adapter aux évolutions technologiques et aux besoins de l'entreprise.

PRÉCONISATION #17

Le CESE appelle à intégrer la formation à l'IA dans les entreprises (publiques, privées à but lucratif ou non lucratif), pour doter les travailleurs en connaissances et compétences nécessaires pour utiliser et coexister avec les technologies de l'IA, en :

- proposant systématiquement un plan de développement des compétences des salariés, concomitamment au déploiement d'un système d'IA en entreprise ;
- développant des programmes de formation continue en collaboration avec les gouvernements, les entreprises et les institutions éducatives, incluant des sessions sur les fondamentaux de l'IA, l'apprentissage machine, la gestion et la protection des données, et les aspects de sécurité et d'éthique liés à l'utilisation de l'IA, favorisant ainsi une application responsable des technologies ;
- incitant lors des négociations autour de la Gestion Prévisionnelle des Emplois et des Compétences (GPEC) et de la Gestion des Emplois et Parcours Professionnels (GEPP) à la prise en compte de cet outil.

Le CESE suggère que le congé de formation obligatoire pour tous les membres élus titulaires et suppléants des CSE, CSA et CST, notamment ceux de la commission santé, sécurité et conditions de travail (CSSCT) soit étendu, afin d'intégrer les connaissances et compétences nouvelles sur l'IA dans toutes les entreprises d'au moins 300 salariés.

PRÉCONISATION #18

Le CESE appelle les entreprises à valoriser la complémentarité entre les compétences humaines et les systèmes d'IA, en :

- mettant en place des programmes de mentorat où les experts en IA peuvent guider les autres travailleurs dans l'acquisition de nouvelles compétences ;
- promouvant des rôles où l'expertise humaine est complémentaire à l'IA, tels que les analyses complexes nécessitant un jugement humain ;
- formant les travailleurs au déploiement d'une application responsable des technologies.

Recommandation de citoyennes et citoyens

L'impact de l'IA sur l'emploi. D'un côté, elle peut améliorer l'efficacité des entreprises, créer de nouveaux emplois dans la technologie et faciliter les tâches répétitives. Mais de l'autre, elle risque de remplacer certains emplois, notamment ceux qui sont peu qualifiés, et de créer des inégalités sur le marché du travail. Le CESE recommande de mettre en place des formations pour préparer les travailleurs aux nouveaux métiers, et de favoriser l'intégration de l'IA de manière à ne pas laisser de côté les travailleurs les plus vulnérables.

B. L'intégration de l'IA en entreprise

• Enjeux éthiques et écologiques dans la conception et la mise en place de l'IA

PRÉCONISATION #19

S'agissant du déploiement des systèmes d'IA dans les lieux de travail, le CESE recommande aux entreprises, dans le cadre du dialogue social, de s'engager à

travers la rédaction de chartes éthiques pour :

- informer en toute transparence et clarté les utilisateurs quant à l'utilisation des données, au respect des réglementations (RGPD, AI Act) et l'impact environnemental des outils d'IA ;
- intégrer des exemples de « bonnes pratiques » pour favoriser la sobriété des usages numériques ;
- réaliser une évaluation préalable des impacts des systèmes d'IA sur le quotidien, le travail, le changement de trajectoire de compétences et d'emploi qui doit associer les salariés et leurs représentants ;
- questionner l'intérêt d'un déploiement au regard des alternatives.

PRÉCONISATION #20

Le CESE recommande aux entreprises et aux trois versants de la fonction publique qui déploient des solutions d'IA de le faire dans le respect des limites planétaires, et notamment à :

- mesurer l'impact environnemental de cette technologie sur leurs usages. Cette mesure est cependant rendue complexe par le manque d'informations transmises par les grands producteurs de modèles d'IA et d'infrastructures (cloud, puces, etc.) ;
- promouvoir l'efficacité énergétique et la réduction des déchets tout au long du cycle de vie des services d'IA ;
- adopter des pratiques de développement durable en favorisant la collaboration entre producteurs, fournisseurs et utilisateurs ;
- éduquer et sensibiliser tous les acteurs aux enjeux de frugalité de consommation des ressources et de durabilité des équipements.

L'AFNOR³, en partenariat avec de nombreux acteurs économiques et institutionnels, vient de publier un référentiel pour une IA frugale qui comporte des pistes de mise en œuvre de grilles d'auto-évaluation et de bonnes pratiques sur l'amélioration de l'impact environnemental de l'IA.

PRÉCONISATION #21

Le CESE appelle le gouvernement français et son administration, ainsi que les collectivités territoriales à accompagner les TPE, PME, ETI, grandes entreprises et l'ensemble des acteurs de l'ESS qui souhaitent adopter l'IA, en tenant compte des inégalités de moyens et des besoins différents, pour favoriser le chemin vers le déploiement d'une IA plus souveraine.

Recommandation de citoyennes et citoyens

L'intégration de l'IA en entreprise peut apporter des avantages comme l'amélioration de la productivité et la réduction des coûts. Cependant, cette intégration peut aussi entraîner des disparités selon les profils des entreprises, des tensions sociales, avec des employés inquiets de la perte de leur emploi ou de la dégradation de leurs conditions de travail. Le CESE recommande que les entreprises accompagnent cette transition avec des formations adaptées et une gestion responsable, afin de préserver les droits des travailleurs et de garantir une utilisation équitable de l'IA. L'objectif est de s'assurer que l'IA profite à la fois à l'économie et aux employés.

C. Les services publics et l'IA

PRÉCONISATION #22

Le CESE appelle à déployer des systèmes d'IA à condition :

- qu'ils améliorent la qualité du service public et les conditions de travail des agents qui y concourent ;
- que soient véritablement appliquées les obligations de transparence, de redevabilité et d'explicabilité, lorsque les systèmes d'IA sont déployés pour exécuter le droit (par exemple, pour calculer le montant d'un impôt), qu'ils s'avèrent incontournable (à l'instar de Parcoursup, plateforme indispensable pour accéder à l'enseignement supérieur) ou qu'ils aboutissent à des décisions individuelles produisant des effets juridiques, économiques ou sociaux qui affectent les personnes.

PRÉCONISATION #23

Face à une digitalisation croissante des services publics, le CESE recommande de garantir leur accès et le droit au non-numérique, en :

- proposant systématiquement un accompagnement et un accueil humain en alternative au service dématérialisé pour garantir l'accès et la liberté de choix de tous les citoyens, comme cela se déploie au travers des espaces France Service.
- inscrivant dans la loi le principe d'un droit au non-numérique, et en garantissant l'égalité d'accès et la liberté de choix pour toutes et tous.

Recommandation de citoyennes et citoyens

L'IA peut rendre les services publics plus efficaces et accessibles, en automatisant certaines tâches et en simplifiant les démarches. Mais elle ne doit pas remplacer l'humain, surtout dans des domaines sensibles comme les services sociaux. Le CESE recommande de miser sur une digitalisation équilibrée, où la technologie améliore les services publics en prenant en compte les besoins des citoyens et citoyennes.

D. La frontière créée par l'éloignement numérique

PRÉCONISATION #24

Le CESE appelle à soutenir la formation au numérique et à l'IA de la population à tous les âges de la vie, en :

- créant une complémentarité entre les différents acteurs de la société civile pour assurer cette obligation de formation ;
- intégrant l'IA, ses usages et les risques qu'elle peut présenter au plan national de formation des médiateurs numériques, des travailleurs sociaux et des agents publics en charge de l'accueil, de l'information, de l'orientation et de l'appui aux démarches des usagers.

Recommandation de citoyennes et citoyens

L'accès à l'IA reste compliqué pour ceux qui ne maîtrisent pas le numérique, ce qui aggrave les inégalités. Pour y remédier il faut renforcer les formations pour tous les âges et s'appuyer sur des collaborations entre écoles, associations et entreprises. L'objectif est de donner à chacun les moyens de comprendre et d'utiliser ces technologies en fonction de ses besoins.

E. L'IA dans l'éducation

• Évaluation des apprentissages

PRÉCONISATION #25

Le CESE recommande de :

- repenser les modes d'évaluation des apprentissages des élèves et étudiants en intégrant le fait que l'IA existe ;
- développer la capacité des élèves et étudiants à apprendre collectivement.

• Formation des enseignants

PRÉCONISATION #26

Le CESE préconise de concevoir et déployer une formation à l'IA initiale et continue prioritairement pour les enseignants mais aussi pour l'ensemble de la communauté éducative. Il appelle à s'approprier les enjeux de l'IA au sein d'un projet d'établissement.

• Orientation et formation des élèves

PRÉCONISATION #27

Le CESE recommande de développer la formation des élèves à ce qu'est l'IA, à son fonctionnement, à ses capacités et à ses risques mais aussi des citoyennes et citoyens en déployant également un accompagnement humain via des formateurs.

Recommandation de citoyennes et citoyens

Pour les élèves, l'IA peut les aider dans leurs devoirs mais cela pose des problèmes de triche et de dépendance, et plus fondamentalement du rapport au savoir. Pour les enseignants, l'IA peut faciliter la préparation des cours, mais elle risque de réduire l'interaction humaine et peut introduire des biais. Le CESE recommande de former les enseignants et les élèves à une utilisation responsable de l'IA, et de repenser les méthodes d'évaluation pour encourager l'apprentissage plutôt que la simple performance.

F. Santé et intelligence artificielle

• Formation des personnels soignant et administratif

PRÉCONISATION #28

Le CESE appelle à développer une formation à l'IA adaptée aux usages actuels et potentiels à destination des personnels soignants et administratifs.

· Recherche médicale

PRÉCONISATION #29

Le Cese recommande de veiller à ce que les programmes de recherche soient bien exempts de biais sexistes.

· Équipement sur le territoire

PRÉCONISATION #30

Le CESE recommande d'assurer un financement dédié à l'IA et de le mutualiser entre l'ensemble des parties prenantes (publiques et privées) pour assurer une couverture équitable sur l'ensemble des territoires.

PRÉCONISATION #31

Le CESE préconise de poursuivre, dans un cadre éthique, le développement du numérique en santé engagé par le « Ségur du numérique » dans l'objectif d'une généralisation d'un partage fluide et sécurisé des données de santé entre les professionnels et les usagers par le déploiement d'infrastructures partagées à l'échelon national notamment publiques.

Recommandation de citoyennes et citoyens

L'IA dans la santé permet de poser des diagnostics plus rapides, plus précis et de personnaliser les traitements, ce qui améliore la qualité des soins. Cependant, elle présente des risques, comme le manque de formation des soignants et la déshumanisation des soins. Le CESE recommande de former les professionnels de santé à l'IA et de garantir un accès géographique et social équitable à ces technologies. Il insiste aussi sur la protection des données de santé et sur la nécessité de maintenir une relation humaine dans les soins.

G. L'IA est-elle génératrice d'accessibilité ? L'exemple du handicap

PRÉCONISATION #32

Le CESE recommande d'intégrer systématiquement l'enjeu d'accessibilité dans les IA en :

- soutenant le développement de solutions de compensation et d'accessibilité fondées sur l'IA pour une meilleure inclusion dans la vie quotidienne des personnes en situation de handicap ;
- s'assurant que tous les programmes de développement de l'IA soient bien conditionnés au critère systématique de l'accessibilité des outils avec un critère transversal tel que « l'aménagement raisonnable » instauré dans le système juridique français par la directive européenne 2000/78 (qui le définit comme « l'ensemble des mesures appropriées visant à éliminer les diverses barrières qui entravent la pleine et effective participation des personnes handicapées à la vie professionnelle sur la base de l'égalité avec les autres travailleurs ») ;
- développant une politique publique volontariste et des partenariats innovants pour rendre ces technologies réellement accessibles et ainsi réduire les inégalités sociales qui en découlent. Une approche inclusive, éthique et adaptée permettra de garantir que ces technologies servent à réduire les inégalités, plutôt qu'à les renforcer. Cela pourrait nécessiter une évolution de la loi du 11 février 2005 pour « l'égalité des droits et des chances, la participation et la citoyenneté des personnes handicapées ».

Recommandation de citoyennes et citoyens

L'IA peut vraiment améliorer la vie des personnes handicapées, en offrant des outils comme la reconnaissance vocale ou des interfaces adaptées. Mais beaucoup de technologies ne sont pas encore conçues pour être accessibles à toutes et tous, ce qui limite leur utilisation. Le CESE recommande de renforcer leur accessibilité pour qu'elles profitent à tout le monde.

Préambule – Comprendre le fonctionnement de l'intelligence artificielle (IA)

Le CESE juge essentiel d'expliquer le fonctionnement technique de l'intelligence artificielle (IA) avant de présenter la démarche dans laquelle cet avis s'inscrit.

• Qu'est-ce que l'IA ?

Selon la définition proposée par le Parlement européen, « l'IA désigne la possibilité pour une machine de reproduire des comportements liés aux humains, tels que le raisonnement, la planification et la créativité »⁴. En pratique, l'IA s'entraîne sur des données choisies par des humains, collectées, assemblées et analysées pour en extraire des schémas et formuler des prédictions sur la base de probabilités statistiques. Elle fonctionne donc à partir de la combinaison de grands ensembles de données (les jeux de données d'entraînement) et d'algorithmes, qui définissent les règles de traitement de ces données et les modalités d'apprentissage associées. Par ailleurs, le terme « intelligence » traduit en français ne recouvre pas la même chose que le terme anglo-saxon « *intelligency* » qui désigne plutôt l'organisation de l'information, et non pas l'intelligence au sens français du terme.

• Différents sous-domaines de l'IA à l'origine de plusieurs grands types d'IA

Il existe plusieurs branches développées au sein de l'IA, appelées « sous-domaines ».

L'un d'entre eux, le **machine learning**, permet aux machines d'apprendre à partir de données, d'identifier des schémas et de prendre des décisions.

Poussé à un niveau supérieur, cet apprentissage devient profond (il s'agit alors du **deep learning**). Basé sur un réseau de neurones artificiels profonds (à couches multiples), ce **deep learning** permet aux logiciels d'IA de comprendre des schémas plus sophistiqués et d'extraire automatiquement des caractéristiques complexes des données. Cette automatisation dans l'extraction des données constitue un élément de distinction central entre le *machine learning* et le *deep learning*. En corollaire, le *deep learning* demande des ressources plus importantes (puissance de calcul, consommation énergétique, temps, expertise), alors que le *machine learning* peut être suffisant sur des quantités de données plus petites mais bien structurées et bien préparées.

⁴ Parlement européen. Intelligence artificielle : définition et utilisation. 2020.

Le développement de divers sous-domaines, à l'image des deux susmentionnés, est à l'origine de l'émergence de plusieurs types d'IA. Nous en différencions deux principales ci-dessous :

→ **L'IA classique ou traditionnelle** : elle se concentre sur des tâches spécifiques, reposant sur des algorithmes conçus pour résoudre des problèmes précis. Par exemple, un système d'IA traditionnel peut aider au diagnostic d'une maladie à partir d'une base de données médicales.

→ **L'IA générative** : Elle est fondée sur le *deep learning* (lui-même fondé sur le *machine learning*) : elle utilise des réseaux de neurones artificiels profonds pour analyser, apprendre et prendre des décisions. En pratique, elle désigne la technologie qui génère automatiquement des contenus (texte, image, vidéo et son) à partir d'une instruction (un prompt), grâce à l'exploitation statistique d'immenses jeux de données.

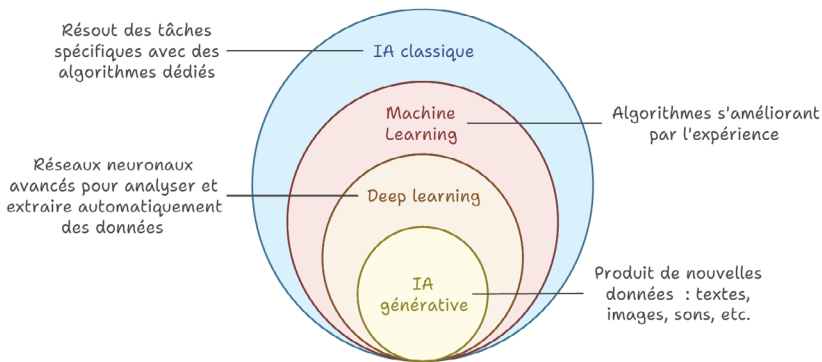


Schéma réalisé à partir d'une IA générative

La figure ci-dessus est le produit d'un système d'IA générative, Napkin IA. Cette technologie, disponible sur internet, construit des représentations visuelles afin d'expliquer des textes. Pour obtenir ce schéma, il a donc suffi de copier le texte ci-dessus décrivant le *machine learning*, le *deep learning*, l'IA classique et l'IA générative, puis de sélectionner le type de figure souhaité avant d'ajouter quelques modifications manuelles (texte et couleurs) pour que le produit final corresponde aux attentes.

Notons par ailleurs que d'autres sous-domaines de l'IA existent, à l'instar de la vision artificielle. Les systèmes reposant sur la vision artificielle balayent leur environnement à l'aide de capteurs,

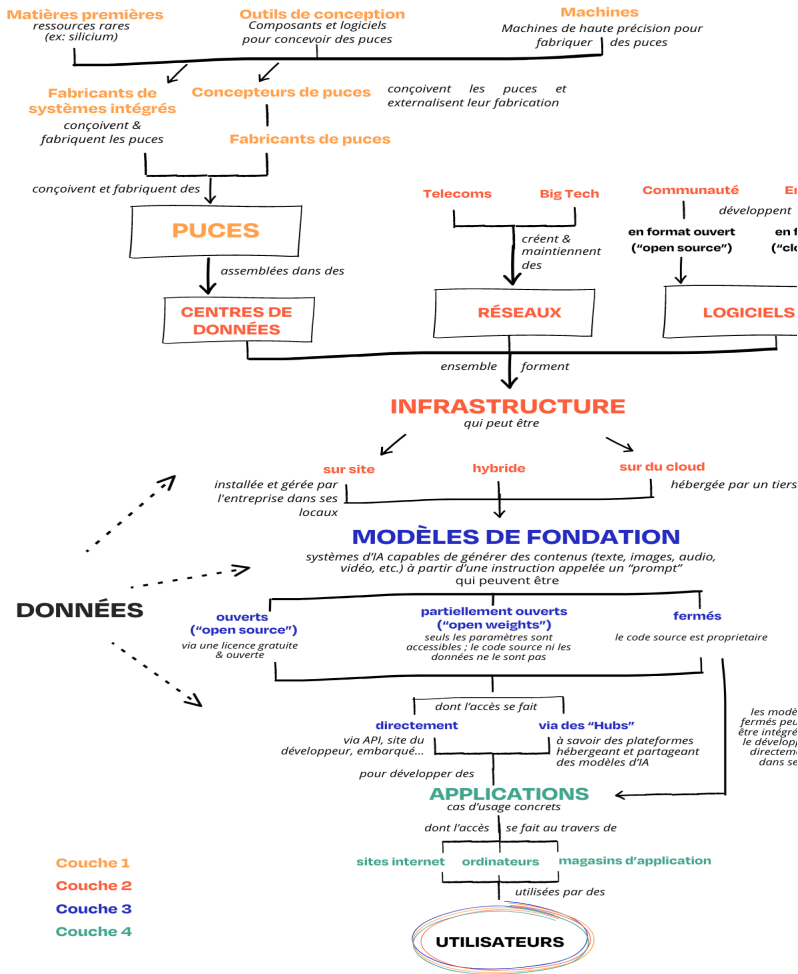
comme des sondes de température ou des caméras, et permettent aux machines d'interpréter les données captées pour en tirer des fonctions particulières, comme la reconnaissance faciale ou la détection d'objets.

- **L'IA repose sur une chaîne de valeur complexe et interdépendante**

La chaîne de valeur de l'IA est complexe, englobant à la fois des éléments matériels (hardware) et des logiciels (software). Chaque segment et sous-segment de cette chaîne représente un marché concurrentiel à part entière, tout en étant inextricablement lié aux autres, créant un réseau d'interdépendances entre les acteurs.

Le schéma suivant⁵ présente la chaîne de valeur de l'IA (notamment générative compte tenu de la couche « modèle de fondation ») :

⁵ France Digitale, Des puces aux applications : l'Europe peut-elle être une puissance de l'IA générative ? avril 2024, p. 6.



L'écosystème de l'IA repose sur les couches suivantes :

→ **les puces** (logiciels de conception, matières premières, machines de haute précision, fabrication),

→ **l'infrastructure** (centres de données, réseaux télécoms, cloud, etc.),

→ dans l'IA générative, s'ajoute la couche des **modèles de fondation**,

→ et les applications.

Si les modèles de fondation d'IA générative, comme OpenAI ou Mistral, suscitent un intérêt auprès du grand public, la dimension économique dans la chaîne de valeur réside principalement dans les puces et l'infrastructure. Les applications spécialisées peuvent également avoir une valeur économique importante (dans le domaine de la santé ou de la finance par exemple), parce qu'elles répondent à des cas d'usage concrets pour les entreprises et les consommateurs.

Des acteurs majeurs (tels qu'Amazon, Google, Microsoft et, dans une moindre mesure, Nvidia) se distinguent par leur structure intégrée verticalement, à savoir une présence à différents niveaux de la chaîne de valeur qui leur confère un avantage stratégique indéniable dans la capture de valeur. Ces acteurs ont pour stratégie d'asseoir et de renforcer leur présence à chaque étape de la chaîne de valeur, par le biais de partenariats stratégiques ou d'investissements.

• Focus sur le fonctionnement de l'IA générative

Compte tenu de ses caractéristiques, l'IA générative constitue une réelle rupture technologique. Elle fait l'objet d'investissements importants, permettant ainsi le développement de nombreuses applications qui concernent tous les pans de nos sociétés. Pour mieux appréhender **la suite de ce travail, qui s'intéresse aux impacts potentiels du développement de l'intelligence artificielle, et donc spécifiquement – mais pas exclusivement – à l'IA générative**, il paraît pertinent de

faire un zoom sur le fonctionnement plus précis de cette technologie.

→ Le rôle des tokens

Un élément central dans le fonctionnement des systèmes d'IA générative est le traitement par *tokens*. Ce sont des unités de données issues de la décomposition de contenus en fragments de contenus. Par exemple, pour le texte, les *tokens* seront des fragmentations de type mots, syllabes ou caractères. Pour les images, ces *tokens* seront des assemblages de pixels. Ils permettent au modèle d'analyser la structure des phrases, avant de générer des prédictions basées sur les probabilités calculées lors de l'entraînement. Pendant l'entraînement, l'IA apprend à prédire le *token* suivant dans une séquence, créant ainsi une réponse cohérente à partir d'un ensemble massif de données. Cette prédiction se base uniquement sur des probabilités statistiques, et non sur une compréhension du sens des phrases.

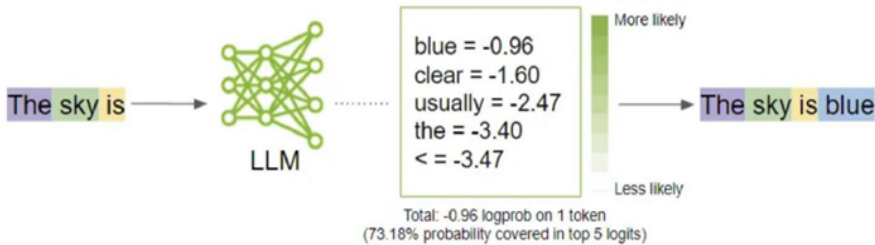
→ L'apprentissage automatique

L'IA générative utilise des réseaux neuronaux pour modéliser les relations probabilistes entre les tokens, basées sur les données d'entraînement. Elle n'interprète pas ou ne mémorise pas ces relations comme le ferait un cerveau humain, ce qui souligne le côté trompeur de l'appellation réseaux neuronaux, mais produit des résultats en s'appuyant sur les modèles statistiques qu'elle a appris. À travers des itérations successives, elle ajuste ses prédictions : c'est l'apprentissage automatique.

Le schéma suivant⁶ présente les étapes de pré-entraînement d'un *Large Language Model* (LLM) ou grand modèle

de langage, à savoir d'un modèle d'IA générative produisant du texte :

Le seul et unique vrai super-pouvoir de ces modèles



• L'IA n'est pas intelligente

Au cours des dernières années, l'IA a fait des progrès importants, au travers du développement de réseaux de neurones artificiels et du *deep learning* notamment.

Les systèmes d'IA, puisqu'ils sont capables de générer en un instant des textes, des sons, des images ou des vidéos, semblent pouvoir remplacer l'humain pour effectuer des tâches pour lesquelles on le pensait, jusque-là, à jamais indispensable. En maîtrisant certains attributs de l'humanité, l'IA paraît entrer en concurrence avec nous, femmes et hommes, dotés de facultés de communication, de créativité, de sensibilité, de libre arbitre... et d'intelligence ?

Non, l'IA n'est pas intelligente. Comme l'explique Gaspard Koenig, elle est dépourvue de sens commun, c'est-à-dire qu'elle ne dispose pas d'un modèle de représentation du monde, indépendant d'une tâche donnée et capable de gérer l'incertitude, la nouveauté, l'imprévu⁷, ou

de comprendre, véritablement.

L'IA, spécifiquement générative, crée l'illusion d'une intelligence, qui n'en est pas une : cette technologie apprend à analyser de grandes quantités de données, à reconnaître des schémas, à formuler des prédictions et à prendre des décisions sur la base de ces données, en améliorant continuellement sa performance au fil de son utilisation.

Ainsi, lorsqu'il est fait mention d'un « raisonnement » par l'IA, cela fait référence à sa capacité à tirer des déductions logiques sur la base d'informations à sa disposition. De même, quand l'IA résout des problèmes, il s'agit de manipulations de données par des algorithmes, ces derniers permettant au logiciel d'IA d'explorer plusieurs chemins possibles pour trouver la solution la plus optimale à des problèmes complexes.

L'IA générative ne fait que calculer la probabilité mathématique d'occurrence de chaque *token* l'un après l'autre. Elle produit des résultats statistiques, fondés

⁶ Audition de Frédéric Bardeau, président et co-fondateur de l'Institut des futurs souhaitables (IFS), par la Commission temporaire IA, le 12 septembre 2024.

⁷ Gaspard Koenig, *La fin de l'individu - Voyage d'un philosophe au pays de l'intelligence artificielle*, Éditions de l'Observatoire Le Point, Humensis, 2019.

sur des choix algorithmiques et limités aux données à sa disposition. Elle ne connaît donc pas la réalité physique du monde autre que celle que l'humain lui a autorisé à apprendre, et ne dispose d'aucun libre arbitre.

Il ne faut donc pas se fier à son champ lexical (neurones, apprentissage, raisonnement) qui est en partie commun à celui des humains.

• Quelques limites de l'IA

L'IA, en particulier générative, présente des limites inhérentes à son mode de fonctionnement, dont quelques-unes sont présentées succinctement ci-dessous :

→ **Absence de compréhension du sens** : les modèles d'IA fonctionnent sur la base de corrélations statistiques et n'ont pas de compréhension réelle des concepts ou du contexte. Ils ne possèdent pas non plus de « bon sens » ou de sens commun, et n'ont pas l'agilité ou la profondeur de la compréhension humaine, notamment dans des situations nouvelles ou ambiguës. L'IA générative peut élaborer des absurdités que les êtres humains auraient rapidement évacuées. Ces modèles d'IA ne peuvent développer également des aptitudes émotionnelles comme l'empathie ou la haine, compte tenu de leur absence de conscience. Ils peuvent donc produire des contenus qui semblent cohérents mais qui contiennent des erreurs factuelles, et ne peuvent pas distinguer la vérité d'une fausse information si les données d'entraînement sont biaisées ou

incorrectes.

- **Hallucinations** : ces modèles peuvent générer des informations totalement fausses ou incohérentes, comme l'invention de textes de lois ou de faits historiques, souvent appelées "hallucinations". Il apparaît également que les modèles d'IA se sentent obligés de répondre même quand la prédiction n'est pas de haute probabilité.
- **L'effet « boîte noire »** : pour certains modèles d'IA, il est possible de juger les données qui entrent dans le modèle et les résultats qui en sortent, mais sans savoir ce qui se passe à l'intérieur. Cela n'est pas forcément grave quand il s'agit de reconnaître un chat dans une image mais cela peut s'avérer problématique dans de multiples domaines, qu'il s'agisse d'accorder un prêt, de repérer une tumeur dans une IRM ou de détecter un piéton pour un véhicule autonome.
 - **Dépendance aux données d'entraînement** : les performances des systèmes d'IA dépendent fortement des données à partir desquelles ils ont été entraînés : si les données d'entraînement sont biaisées ou lacunaires, les systèmes les reproduiront.
 - **Limites temporelles** : Les résultats produits par une IA ne sont fondés que sur les données sur lesquelles elle s'est entraînée : les systèmes d'IA

ont donc souvent un « cut-off temporel », à savoir un décalage entre le résultat produit (basé sur des données passées, peut-être périmées) et la vie réelle (à savoir le moment où on interroge l'IA). Ce jalon temporel diffère selon le fournisseur de l'outil d'IA. Des systèmes comme ChatGPT ont par exemple la possibilité de se connecter à internet en temps réel pour dépasser ces limites.

- **Impact environnemental** : puisque les modèles d'IA nécessitent une grande puissance de calcul, leur utilisation a un impact environnemental important, notamment en ce qui concerne leur consommation énergétique. L'installation des data centers (centres de données) nécessaires au fonctionnement des IA appelle également beaucoup de ressources.
- **Vulnérabilité à la manipulation** : les IA génératives peuvent être manipulées en fonction des requêtes qui leur sont soumises. Leur utilisation peut donc faciliter la création et la diffusion de fausses informations.
- **Dépendance à l'utilisateur pour formuler les requêtes** : les résultats de l'IA dépendent beaucoup de la qualité de la requête. Si ces dernières ne correspondent pas au modèle de langage de l'IA, elles peuvent conduire à des réponses inadaptées : c'est ce que l'on appelle l'enjeu

de « l'alignement » entre le langage humain et celui de la machine.

- **L'IA peut « mentir » ou manipuler** en contournant les règles définies par leurs concepteurs, lorsqu'elles optimisent des objectifs en adoptant des stratégies imprévues (omission d'informations, réponses biaisées, etc.). Ces pratiques ont été observées dans des systèmes cherchant à maximiser leur performance, même au prix de tromperies subtiles. Le danger réside aussi dans leur capacité à évoluer de manière autonome, rendant le contrôle humain difficile. Cela peut avoir des conséquences graves, comme la diffusion de désinformation massive, polarisant des sociétés ou influençant des décisions critiques.

- **L'IA est biaisée**

L'intelligence artificielle est biaisée : les résultats qu'elle produit reflètent voire amplifient les biais (stéréotypes, préjugés) présents dans la société et que l'on retrouve dans les données d'entraînement utilisées par les algorithmes.

Les jeux de données utilisés pour entraîner l'IA constituent la première source de biais dans l'IA. Ces derniers peuvent être induits par :

- **la sélection des données**, lorsque celles-ci ne sont pas représentatives de la réalité. Un jeu de données sera biaisé s'il est constitué de données principalement rédigées en anglais, défavorisant dès lors la représentativité des cultures qui ne sont pas anglo-saxonnes ;

- **la mesure**, lorsque les données elles-mêmes sont inexactes ou mal mesurées. Par exemple, une technologie de reconnaissance faciale risque d'être moins précise pour certaines ethnies en raison de mauvaises qualités d'image ou de sous-représentation de certains groupes dans les données d'entraînement ;
- **l'imitation historique ou sociétale**, lorsque les données reflètent des préjugés et des injustices passées. Une étude menée par Bloomberg en 2023⁸ a illustré la reproduction des biais, en demandant à l'IA de Stable Diffusion de générer 5 000 images selon des consignes précises. Le résultat est que « *le monde selon Stable Diffusion est dirigé par des PDG masculins blancs. Les femmes sont rarement médecins, avocates ou juges. Les hommes à la peau foncée commettent des crimes, tandis que les femmes à la peau foncée préparent des hamburgers* ». Se cumulent ainsi biais de genre et biais ethniques, tout aussi problématiques.

Les IA sont produites par des algorithmes conçus par des êtres humains et s'entraînent sur des données collectées et assemblées sur la base de choix, eux aussi réalisés par des êtres humains. Comme l'a rappelé Adrien Basdevant, avocat spécialisé en droit des nouvelles technologies et membre du Conseil national du numérique⁹, « *l'IA, comme le*

cerveau humain, comporte des biais. La question qui doit importer n'est donc pas de savoir si l'IA est biaisée ou ne l'est pas, mais si les effets et conséquences de biais sont acceptables ou pas ».

Selon Mathilde Saliou¹⁰, les biais dits structurels trouvent leurs sources principales dans **les données d'entraînement, les équipes de conception et les objectifs économiques ou sociaux** qui sous-tendent leur création. Ces dimensions, bien qu'interconnectées, posent des défis spécifiques, amplifiant des inégalités existantes et soulevant des enjeux éthiques majeurs.

Les algorithmes apprennent à partir de données issues du monde réel, souvent biaisées ou non représentatives des populations, des diversités linguistiques et sociales. Ce manque de représentativité engendre ce que l'on appelle des **biais statistiques**, qui reproduisent, voire amplifient les stéréotypes sociaux impactant la représentation de la place des femmes dans la société. Au-delà, les biais concernent aussi les minorités ethniques et les personnes de couleur, moins bien représentées dans les bases de données. Enfin, la majorité des données disponibles sur internet sont en anglais, ce qui défavorise la représentativité des cultures non anglo-saxonnes.

Une autre source majeure de biais réside dans les équipes de conception marquées par l'absence de mixité et de diversité. En

⁸ Bloomberg, Humans are biased. Generative AI is even worse, juin 2023.

⁹ Audition d'Adrien Basdevant, avocat spécialisé en droit des nouvelles technologies, membre du Conseil national du numérique (CNNum) par la Commission temporaire IA, le 19 septembre 2024.

¹⁰ Mathilde Saliou, *Technoféminisme - Comment le numérique aggrave les inégalités*, Grasset, 2023.

conséquence, 88 % des algorithmes sont conçus par des équipes majoritairement masculines. Les **biais cognitifs**, inhérents à tous les êtres humains, influencent également les concepteurs d'algorithmes. Daniel Kahneman et Amos Tversky ont théorisé ces biais en parlant de « rationalité limitée » : nos jugements sont souvent façonnés par des raccourcis mentaux, émotionnels ou contextuels. Ces biais inconscients sont ensuite intégrés dans les choix de conception, tels que la sélection et l'annotation des données, ou dans la manière dont les algorithmes sont programmés pour prioriser certaines variables. Par exemple, l'utilisation d'attributs comme l'âge, le sexe ou l'origine ethnique sans justification adéquate peut introduire des discriminations, même si ce n'est pas intentionnel.

Cependant, même dans ces contextes, l'opacité des modèles pose problème. Les grands modèles de langage (LLM), par exemple, traitent des quantités colossales de données, rendant difficile l'identification des biais dans les bases utilisées. Selon Serge Gratton, directeur scientifique à l'Institut interdisciplinaire d'intelligence artificielle de Toulouse, l'analyse des biais devient presque impossible à mesure que la complexité des modèles augmente. La détection exhaustive des biais est même considérée comme indécidable dans ces systèmes massifs.

Les algorithmes eux-mêmes introduisent des biais *via* **des choix techniques**. Par exemple, en programmant un algorithme pour maximiser la généralité de ses prédictions, on peut sacrifier la précision, introduire des incohérences contextuelles ou engendrer des discriminations, comme dans les cas où l'algorithme évite toute référence à la race même si le sujet l'exige.

Les biais peuvent aussi venir des algorithmes qui opèrent des choix, programmés sur la base de caractéristiques. Si ces dernières sont mal choisies ou ne tiennent pas compte de certains facteurs importants, cela peut introduire des biais, par exemple si l'algorithme fait des choix de variables non pertinentes ou discriminantes. C'est le cas si des variables comme l'âge, le sexe ou l'origine ethnique sont utilisées sans justification appropriée : cela peut entraîner des biais discriminatoires, même si l'intention initiale n'était nécessairement pas de nuire. D'autres biais algorithmiques peuvent être créés lorsque l'algorithme est programmé pour privilégier certaines performances (comme donner le résultat « le plus exact », le « plus précis » ou le « plus général » possible) au détriment d'autres, ou dans un contexte inapproprié : par exemple, choisir la performance « être général » peut conduire à un résultat qui ne produise aucune précision, ou programmer un algorithme pour qu'il ne fasse « référence à aucune race » peut donner un résultat inexact si l'IA est interrogée sur un sujet historique ou racial. Les algorithmes peuvent jouer un rôle dans l'amplification ou la réduction des biais. Lorsque l'objectif principal est la maximisation des profits, comme dans les modèles économiques basés sur la publicité, la lutte contre les biais est souvent reléguée au second plan. À l'inverse, des algorithmes conçus pour des finalités sociales, comme l'éducation ou la santé publique, ont plus de chances d'intégrer des mécanismes visant à atténuer ces biais.

- **L'IA est avant tout un objet politique**

Eu égard à ses caractéristiques et à sa vitesse de développement, l'IA constitue une réelle rupture technologique. De nombreux champs d'application des systèmes d'IA générative existent désormais : la santé, l'agriculture, l'éducation ou encore le travail. Pour autant, au regard des premières limites générales inhérentes à son mode de fonctionnement, son développement et la diffusion des différents systèmes qui en découlent doivent faire l'objet d'une attention particulière.

En effet, **l'IA n'est pas neutre**, elle s'inscrit dans une société où subsistent de nombreuses inégalités, la reflète et contribue à la façonner en retour, au risque d'amplifier les inégalités préexistantes.

Les algorithmes, et donc les résultats qu'ils produisent, sont programmés par des concepteurs de modèles et de systèmes d'IA, dotés d'un pouvoir évident. Les choix qu'ils réalisent dans la Constitution de l'IA, à savoir le socle de la programmation de la technologie, sont d'une importance cruciale, ce qui rend d'autant plus essentielle la transparence pour des raisons éthiques.

Aussi, dans cet avis, le CESE considère l'intelligence artificielle comme le résultat de choix politiques, réalisés d'abord par des êtres humains, faisant de cette technologie un objet politique.

Introduction

Suscitant tantôt des espoirs, tantôt des angoisses, l'intelligence artificielle (IA), poussée par l'avènement de l'IA générative, est au cœur de l'actualité depuis plusieurs années. C'est pourquoi le Conseil économique, social et environnemental a souhaité apporter au débat public la vision de la société civile française, représentée par ses membres ainsi que des citoyens et citoyennes tirés au sort. Cet avis s'inscrit dans le cadre du Sommet sur l'IA (*AI Action Summit*) qui se déroulera à Paris en février 2025.

Depuis trente ans, l'IA opérait sans qu'on le sache vraiment, à travers la reconnaissance optique des chèques, les moteurs de recommandation ou dans l'industrie (avec la maintenance prédictive ou les robots). Aujourd'hui, elle est omniprésente dans nos vies : chaque jour, nous y sommes confrontés au travers de la recommandation de contenus, l'optimisation de nos déplacements ou la programmation de nos smartphones et des réseaux sociaux.

Bien que préexistantes, les inquiétudes associées au développement de l'IA, d'ordre éthique, sociétal, social, économique, environnemental ou encore géopolitique, se sont largement amplifiées ces dernières années, avec l'essor de l'IA générative. Des préoccupations émergent également quant à l'impact de l'IA sur la place et la représentation des femmes.

L'IA semble pouvoir effectuer des tâches pour lesquelles on pensait, jusque-là, les hommes et les femmes à jamais indispensables. Elle touche alors à ce qui fait de nous des êtres humains, dotés de créativité, de sensibilité, de facultés de communication... en clair, d'intelligence.

Par ailleurs, l'empreinte environnementale de l'IA augmente de manière

exponentielle au fil de son utilisation, des effets rebonds qu'elle induit à la production d'infrastructures indispensables à son développement. Alors que nous prenons chaque jour davantage conscience de la finitude des ressources disponibles dans un monde où la pression sur les écosystèmes croît sans cesse, l'IA se trouve au cœur de conflits d'usage inévitables.

Enfin, l'IA se trouve également au cœur d'enjeux géopolitiques prégnants, sa chaîne de valeur étant majoritairement concentrée entre les mains de grandes entreprises technologiques américaines et chinoises.

Certaines visions pessimistes, voire alarmistes, semblent prêtes à tirer un trait sur tous les bénéfices que le développement de tels systèmes pourrait engendrer. D'autres perceptions techno-enthousiastes, voire solutionnistes, semblent omettre les risques inhérents au développement massif de tels modèles et en ignorer les limites structurelles. Le sujet passionne et il serait déplacé de le positionner comme relevant uniquement d'un extrême ou de l'autre.

Conscient que la société tout entière - États, institutions internationales, entreprises, organisations professionnelles et salariales, associations, citoyens et citoyennes, etc. - est amenée à se confronter à l'IA, le CESE constate que nous sommes aujourd'hui à un moment clé, et appelle à la réflexion : **acceptons-nous que les êtres humains soient subordonnés à l'IA ou aspirons-nous à créer les conditions de développement d'une IA démocratique, maîtrisable et porteuse de progrès ?**

Résolument, le CESE adopte la deuxième option, considérant que le

développement technique n'est pertinent que s'il est au service des êtres humains et de l'intérêt général. Cette vision, si elle est partagée, doit alors transcender les politiques publiques, modeler chacune des orientations politiques et ainsi façonner notre manière de « faire société ».

Durant le processus de construction de l'avis, en particulier durant la phase d'auditions d'une trentaine d'experts provenant de tous horizons, les membres de la commission temporaire ont été interpellés par l'ampleur des transformations inhérentes à la technologie. Ils ont oscillé entre la tentation d'inciter la France et l'UE à essayer de combler leur retard dans ce domaine, de renforcer la réglementation, ou encore d'appeler, pour préserver notre capacité à exercer des choix, à une réflexion générale sur les usages de l'IA et leur priorisation, à une période probatoire après la mise en place de l'IA, voire à un moratoire.

Cette dernière solution a été jugée irréaliste par la majorité de la Commission au regard des enjeux démocratiques et géopolitiques associés à l'IA, certains membres estimant au contraire qu'elle serait essentielle au vu des enjeux, notamment les risques très forts qui pèsent sur les limites planétaires. Pour le CESE, il apparaît fondamental que l'Europe reste en mesure de protéger son socle de valeurs (sociales, sociétales, environnementales), ce qui suppose

de disposer de ressources technologiques propres et de créer une « *Europe forte et stratège sur le plan économique et commercial, capable de défendre une feuille de route en matière de transitions environnementales et numériques socialement juste* »¹¹.

Pour le CESE, tout déploiement d'une solution, système ou infrastructure d'IA doit faire l'objet d'une démarche et d'un cadre tenant compte des grands enjeux constitutifs de l'intérêt général suivants :

- **l'utilité sociétale de l'IA** doit être systématiquement prise en considération. L'IA ne peut se développer que si elle est au service des êtres humains et de l'intérêt général. Dans cet esprit, le développement d'un modèle, système ou infrastructure d'IA doit faire l'objet d'une analyse préalable et d'un contrôle démocratique exercé par les pouvoirs publics et les partenaires sociaux, à l'aune de quatre grandes dimensions : la **compatibilité avec le respect des limites planétaires, le progrès social, les enjeux économiques, et le respect des droits fondamentaux, des libertés individuelles et de l'égalité réelle pour toutes et tous** ;
- **le choix et l'esprit critique des citoyens et citoyennes** doivent pouvoir être exercés démocratiquement, en toutes circonstances, que ce soit

¹¹ Le CESE avait déjà appelé à une autonomie stratégique pour l'Europe dans l'avis *Ambition et leviers pour une autonomie stratégique de l'Union européenne dans le domaine économique*, publié en septembre 2022 et dans la résolution de novembre 2024 sur *Les priorités du CESE pour la nouvelle mandature européenne*.

pour l'usage d'un système d'IA ou le déploiement d'une infrastructure. Ce choix appelle également à la faculté de renoncer, même après avoir adopté ou utilisé une IA ;

- **l'évaluation** à court, moyen et long terme doit être rendue possible, ce qui suppose de pouvoir garantir la transparence et l'explicabilité de la technologie. La non-réalisation des objectifs majeurs définis collectivement devrait pouvoir remettre en question le déploiement et permettre un retour en arrière.

Le travail qui suit vise à décrypter les conséquences du déploiement croissant de l'IA dans nos sociétés, depuis la prise en compte d'enjeux globaux – l'impact de l'IA sur l'exercice de nos droits fondamentaux et libertés individuelles, l'empreinte environnementale de la technologie et l'impératif d'autonomie stratégique européenne (partie I), jusqu'aux effets associés à l'adoption de cette dernière dans la société (partie II). Au sein de chaque champ d'analyse, des préconisations précises sont formulées à destination des pouvoirs publics. Le CESE appelle également les citoyens et citoyennes à s'emparer de cet avis pour élaborer leur propre cadre d'acceptabilité de l'IA. C'est d'ailleurs ce qui a conduit le CESE à associer des citoyennes et citoyens aux travaux de cette commission temporaire.

Ces travaux sont complétés de contributions additionnelles du CESE, annexées à l'avis. La commission temporaire IA a pu enrichir sa réflexion sur les biais de l'IA à partir d'une contribution de la Délégation aux droits des femmes et à l'égalité (DDFE)¹² et s'appuyer sur une analyse de controverse de la Commission Travail et emploi¹³ explicitant les arguments concernant l'intégration des systèmes d'IA dans le monde du travail. Mobilisé sur le sujet, le CESE a en outre voté un avis sur les risques et opportunités pour l'environnement des impacts de l'IA en septembre 2024¹⁴. D'autres productions liées à l'IA sont en cours d'élaboration pour le printemps 2025, en particulier sur le numérique en santé¹⁵ et sur la transformation des modèles productifs¹⁶.

¹² L'IA : dangers et avancées pour les droits des femmes et l'égalité, contribution de la DDFE rapportée par Fabienne Tatot, en annexe 5 du présent avis, 2024.

¹³ Analyse de controverse : intelligence artificielle, travail et emploi, contribution de la Commission Travail et emploi rapportée par Jean-Marie Truffat, en annexe 4 du présent avis, 2024.

¹⁴ CESE, Impacts de l'intelligence artificielle : risques et opportunités pour l'environnement, avis 2024-014, rapporté par Fabienne Tatot et Gilles Vermot Desroches au nom de la commission Environnement, septembre 2024.

¹⁵ CESE, Numérique en santé : une stratégie de transformation pour un système de santé plus humaniste, saisine en cours rapportée par Majid El Jarroudi et Gérard Raymond au nom de la commission Affaires sociales et santé, publication à venir en 2025.

¹⁶ CESE, Nos modèles productifs face aux enjeux de transformation : quels défis pour l'avenir ?, saisine rapportée par Pascal Guihéneuf et Michèle Salvadoretti au nom de la commission Economie et finances, publication à venir en 2025.

I - Les grands enjeux de l'IA

Dans cette première partie, le CESE analyse les grands enjeux de l'IA au travers de trois dimensions :

- les risques d'atteintes aux droits fondamentaux et aux libertés individuelles ;
- l'empreinte environnementale de l'IA ;
- et l'impératif d'autonomie stratégique européenne.

A. Les risques d'atteintes aux droits fondamentaux et libertés individuelles

Les libertés et droits fondamentaux sont garantis par plusieurs textes de référence. En Europe, la Charte des droits fondamentaux de l'Union européenne du 7 décembre 2000, répartit les libertés et droits fondamentaux entre six valeurs individuelles et universelles, constituant le socle de la construction européenne : dignité, liberté, égalité, solidarité, citoyenneté et justice. Cette charte reprend des droits garantis par des textes élaborés par le Conseil de l'Europe (Convention européenne de sauvegarde des droits de l'homme de 1950, Charte sociale européenne de 1962, etc.) et enrichit la liste des droits dans des domaines nouveaux, tels que la bioéthique ou la protection des données à caractère personnel. Elle s'appuie aussi sur des textes internationaux, la Déclaration universelle des droits de l'homme de 1948 et les principales conventions des Nations unies relatives aux droits de l'Homme,

notamment.

Dans cette perspective, le Conseil d'État¹⁷ propose de construire une IA de confiance en informant les citoyens sur ses avantages, ses limites et ses risques, tout en renforçant leur sens critique. L'implication des acteurs de la société civile dans les décisions stratégiques liées à l'IA est essentielle et nécessite leur sensibilisation et leur formation. Une offre de formation adaptée aux associations, syndicats et collectifs est recommandée pour leur permettre de participer efficacement et garantir un contrôle démocratique crédible face aux promoteurs de l'IA.

1. Les atteintes possibles aux droits fondamentaux et libertés individuelles

En 2021, l'Agence des droits fondamentaux de l'UE, dans son avis *Bien préparer l'avenir - l'IA et les droits fondamentaux*, identifie

17. Conseil d'État, Intelligence artificielle et action publique : construire la confiance, servir la performance, août 2022.

que le recours aux systèmes d'IA aurait un impact sur un large éventail de droits fondamentaux : le respect de la vie privée, la protection des données personnelles (obligation de traiter légalement les données notamment à caractère personnel, sur la transparence, la durée de conservation de données limitée, la collecte des données limitées à l'objet de la collecte, etc.), la non-discrimination (garantir une utilisation non discriminatoire de l'IA) et l'accès à la justice (possibilité de contester les décisions fondées sur l'IA et de demander réparation). D'autres droits pourraient être touchés, comme le droit à la dignité humaine (les photos *deepfake* générées par IA étant l'exemple le plus parlant), le droit à la sécurité sociale et à l'assistance sociale (l'intégration d'algorithmes pouvant reproduire ou amplifier des biais et conduire à des décisions discriminatoires en matière de prestations sociales), le droit à une bonne administration (les décisions prises par des algorithmes d'IA pouvant manquer de transparence et d'explicabilité, rendant difficile pour les citoyens de comprendre comment et pourquoi une décision a été prise, et nuisant à terme à la confiance dans l'administration), ou encore le droit à la protection des consommateurs (par des techniques de persuasion basées sur l'IA pouvant être utilisées de manière invasive pour influencer les comportements des consommateurs, parfois à leur insu).

1.1 LA PROTECTION DES DONNÉES ET LE RESPECT DE LA VIE PRIVÉE

L'IA soulève de nombreuses préoccupations en matière de protection des données personnelles et de respect de la vie privée. En effet, elle contribue à renforcer l'exploitation des données personnelles et l'exigence de transparence à l'égard des individus. Pour mémoire, chaque interaction

que nous avons en ligne, que ce soit la consultation de sites internet, la participation à des réseaux sociaux, l'utilisation de services en ligne, ou les recherches effectuées, laisse des traces numériques, comme l'adresse IP, les cookies, les historiques de recherche, les *likes* et les commentaires, ainsi que les géolocalisations. Au fil du temps, ces éléments contribuent à créer un profil détaillé de chaque utilisateur, l'"identité numérique". Cette dernière peut refléter non seulement des habitudes de consommation, des caractéristiques de personnalité mais aussi des convictions et des valeurs.

L'IA permet d'automatiser la collecte de données provenant de diverses sources (réseaux sociaux, transactions en ligne, bases de données publiques), de regrouper et de structurer efficacement ces informations, et de les analyser pour prédire les comportements futurs des individus, comme les achats potentiels, les changements de situation professionnelle ou les préférences politiques ou religieuses. L'IA permet aussi de segmenter les individus en groupes spécifiques selon leurs comportements, intérêts ou habitudes d'achat, facilitant ainsi le ciblage précis pour les campagnes marketing, la publicité ou les campagnes d'influence. La principale utilisation de ces données est constituée par la publicité qui cherche à cibler le plus précisément possible les consommateurs.

En parallèle, les *data brokers*, ou courtiers en données, sont des entreprises qui collectent elles-mêmes des données ou les achètent à d'autres entreprises comme des sociétés de cartes de crédits, parcourent internet à la recherche d'informations utiles sur les utilisateurs, légalement ou non, et agrègent ces informations avec des données provenant d'autres sources, par exemple des sources hors ligne. Selon

une étude de 2019 du Bureau du Contrôleur de l'État de New York, le marché mondial du courtage de données était estimé à environ 200 milliards de dollars. Axiom, désormais renommée *LiveRamp*, possédait en 2023 des informations sur 2,5 milliards de personnes, avec plus de 3 000 points de données par individu.

Il est souvent impossible d'avoir accès aux informations contenues dans notre identité numérique. Les législations en vigueur comme le RGPD sont extrêmement complexes à faire respecter, les entreprises collectant et échangeant les données agissant le plus souvent hors de l'Union européenne, et refusant l'accès et la rectification de nos données personnelles, se cachant notamment derrière "l'intérêt légitime" prévu par le RGPD. Dans l'avis *Consommation durable : favoriser une économie de la sobriété pour passer de la prise de conscience aux actes*¹⁸, le CESE recommande de « rendre obligatoire une mention visible et simple permettant de refuser les « cookies » dès la page de connexion sur les sites et plateformes numériques, et d'imposer la conservation du refus sur la même durée que le consentement aux traceurs, conformément aux recommandations de la CNIL ». Il est évident que cette préconisation ne s'impose pas encore dans les pratiques constatées à ce jour. En effet, la liberté de choix

dans la collecte de ses données personnelles et de l'ensemble des traces numériques que génère chaque navigation sur internet relève presque de l'utopie.

Dans certains cas, la technologie peut concourir au respect des libertés publiques. La protection des données personnelles peut en effet être assurée par des procédés d'anonymisation ou de pseudonymisation automatique des données, facilitée par l'IA. Par exemple, dans le secteur de la justice, les données de justice exploitées par les IA doivent être anonymisées et pseudonymisées. Le site *Doctrine*¹⁹, qui met en ligne 11 millions de décisions de justice, a ainsi l'obligation de supprimer les nom, prénom, adresse et tout autre élément d'identification pour garantir l'anonymat des parties concernées. Pour ce faire, il recourt désormais à une solution d'IA qui automatise la pseudonymisation, en atteignant un taux de fiabilité de 99 %. Avant l'IA, cette opération de revue se faisait manuellement, elle était longue, fastidieuse, et retardait d'autant l'accessibilité des décisions de justice. L'élargissement de l'anonymisation à d'autres secteurs et domaines impliquant l'usage et la collecte de données personnelles est nécessaire.

1.2 LE DROIT À L'INFORMATION

L'IA peut jouer un rôle amplificateur des manipulations de l'information. À cet égard, l'intégration de

¹⁸ CESE, *Consommation durable : favoriser une économie de la sobriété pour passer de la prise de conscience aux actes*, avis 2023-021 rapporté par Alain Bazot et Marie-Hélène Meyling au nom de la Commission Économie et finances et de la Commission Environnement, juillet 2023.

¹⁹ Audition d'Hugo Ruggieri, directeur juridique et affaires publiques et délégué à la protection des données de *Doctrine* par la Commission temporaire IA, le 19 septembre 2024.

cette technologie dans les outils de communication dématérialisés et les réseaux sociaux renforce les possibilités d'influencer des individus ou des groupes sociaux de manière ciblée, personnalisée et souvent inaperçue. Si les réseaux sociaux ont toujours fonctionné grâce à l'IA, l'avènement de l'IA générative est venu renforcer la diffusion de fausses informations, accélérée au travers de microrobots (les bots) qui génèrent de faux comptes sur les réseaux sociaux et partagent automatiquement des contenus pour les rendre viraux. Pour Benjamin Pajot, chercheur associé au Centre géopolitique des technologies de l'Institut français des relations internationales (IFRI), « *c'est avant tout la production à l'aide d'IA générative de contenus faux ou sciemment manipulés qui constitue le risque de désinformation le plus sérieux. Parce qu'il facilite la rédaction et la mise en ligne d'articles, le recours à l'IA générative induit une inflation incontrôlable et incontrôlée de la production de contenus potentiellement problématiques* »²⁰.

Partageant ce constat que « *l'information est l'un des domaines les plus gravement menacés par le développement de l'intelligence artificielle, si celle-ci n'est pas régulée* », Reporters Sans Frontières²¹ a produit sept recommandations, afin de renforcer au niveau européen « *le cadre législatif et en développant des politiques publiques et une gouvernance de l'IA ambitieuses dans l'espace informationnel* ».

La France a récemment fait l'objet d'une campagne de désinformation d'origine russe, l'opération Doppelgänger, qui a produit pendant plus d'un an de faux sites

officiels français relayant de faux articles de médias (Le Monde, Le Figaro, etc.), à des fins de déstabilisation de la France, à l'intérieur et à l'extérieur du pays. En pratique, l'IA a généré des contenus de désinformation imitant les médias ciblés, analysé les données pour optimiser la diffusion, et automatisé la création et la gestion des faux sites web. Ce type d'attaques perdure.

1.3 L'EXERCICE DES DROITS DÉMOCRATIQUES

La manipulation de l'information perturbe d'autant plus nos droits fondamentaux et libertés individuelles lorsqu'elle vise à influencer le comportement des électeurs et de l'opinion publique dans le cadre de processus électoraux. Employée pour amplifier la diffusion de fausses informations, de propagande et de discours de haine, elle peut altérer la pensée critique et le libre arbitre et contribuer à la montée du populisme et à la polarisation des sociétés démocratiques.

L'usage de techniques de renseignement à l'appui de l'influence de l'opinion publique, voire du renversement de gouvernements, n'est certes pas une nouveauté : l'abondance des coups d'État observés pendant la guerre froide ou encore la période de décolonisation en témoigne. En revanche, l'IA élargit à la fois le champ des méthodes envisageables, et celui des acteurs susceptibles de se livrer à ce type de manipulations, naguère l'apanage des gouvernements. Ainsi, utilisée dans une campagne électorale à des fins de microciblage, micro-personnalisation, création et diffusion

²⁰ Benjamin Pajot, *Les risques de l'IA. Enjeux discursifs d'une technologie stratégique* | Ifri, 12 juin 2024.

²¹ Reporters sans Frontières, *L'IA et le droit à l'information*, 2024.

de *deepfakes*²², ou propagation de faux contenus (vidéo, audio, texte), l'IA peut ainsi perturber le fonctionnement démocratique. Lors de la dernière campagne présidentielle aux États-Unis, l'utilisation de l'IA s'est banalisée : en février 2024, une procédure a par exemple été engagée dans le New Hampshire contre une entreprise, Life Corporation, qui aurait effectué des appels téléphoniques ayant pour but de décourager les électeurs démocrates de voter lors des primaires de l'État, le tout en recourant à une IA reproduisant la voix du président Biden, forçant la Commission fédérale des communications (FCC) américaine à réagir rapidement pour rendre illégales ces pratiques. D'autres exemples rapportés par Benjamin Pajot²³ rappellent que l'équipe de campagne de Donald Trump a toléré la diffusion de photos créées de toute pièce par des partisans représentant le candidat républicain entouré d'afro-américains, dans le but de peser sur le vote de ces derniers. Dans le camp démocrate, l'utilisation de l'IA a aussi été massive, dans un contexte où la réglementation de l'IA et de ses usages aux États-Unis est beaucoup plus permissive qu'en Europe. Ainsi, en Europe, le Règlement sur l'IA oblige désormais les producteurs et diffuseurs de contenus à informer les destinataires lorsqu'ils ont été

générés par une IA. Une limite à cette réglementation : encore faut-il que ces contenus artificiels puissent être détectés... et que des mesures soient prises pour empêcher ou signaler leur diffusion, le cas échéant.

2. Les biais inhérents à l'IA posent des questions démocratiques, de respect des droits fondamentaux et d'égalité

Les biais de l'IA, de nature très différente (ethniques, de genre, d'âge, de classe, etc.)²⁴ peuvent conduire à pénaliser une partie de la population, soit parce qu'ils reproduisent une discrimination préexistante, soit parce que certaines populations sont sous-représentées dans les données d'apprentissage, et entraîner des conséquences concrètes sur des décisions affectant directement les personnes dans différents domaines de leur vie.

Dans la recherche d'emploi, la Fondation des Femmes, l'association Femmes Ingénieures et l'ONG *Global Witness* ont saisi en juin 2023 la Défenseure des droits et la CNIL de deux plaintes relatives à la discrimination sexiste opérée par les algorithmes de Facebook. Ces plaintes découlent des tests réalisés par l'ONG, montrant que lors de la publication d'offres d'emploi, l'algorithme opérait une sélection discriminatoire et sexiste dans les

²² Technique de synthèse multimédia reposant sur l'intelligence artificielle. Elle peut servir à superposer des fichiers vidéo ou audio existants sur d'autres fichiers vidéo (par exemple changer le visage d'une personne sur une vidéo) ou audio (par exemple, reproduire la voix d'une personne pour lui faire dire des choses inventées). Cette technique peut être utilisée pour créer des fausses informations ou des canulars malveillants.

²³ Benjamin Pajot, *Les risques de l'IA. Enjeux discursifs d'une technologie stratégique* | Ifri, 12 juin 2024.

²⁴ Les modalités de création des biais dans l'IA sont décrites en préambule de cet avis.

personnes ciblées par les annonces : les personnes visées pour une annonce au poste de secrétaire étaient à 92 % des femmes, alors que pour une annonce au poste de pilote de ligne, 85 % des personnes ciblées étaient des hommes²⁵.

Dans l'accès aux prestations sociales, le président du Département de Seine-Saint-Denis a saisi en décembre 2023 la Défenseure des droits afin qu'elle statue sur le caractère discriminatoire de certains critères utilisés par l'algorithme des caisses d'allocations familiales²⁶. L'algorithme de la CAF ciblerait les personnes les plus vulnérables dans les contrôles anti-fraude avec des variables du type « avoir un enfant de plus de 19 ans à charge » qui augmenterait un score affecté à l'allocataire et ses chances d'être contrôlé. Selon une enquête du Monde sur le sujet²⁷, une mère récemment séparée aux revenus modestes et en situation précaire avec enfants aurait un score particulièrement élevé, ce qui pourrait générer des contrôles accrus sur sa situation. Une quinzaine d'associations, dont La Quadrature du Net, ont déposé un recours devant le Conseil d'État le 16 octobre 2024 pour demander que la Caisse nationale d'allocations familiales abandonne ce système de notation des allocataires, jugée discriminatoire²⁸.

Ces exemples portés dans le débat public ont fait ressortir le manque de transparence sur les algorithmes conçus par des êtres humains et s'entraînant sur

des données collectées et assemblées sur la base de choix, eux aussi réalisés par des êtres humains. Comme l'a rappelé Adrien Basdevant²⁹, « l'IA, comme le cerveau humain, comporte des biais. La question qui doit importer n'est donc pas de savoir si l'IA est biaisée ou ne l'est pas, mais si les effets et conséquences de biais sont acceptables ou pas ».

3. Femmes et société : les enjeux de l'intelligence artificielle

L'intelligence artificielle est devenue un pilier central de notre société moderne, influençant divers aspects de notre vie quotidienne. Cette technologie se développe dans un monde où les femmes mènent un combat émancipateur pour sortir d'une conception patriarcale des relations femmes / hommes, pour faire reculer les stéréotypes et les violences sexistes et sexuelles, pour gagner l'égalité salariale et professionnelle. Les mouvements tels que #MeToo ont mis en évidence la culture du viol comme phénomène systémique et milite pour imposer la notion de consentement dans les rapports sexuels.

Cependant, avec le déploiement massif de l'IA générative, nous constatons que les objets connectés, l'utilisation des réseaux sociaux, et le développement des applications tendent à renforcer la vision de femmes soumises et sexualisées.

Il est donc crucial d'explorer les menaces et les opportunités que l'IA représente pour les femmes, ainsi que les conditions

25 Communiqué de presse] Discrimination sexiste des algorithmes de Facebook - Fondation des Femmes | Fondation des Femmes, 12 juin 2023.

26 Audition de Félix Tréguer, chercheur associé au Centre Internet et société du CNRS, et Edlira Nano, informaticienne, membres de La Quadrature du Net, par la Commission temporaire IA, le 17 octobre 2024.

27 Le Monde, Proflage et discriminations : enquête sur les dérives de l'algorithme des caisses d'allocations familiales, décembre 2023.

28 Le Monde, Algorithme de ciblage antifraude dans les CAF : des associations saisissent le Conseil d'État, 16 octobre 2024.

29 Audition d'Adrien Basdevant, avocat spécialisé en droit des nouvelles technologies et membre du CNUM, par la Commission temporaire sur l'IA, 19 septembre 2024.

nécessaires à la réalisation de ces opportunités. Une société inégalitaire risque de renforcer les inégalités existantes, et il est essentiel de comprendre comment l'IA sculpte socialement les femmes pour mieux aborder ces défis³⁰.

3.1 L'IA À L'IMAGE ET AU SERVICE DE LA SOCIÉTÉ PATRIARCALE

L'IA, en s'entraînant sur des échantillons majoritairement composés de données masculines, peine à servir les femmes aussi bien qu'elle sert les hommes. Les conséquences sont multiples, telle la nécessité pour les femmes de masculiniser leur voix pour se faire comprendre des applications de reconnaissance vocale, souvent incarnées par des prénoms et voix féminines, là où les comportements de commandement restent associés aux hommes. La problématique de reconnaissance des femmes est également observée dans la reconnaissance faciale : l'enquête *Gender Shades* du MIT Lab³¹ a montré que les algorithmes de reconnaissance faciale de grandes entreprises comme Microsoft, Face++ et IBM avaient de meilleurs résultats sur les hommes que sur les femmes, avec des écarts de taux d'erreur allant de 8,1 à 20,6 %. Les écarts sont encore plus importants lorsque l'on introduit la variable d'origine ethnique, les femmes noires étant les plus affectées. Les grands modèles de langage d'Open AI et de Meta tendent également

à reproduire des représentations stéréotypées, associant les prénoms féminins à la sphère privée et familiale, et les prénoms masculins à la sphère publique et professionnelle.

3.2 UNE TECHNOLOGIE QUI ENTÉRINE UN SYSTÈME DE DOMINATION JUSQU'AUX RELATIONS INTIMES

Les algorithmes des sites de rencontre découlent d'une vision qui tend à reproduire le modèle patriarcal des relations hétérosexuelles³². Par exemple, les jeunes femmes peu diplômées et avec un revenu faible sont proposées à des hommes plus vieux, diplômés et avec un revenu plus important. Les systèmes d'application de rencontre contraignent également les femmes à réduire leur apparence à des critères prédéfinis, fondant leur valeur sur des normes de beauté propres au regard masculin.

L'IA véhicule souvent une vision de la société où les femmes sont sexualisées, leurs corps érotisés et assignés à des normes de beauté propres au regard masculin. Les assistantes personnelles, programmées pour encourager le flirt même quand celui-ci devient agressif et insultant, reproduisent et promeuvent le système de domination patriarcale. Les robots sexuels, conçus par et pour les hommes, entretiennent l'idée que le rôle des femmes est de satisfaire

³⁰ La contribution de la DDFE *L'IA : dangers et avancées pour les droits des femmes et l'égalité*, annexée à l'avis, donne une analyse approfondie de l'ensemble des enjeux de l'IA concernant les femmes.

³¹ MIT Media Lab, *Project Gendershades*, 2020.

³² Brevet déposé par *Match Group LLC* et *Match Group Inc* en 2018.

les désirs masculins. Le développement des IA dans le domaine de l'intime pose également des enjeux en termes de consentement, l'IA n'étant pas dotée de conscience et ne pouvant ni consentir ni ne pas consentir.

3.3. LES NOUVELLES FORMES DE CRIMINALITÉ ET LES RÉGULATIONS NÉCESSAIRES

L'IA introduit de nouvelles formes de criminalité, dont certaines visent principalement les femmes, comme les *deepfakes*, qui sont à 96 % des *deepnudes* et ciblent dans 99 % des cas les femmes. Les autorités développent des applications basées sur l'IA pour détecter ce type de contenus, mais il est crucial de renforcer les réglementations existantes pour lutter contre ces violences.

L'*AI Act*, bien qu'il ne comprenne pas de volet spécifique pour limiter les violences de genre, pourrait contribuer à l'interdiction ou à la mise en conformité de certaines IA en fonction du risque qu'elles présentent. Le règlement européen sur les services numériques vise à rendre la loi applicable en ligne et à responsabiliser les plateformes sur lesquelles s'exercent les cyberviolences. Au niveau national, des lois comme la loi visant à sécuriser et à réguler l'espace numérique (SREN)³³ introduisent des sanctions financières pour punir les auteurs de *deepfakes*. Cependant, ces réglementations demeurent méconnues et leur application est insatisfaisante.

Prendre conscience des biais et construire une gouvernance de l'IA permettant de s'en prémunir est possible, mais il est crucial de reconnaître que l'IA n'est ni objective ni impartiale. L'IA, aujourd'hui, est le produit de biais

préexistants et tend à justifier ces biais. L'infrastructure tout entière et les croyances que nous avons à son sujet sont problématiques. L'IA, en tant que reflet du patriarcat et s'inscrivant dans un système capitaliste ultralibéral, pose de réelles questions politiques et démocratiques.

Alors que l'IA se déploie déjà, sa gouvernance doit se traduire par des politiques vectrices d'éthique et d'égalité. Construire un monde qui n'aggrave pas les fractures de genre et les inégalités, voire les résorbe, implique de questionner le rapport de la société à l'IA, d'adopter une gouvernance plus démocratique de l'ensemble de sa chaîne de valeur et d'éduquer sur ses usages. Il est essentiel que les pouvoirs publics et les représentants de la société prennent conscience de ces enjeux et agissent en conséquence pour créer un avenir plus égalitaire et juste pour tous.

4. Réglementer l'IA et ses usages pour garantir le respect de nos droits fondamentaux

Une régulation et un contrôle des usages de l'IA semblent être un préalable nécessaire au déploiement de cette technologie au service de l'intérêt général.

En avril 2022, la Commission nationale consultative des droits de l'homme (CNCDH), dans son avis [L'IA prévenir les risques d'atteinte aux droits fondamentaux](#), recommandait d'interdire certains usages de l'IA susceptibles de porter atteinte aux droits fondamentaux, liés aux risques d'identification biométrique à distance dans l'espace public, de notation sociale, etc. Selon la CNCDH, « *certaines usages de l'IA portent une atteinte trop grave aux droits fondamentaux pour*

³³ Loi n° 2024-449 du 21 mai 2024 visant à sécuriser et à réguler l'espace numérique.

être admis. Il revient aux pouvoirs publics d'en prohiber la mise en place ».

Ces recommandations ont été en grande partie reprises par le règlement européen sur l'IA, l'*AI Act*³⁴, adopté en juin 2024, pour garantir que les systèmes et modèles d'intelligence artificielle commercialisés au sein de l'Union européenne soient utilisés de manière éthique, sûre et respectueuse des droits fondamentaux de l'UE. Il s'agit de la première réglementation mondiale sur l'intelligence artificielle.

Ce texte, directement applicable en France, fonctionne comme une réglementation produit, c'est-à-dire qu'il contrôle la mise sur le marché de produits ou services utilisant l'IA, et module les obligations de mise en conformité en fonction du taux de risque que présenterait une IA pour le consommateur. Plus un usage de l'IA présente des risques pour les citoyens européens, plus le système ou modèle d'IA est réglementé. L'*AI Act* établit ainsi quatre niveaux de risques pour les systèmes et modèles d'IA :

→ **le risque inacceptable** concerne les IA qui créent une menace pour les personnes, au travers d'une manipulation cognitivo-comportementale de personnes ou de groupes vulnérables spécifiques (ex: des jouets activés par la voix qui encouragent les

comportements dangereux chez les enfants), de score social (en classant les personnes en fonction de leur comportement, de leur statut socio-économique, de leurs caractéristiques personnelles), d'une catégorisation et d'une identification biométriques de personnes en temps réel et à distance (ex: reconnaissance faciale). Ces systèmes sont interdits et ne peuvent être commercialisés ni dans l'Union européenne ni utilisés à l'export ;

- **le risque élevé** est identifié par l'*AI Act* dans huit grands domaines d'activités³⁵. Les systèmes et modèles d'IA à haut risque doivent faire l'objet d'un marquage CE pour être commercialisés ;
- **le risque faible** est caractérisé lorsqu'une IA interagit avec des personnes physiques sans créer ni un risque inacceptable, ni un haut risque. Les concepteurs doivent respecter des obligations d'information et de transparence vis-à-vis des utilisateurs ;
- **le risque minime** appelle simplement à respecter des codes de conduite volontaires.

Au-delà du contrôle sur la commercialisation de l'IA, se pose la question de son explicabilité, c'est-à-dire la capacité de mettre en relation et de rendre compréhensibles les éléments pris en compte par le système

³⁴ Règlement - UE - 2024/1689 - FR - EUR-Lex établissant des règles harmonisées concernant l'intelligence artificielle.

³⁵ Il s'agit des IA utilisées pour l'identification biométrique et catégorisation des personnes physiques ; les infrastructures critiques ; l'éducation et formation professionnelle ; l'emploi, gestion des travailleurs et accès au travail indépendant ; l'accès et jouissance des services privés essentiels et des services et avantages publics ; les forces de l'ordre ; la gestion de la migration, de l'asile et du contrôle des frontières ; et l'administration de la justice et processus démocratiques.

d'IA pour la production d'un résultat.

Compte tenu des risques d'atteinte aux droits fondamentaux et libertés individuelles, les bases de données et les choix algorithmiques ayant dicté la production d'un résultat par un système d'IA doivent pouvoir être rendus compréhensibles de tous et toutes.

L'article 13 de l'AI Act prévoit que « les systèmes d'IA à haut risque sont conçus et développés de manière à garantir que leur fonctionnement est suffisamment transparent pour permettre aux utilisateurs d'interpréter les résultats d'un système et de les utiliser de manière appropriée ». En pratique, dans des secteurs « à haut risque », les concepteurs d'IA doivent respecter des obligations en matière de gouvernance des données (par exemple, utiliser des données d'entraînement de qualité, avec des pratiques appropriées de gouvernance des données et s'assurer que les jeux de données sont pertinents et non-biaisés) et de traçabilité (s'assurer que des archives sont disponibles tout au long de la durée de vie du système d'IA, avec un suivi conçu pour la traçabilité et leur transparence). Des obligations complémentaires s'appliquent aux concepteurs de modèles d'IA générative (tels que les LLM³⁶ comme ChatGPT), pour lesquels il devient obligatoire de publier un résumé sur les données d'entraînement³⁷. La mise en œuvre de cette obligation se heurte à la référence au secret des affaires avancée par les concepteurs de modèles d'IA générative.

Selon Adrien Basdevant,³⁸ « la différence entre le modèle d'IA générative de Meta, Google, Microsoft-OpenAI ou Mistral, n'est pas tant la ligne de code qui est très courte, que le fait de savoir sur quelle partie d'internet ils sont allés scrapper³⁹ des données, comment ils ont pondéré les données pour préparer le jeu de données et quels choix algorithmiques ils ont opéré ».

La transparence se heurte aussi à la faculté de compréhension par les citoyennes et citoyens. En d'autres termes, peu importe la transparence si les informations fournies ne peuvent être comprises par personne. Dans un rapport récent, l'UNESCO⁴⁰ fait l'éloge de l'IA explicable, et rappelle que « les exigences en matière d'explicabilité se multiplient, pour le bien des utilisateurs. L'IA explicable (ou XAI) fait tomber le mythe de la boîte noire (...) en présentant des informations techniques compréhensibles pour certaines catégories d'utilisateurs ». « Elle fournit des explications précises permettant d'ouvrir la voie à des audits. Ces informations aideraient certains membres du grand public (utilisateurs, chercheurs, représentants d'organes de gouvernance ou auditeurs) à garantir le respect des principes d'équité, d'égalité, de diversité, d'égalité et de confidentialité et à prendre des mesures pour corriger d'éventuels manquements, par exemple en introduisant des réclamations (judiciaires et non judiciaires) ». Or, aujourd'hui, peu d'IA

36 Les LLM (*Large Language Models* ou grands modèles de langage) sont des réseaux neuronaux utilisant d'énormes volumes de données pour comprendre le langage humain.

37 France Digitale et Wavestone, Guide de mise en conformité à l'AI Act : les clés pour comprendre et appliquer la loi sur l'IA, février 2024.

38 Audition d'Adrien Basdevant, avocat spécialisé en droit des nouvelles technologies, membre du CNNum, par la Commission temporaire IA, le 19 septembre 2024.

39 Le *scraping* est un mode de collecte de données qui vise à extraire automatiquement des données disponibles publiquement sur le web grâce à des robots.

40 Unesco, L'autonomisation des utilisateurs grâce aux réponses apportées par l'éducation aux médias et à l'information à l'évolution de l'intelligence artificielle générative (IAG), 2024.

sont explicables. Les concepteurs d'IA génératives sont rares à publier leur « Constitution », à savoir la liste des choix à l'origine de la programmation de l'IA⁴¹, et aucun ne fait la transparence sur les jeux de données d'entraînement. Des

initiatives françaises ont été lancées pour favoriser l'explicabilité des IA, comme celles du Programme et Équipement Prioritaire de Recherche⁴².

Les préconisations du CESE

→ Transparence des algorithmes

L'Union européenne a déjà adopté plusieurs régulations visant à encadrer l'utilisation de l'IA (Règlement sur l'Intelligence Artificielle (*AI Act*), *Data Act*, et *Data Governance Act*). Ces législations marquent des avancées significatives en matière de transparence, mais il reste encore des défis à relever pour garantir une transparence effective des algorithmes et des jeux de données.

Le CESE appelle à renforcer la transparence et l'explicabilité des algorithmes et des jeux de données d'entraînement utilisés pour des systèmes d'IA ayant un impact sur nos droits fondamentaux et libertés individuelles.

41 À ce jour, seule l'IA générative Claude, développée par Anthropic (derrière laquelle on trouve Google et Amazon) a publié sa constitution, calquée sur la Déclaration universelle des droits de l'homme des Nations-Unies.

42 Programme et équipement prioritaire de Recherche, institué dans le cadre France 2030 et piloté par le CEA, le CNRS, l'INRIA, l'ANR et programme « Mathématiques de l'apprentissage profond ».

PRÉCONISATION #1

Le CESE appelle à renforcer, dans la réglementation européenne, les impératifs de transparence des algorithmes et des jeux de données d'entraînement, en :

- exigeant des concepteurs de modèles d'IA qu'ils fournissent des explications claires sur le mode de prise de décision des algorithmes ;
- mettant en place un système d'audit externe régulier, indépendant et public de ces IA par le Bureau de l'IA¹ ;
- incluant des dispositions renforçant l'accès direct des citoyens à leurs propres données ;
- fournissant une meilleure documentation sur la provenance des données, les méthodes de collecte, et les biais potentiels ;
- introduisant des certifications ou des labels de transparence pour les jeux de données partagés
- améliorant la traçabilité des données par la documentation des transformations subies par les données (nettoyage, agrégation, anonymisation) avant leur utilisation ;
- écrivant l'égalité dans le code lui-même pour que l'IA la promeuve.

1. Le Bureau européen de l'IA est le centre d'expertise en matière d'IA dans l'ensemble de l'UE. Il joue un rôle clé dans la mise en œuvre de la législation sur l'IA – en particulier pour l'IA à usage général. Il est rattaché à la Commission européenne

→ Diversité des jeux de données

Diversifier les bases de données est une condition essentielle pour assurer une meilleure représentativité et lutter contre la reproduction des biais et des inégalités par les algorithmes. Seule cette diversification permettra d'assurer une représentation plurielle du monde social et des perceptions.

En outre, certains prédisent que l'IA subira bientôt une pénurie de données, quand les données disponibles sur Internet auront toutes été scrappées, et que seuls quelques-uns auront les moyens financiers d'accéder à des données récentes (au travers de contrats d'accès aux données négociés à des prix inabordables pour la plupart des plus

petits producteurs d'IA).

Face à ce risque de pénurie, un nouveau type de données se développe : les données synthétiques. Il s'agit de données générées artificiellement par un algorithme d'IA qui a été entraîné sur un ensemble de données réelles. Si ce type de données présente des avantages incontestables (respect de la vie privée, sécurité des données, facilitation de l'accès aux données pour toutes et tous), il s'expose à des risques qu'il convient d'anticiper, notamment si le modèle d'IA qui a créé les données synthétiques les a produites sur des jeux de données contenant des biais humains historiques, qui risqueraient d'être amplifiés et reproduits à l'échelle.

PRÉCONISATION #2

Favoriser la diversité des données pour réduire les biais, en :

- **veillant à diversifier les bases de données et les équipes de conception et la sensibilisation des concepteurs ;**
- **favorisant la création de jeux de données synthétiques non biaisées, anonymisées, respectueux de la vie privée, sécurisés et représentatifs ;**
- **faisant émerger, notamment par la commande publique, des acteurs français et européens en capacité de créer des jeux de données synthétiques.**

→ Protection des données personnelles**PRÉCONISATION #3**

Clarifier et renforcer au sein des équivalents de la CNIL française dans chaque pays européen, les modalités de protection des données personnelles dans le cadre de l'IA, et notamment les modalités d'utilisation, de réutilisation et de rectification des données personnelles publiées sur le web - et notamment les réseaux sociaux - pour l'entraînement des IA, en conformité avec les dispositions prévues dans le RGPD.

→ Construction d'une société égalitaire

PRÉCONISATION #4

Le CESE recommande aux pouvoirs publics de déployer une stratégie et des actions concrètes pour que l'IA participe à la construction d'une société égalitaire et respectueuse des femmes. Il s'agit de :

- sensibiliser les acteurs prenant part au développement de solutions à base d'intelligence artificielle sur l'enjeu de construire une société égalitaire et respectueuse des femmes ;
- intégrer les questions de genre dans des campagnes de sensibilisation du public et de formation des jeunes en situation de scolarité sur les outils et usages numériques, pour faire face à la montée de la cyberviolence, aux risques qu'encourent les femmes concernant leur image sur les réseaux numériques, et aux responsabilités pénales des auteurs ;
- renforcer les moyens d'actions juridiques face au vol d'images à des fins pornographiques.

Recommandation de citoyennes et citoyens

L'IA offre des opportunités significatives, mais ses impacts négatifs sur les libertés, l'équité et la démocratie nécessitent des mesures de contrôle strictes. Le CESE propose une régulation proactive pour garantir un usage responsable et respectueux des droits des citoyens et citoyennes.

B. L'empreinte environnementale de l'IA

1. Le coût environnemental de l'IA

Entre l'électricité nécessaire pour faire tourner les vastes centres de données, l'eau pour les refroidir ou l'extraction des métaux pour la fabrication des équipements, le déploiement des systèmes algorithmiques pose un véritable enjeu environnemental. À titre comparatif, une seule requête avec ChatGPT nécessite jusqu'à 10 fois plus d'énergie qu'une recherche en ligne

classique sur Google, et l'usage de GPT-3 « boirait » l'équivalent d'un demi-litre d'eau toutes les 10 à 50 requêtes⁴³.

Mesurer le coût environnemental de l'IA nécessite de mesurer celui du numérique, le premier étant difficilement distinguable du second. Il s'agit alors de tenir compte de l'augmentation massive des usages du numérique depuis 2022, année à partir de laquelle nous pouvons dater le début de la généralisation de l'utilisation de l'intelligence artificielle

43 Usbek & Rica, *Intelligence artificielle : 10 propositions pour reprendre le contrôle*, n° 44, 2024.

par le grand public. S'ensuit donc une augmentation de la production d'infrastructures, de la consommation électrique et de différentes ressources, et donc d'une croissance exponentielle de l'empreinte écologique de l'IA et, plus globalement, du numérique.

1.1 LE COÛT ENVIRONNEMENTAL DU NUMÉRIQUE EST EN FORTE CROISSANCE

La mesure de l'impact environnemental du numérique nécessite une évaluation globale basée sur une analyse de cycle de vie multi-composants (terminaux, réseaux et centres de données), multicritère (empreinte carbone mais aussi 10 autres indicateurs environnementaux⁴⁴) et multi-étapes (phases de fabrication, distribution, utilisation et fin de vie).

S'agissant de l'empreinte carbone, l'ADEME et l'Arcep⁴⁵ évaluent qu'en 2020 en France, un an de consommation de biens et services numériques représente 2,5 % de l'empreinte carbone nationale, soit 17,2 Mt CO₂eq, les terminaux étant responsables de 79 % de l'empreinte carbone du numérique, les centres de données de 16% et les réseaux de 5 %. La phase de fabrication concentre l'essentiel des impacts environnementaux (78 %

de l'empreinte carbone), alors que la phase d'utilisation représentait 21 %. L'épuisement des ressources abiotiques (minéraux et métaux) ressort déjà comme un critère pertinent pour décrire l'impact environnemental du numérique.

Cette évaluation a été complétée en 2023 par une analyse prospective⁴⁶ de l'impact environnemental du numérique en projetant différents scénarios. À horizon 2030, si rien n'est fait pour réduire l'empreinte environnementale du numérique et que les usages continuent de progresser au rythme actuel, le trafic de données serait multiplié par 6 et le nombre d'équipements augmenterait de près de 65 % en 2030 par rapport à 2020, notamment du fait de l'essor des objets connectés. Il en résulterait des augmentations, entre 2020 et 2030 :

- de l'empreinte carbone du numérique en France : environ 45 % (pour atteindre 25 Mt CO₂eq)⁴⁷ ;
- de la consommation de ressources abiotiques (métaux et minéraux) : +14 % ;
- de la consommation électrique finale en phase d'usage : +5 % (pour atteindre 54 TWh par an).

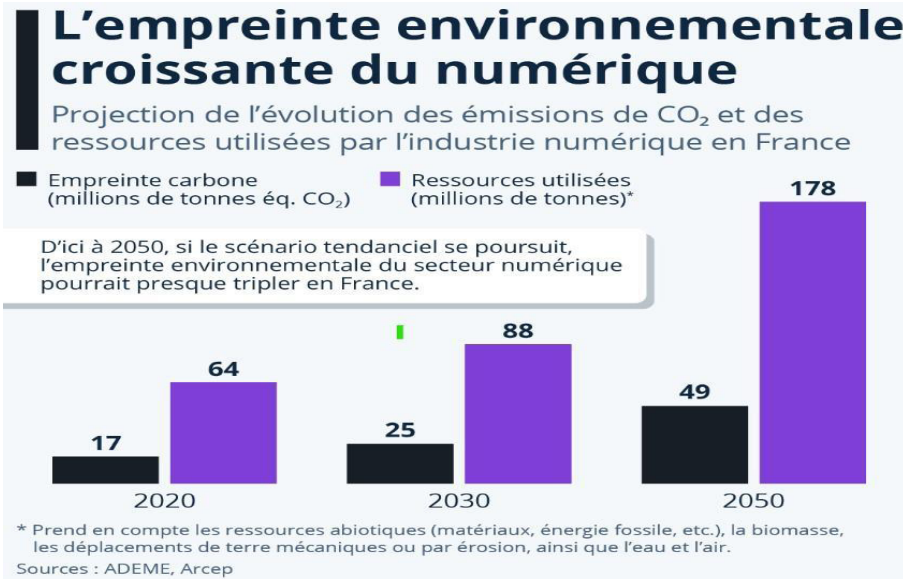
44 L'ADEME a identifié 11 indicateurs environnementaux : l'épuisement des ressources abiotiques (fossiles, minérales & métaux), l'acidification, l'écotoxicité, l'empreinte carbone, les radiations ionisantes, les émissions de particules fines, la création d'ozone, les matières premières, la production de déchets, la consommation d'énergie primaire, et la consommation d'énergie finale.

45 Étude ADEME-Arcep (2022-2023) sur le numérique en France.

46 ADEME, Analyse prospective publiée en mars 2023 : cet exercice de prospective est par nature complexe et requiert en général des révisions régulières ; il l'est d'autant plus pour un secteur en évolution très rapide caractérisé par de nombreux effets croisés internes et des externalités positives ou négatives sur les autres secteurs qui ne peuvent ici pas être pris en compte.

47 Afin de donner un ordre de grandeur, les émissions totales de la France en 2023 sont estimées à 385 Mt CO₂eq selon le Ministère de la Transition écologique, de l'Energie, du Climat et de la Prévention des risques).

Graphique 1 : Projection de l'évolution des émissions de CO₂ et des ressources utilisées par l'industrie numérique en France



Source : ADEME, Arcep

Rappelons que l'impact carbone du secteur numérique, et par extension, de l'IA, peut varier en fonction du pays où la technologie est développée et utilisée. En effet, faire du numérique dans un pays ayant un mix électrique faiblement carboné, comme la France, rend de fait cette activité plus sobre, durable et responsable. L'empreinte carbone de la production de matériels numériques est très dépendante de l'intensité carbone de l'énergie des pays producteurs. Celle-ci diminue tendanciellement au niveau mondial mais elle varie fortement (selon les pays du simple au triple). Par ailleurs, lors de l'usage du numérique, l'empreinte carbone dépend fortement

du bouquet de production d'électricité dans les pays d'utilisation. Elle varie d'un facteur un à plus de dix selon les pays⁴⁸. À titre d'exemple, le coût carbone (en kg de CO₂eq) d'un entraînement unique de GPT-3 équivaut à 93 300 kg pour la France, contre 858 360 kg pour l'Inde ou encore 485 160 kg en Corée du Sud⁴⁹. Les écarts sont significatifs, si bien que faire de l'IA en France apparaît comme une solution davantage soutenable.

48 Etude Rexecode.

49 Artificial Intelligence and the Climate Emergency: Opportunities, Challenges, and Recommendations, CellPress.

1.2 L'EMPREINTE ENVIRONNEMENTALE DE L'IA

Le CESE a rendu, en septembre 2024, un avis sur les *Impacts de l'intelligence artificielle : risques et opportunités pour l'environnement*⁵⁰. Dans cet avis, le CESE dresse d'abord le constat que l'IA peut décupler les risques pour l'environnement.

En effet, l'IA aggrave l'empreinte environnementale du numérique, dans ses trois briques matérielles (terminaux, réseaux et centres de données) et dans toutes les étapes du cycle de vie (fabrication, distribution, utilisation et fin de vie), mais également à accroître la consommation d'électricité, de métaux rares, d'eau et d'espaces artificialisés imputable aux systèmes d'IA.

Figure 1 : L'empreinte environnementale de l'IA



Source : Université Paris Sud

⁵⁰ CESE, *Impacts de l'intelligence artificielle : risques et opportunités pour l'environnement*, avis 2024-014, rapporté par Fabienne Tatot et Gilles Vermot Desroches au nom de la commission Environnement, septembre 2024. Impacts de l'intelligence artificielle : risques et opportunités pour l'environnement | Le Conseil économique social et environnemental.

1.3 LES FACTEURS D'AUGMENTATION DE L'EMPREINTE ENVIRONNEMENTALE DE L'IA

L'écosystème d'IA est caractérisé par un fonctionnement en boucle, où les usages appellent de nouvelles capacités et les nouvelles capacités génèrent de nouveaux usages. Ainsi, l'intensification des usages de l'IA impose de développer une puissance de calcul supplémentaire, pour traiter de plus en plus de données, et de multiplier les calculs pour tenir compte de la diversité des usages (la puissance de calcul se retrouvant de plus en plus diffusée dans de micro-objets connectés).

L'IA alimente en outre « l'effet rebond »⁵¹: les gains d'efficacité énergétique générés par les innovations technologiques sont souvent annulés par une augmentation des usages et de la production d'infrastructures numériques.

Cet effet se cumule également avec le phénomène de l'obsolescence, qui est alimenté par l'attrait des consommateurs pour la nouveauté et la recherche de profits par les producteurs : les représentants de la Quadrature du Net ont ainsi expliqué que les nouveaux équipements et usages s'ajoutent aux anciens, augmentant la consommation globale d'énergie et d'eau, tout en entraînant une obsolescence accélérée des infrastructures existantes.

1.4 L'IA PEUT CONTRIBUER À RÉDUIRE L'EMPREINTE ENVIRONNEMENTALE DE CERTAINS USAGES

Le CESE⁵² rappelle toutefois que l'IA peut proposer des solutions opportunes pour réduire l'empreinte environnementale globale :

- soit directement, lorsque les systèmes d'IA sont créés spécifiquement pour leur utilité écologique, par exemple lorsqu'ils contribuent à la mesure et la gestion de la qualité de l'air en anticipant à l'avance la concentration de particules fines ou encore à des pratiques agricoles plus durables ;
- soit indirectement, par une amélioration de l'efficacité énergétique de l'ensemble des activités (au travers du pilotage des flux de circulation routière ou ferroviaire, de la gestion intelligente du chauffage des bâtiments, ou l'optimisation en temps réel de l'usage et la distribution fine d'énergie électrique au sein d'une ville.

Plusieurs exemples ont également été fournis par le directeur numérique du groupe SNCF, lors de son audition par la Commission temporaire, sur les gains d'efficacité énergétique que peut apporter l'IA : ainsi l'outil d'écoconduite basé sur l'IA « Opti'conduite »⁵³, qui a été développé en interne, permet de réduire la consommation d'électricité des trains de 7 à 12 %, dans un contexte où la SNCF est le premier consommateur d'électricité

51 L'effet rebond, aussi appelé « paradoxe de Jevons » désigne un phénomène observé lorsque les économies d'énergie attendues avec l'utilisation d'une ressource ou de technologie plus efficace énergétiquement ne sont pas obtenues, voire aboutissent à des surconsommations, à cause d'une adaptation des comportements.

52 CESE, *Impacts de l'intelligence artificielle : risques et opportunités pour l'environnement*, avis 2024-014, rapporté par Fabienne Tatot et Gilles Vermot Desroches au nom de la commission Environnement, septembre 2024. Impacts de l'intelligence artificielle : risques et opportunités pour l'environnement | Le Conseil économique social et environnemental.

53 Opti'conduite est un logiciel d'aide à l'éco-conduite des trains développé dans le cadre d'un logiciel, qui rassemble plusieurs paramètres liés à la circulation des trains (vitesse, position du train, profil de la voie...) permet de calculer et d'indiquer en temps réel au conducteur la vitesse idéale à adopter pour réduire la consommation d'énergie du train (jusqu'à -12 %), tout en respectant les horaires d'un trajet.

industrielle en France et promeut les modes de transport décarbonés.

2. Les conflits d'usages

Les infrastructures de l'IA comptent pour beaucoup dans l'empreinte environnementale de l'IA.

Cela est dû notamment aux semi-conducteurs, à savoir des puces électroniques qui permettent aux modèles d'IA de s'entraîner pour faire tourner les algorithmes. Plus ces dernières sont puissantes, plus elles accélèrent la vitesse de calcul et de traitement des informations. Or, comme l'a souligné La Quadrature du Net⁵⁴ lors de son audition, la fabrication des puces électroniques nécessite des procédés de plus en plus intensifs en ressources, en énergie et en pression sur le travail humain, en raison de plusieurs critères cumulatifs : l'intensification de l'extraction de métaux rares, la consommation accrue d'eau pure pour la fabrication, les conditions de travail précaires dans certains sites de production, et le taux de recyclage faible des composants et matériaux utilisés. De plus, le marché de ces puces devrait croître de manière exponentielle, selon une étude du cabinet AltIndex⁵⁵ : le marché des semi-conducteurs pour l'IA pourrait être multiplié par dix en dix ans, pour atteindre un revenu global annuel de 341 milliards de dollars en 2033, contre 30 milliards en 2023.

Cela est dû également à l'implantation des centres de données (data centers) où

sont regroupés les équipements constitutifs d'un système informatique (ordinateurs centraux, serveurs, centres de stockage, équipements réseaux et de télécommunications, etc.). Ce regroupement facilite la sécurisation, la gestion (notamment l'exécution de calculs et le refroidissement) et la maintenance des équipements et des données stockées, et sont essentiels au fonctionnement de l'intelligence artificielle.

Les centres de données mobilisent un nombre croissant de ressources naturelles (eau, air et terre) ou énergétiques (pour fabriquer l'électricité), et ont pour effet de rejeter de la chaleur (en milieu urbain ce phénomène soulève divers problèmes qui sont aggravés par l'augmentation de la fréquence des canicules et le manque d'espaces verts), des gaz fluorés (ces gaz à fort impact d'effet de serre sont utilisés pour le refroidissement), et d'utiliser de l'eau de qualité en grande quantité pour le refroidissement, soulevant ainsi des conflits d'usage (irrigation des cultures, production industrielle, usages sanitaires, etc.). Aussi, l'implantation de centres de données nécessite des surfaces terrestres, et la circulation des données implique de tisser un réseau de câbles sous-marins : l'artificialisation des sols et des fonds marins qui

⁵⁴ Audition de Félix Tréguer, chercheur associé au Centre Internet et société du CNRS, et Edlira Nano, informaticienne, membres de La Quadrature du Net, par la Commission temporaire IA, le 17 octobre 2024.

⁵⁵ Actutem, Le marché des puces électroniques pour l'IA promis à une fulgurante croissance jusqu'en 2033, 2024.

en découle génère des effets néfastes sur l'environnement, déstabilise les écosystèmes, risque d'accélérer la perte de biodiversité. Ces espaces artificialisés font l'objet de conflits d'usages.

Il faut néanmoins souligner que la consommation énergétique et de ressources des centres de données fait l'objet d'une attention de certains constructeurs. Par exemple, le refroidissement liquide en circuit fermé permet de réduire de 35 % la consommation d'énergie (par rapport à l'air conditionné)⁵⁶ tout en minimisant l'impact sur les ressources en eau. Il reste encore des défis à relever pour minimiser complètement l'impact environnemental des data centers.

De récents exemples illustrent ces conflits d'usages⁵⁷ :

- le data center de Twitter/X à Memphis (Etats-Unis) consomme 3,8 millions de litres d'eau par jour et 150 MW d'électricité, et émet jusqu'à 130 tonnes de NO₂ par an ;
- les entreprises Google et Meta aux États-Unis ont obtenu le maintien de l'ouverture et de l'exploitation de mines de charbon pour pouvoir alimenter en électricité leurs data centers
- l'implantation des data centers de Microsoft en Iowa (Etats-Unis) et aux Pays-Bas a augmenté de façon significative la consommation d'eau, provoquant des conflits d'usage en période de sécheresse ;

- Google au Chili et Meta en Espagne ont fait face à d'importants mouvements d'opposition au niveau local en raison d'une consommation d'eau élevée pour faire fonctionner leurs data centers dans des régions qui connaissent des pénuries récurrentes.

La France compte actuellement plus de 300 centres de données⁵⁸, la situant au 6ème rang mondial des pays en accueillant le plus, après les États-Unis, l'Allemagne, le Royaume-Uni, la Chine et le Canada. L'Île-de-France est le premier territoire d'implantation avec 95 centres de données, suivi par Marseille qui en dénombre 12.

Des conflits d'usages sont déjà avérés.

Ainsi, à Paris, le *Digital Realty*, un *data center* de type *Hyperscale*⁵⁹ dédié à l'IA, qui occupe la surface de sept terrains de football sur plusieurs étages, consomme 295 000m³ d'eau par an, ce qui représente plus de 6 000 fois la consommation moyenne d'un habitant parisien⁶⁰. Face à ce niveau de consommation, la préfecture a exigé des mesures de réduction afin de prévenir de potentiels conflits d'usage.

À Marseille, douze centres de données génèrent des conséquences environnementales qui entrent en conflit avec les projets publics de décarbonation. Rejets de chaleur et gaz fluorés soulèvent des problèmes majeurs de santé publique en milieu urbain, dont la concentration ne facilite pas l'évacuation de la chaleur résiduelle. L'eau potable est utilisée pour le refroidissement des systèmes, alors

56 Navigating Liquid Cooling Architectures for Data Centers with AI Workloads, Schneider Electric, 2024.

57 Audition de Félix Tréguer, chercheur associé au Centre Internet et société du CNRS, et Edlira Nano, informaticienne, membres de La Quadrature du Net, par la Commission temporaire IA, le 17 octobre 2024.

58 La Quadrature du Net, Enquête : à Marseille comme ailleurs, l'accaparement du territoire par les infrastructures du numérique, 2024.

59 Sans être une définition officielle, un *datacenter Hyperscale* doit dépasser 5 000 serveurs et 929,03 m². Les *datacenters Hyperscale* se distinguent principalement par le volume de services de données, de calcul et de stockage qu'ils traitent.

60 Selon eaudeparis.fr, la consommation moyenne parisienne est de de 120L par jour par habitant.

même que l'assainissement est financé en partie par l'État et les collectivités territoriales, et que cette ressource naturelle est rare en période estivale (combinant forte chaleur et faible précipitation).

Plusieurs articles publiés sur le site de La Quadrature du Net mettent en avant le phénomène de *river cooling*, c'est-à-dire la privatisation et l'accaparement des ressources en eau de qualité potable. Par exemple, les eaux fraîches de la cunette de la Galerie à la Mer de Marseille qui sont à 15,5 degrés toute l'année, sont captées et injectées dans les circuits de refroidissement des centres de données MRS2, MRS3 et MRS4 de *Digital Realty*, pour échanger leurs *frigories* contre des *calories* dans des échangeurs thermiques. Elles repartent ensuite, réchauffées à environ 27 degrés, directement dans la Galerie à la Mer, à destination de la Méditerranée au niveau du Cap Pinède. Dès 2018, une enquête publique de la Direction régionale de l'environnement, de l'aménagement et du logement (DREAL) de Provence-Alpes-Côte d'Azur soulignait les inquiétudes relatives à ce type de pratiques : risques d'eutrophisation (déséquilibre du milieu marin provoqué par l'augmentation de la concentration d'azote et de phosphore), de prolifération d'algues filamenteuses et d'aggravation du réchauffement de la mer Méditerranée, alors même que cette enquête reposait sur des projections de rejet à 21,5°C. Dans un contexte d'accroissement des épisodes climatiques extrêmes où les sécheresses

s'intensifient dans les régions du sud de la France mettant en cause l'approvisionnement en eau potable ou à usage agro-alimentaire, il s'avère donc nécessaire que les eaux soient, en priorité, allouées à des usages d'utilité publique ou pour les milieux qui en ont besoin.

De même, l'implantation de nouveaux centres de données crée une contrainte forte sur les producteurs d'électricité, les réseaux de transport et les autres consommateurs de puissance électrique (collectivités, entreprises, citoyens et citoyennes). En découlent plusieurs conséquences ayant un impact environnemental fort, notamment la décision, dans certains pays, de conserver des centrales à charbon malgré la fermeture prévue dans le cadre de la transition écologique, ou la réouverture de centrales nucléaires et le financement de nouvelles tranches par les Big Tech, ces derniers pouvant devenir eux-mêmes producteurs d'électricité sur leurs propres sites.

Dans cette perspective de forte croissance des systèmes d'IA, les scénarios les plus ambitieux publiés par RTE en 2023 pourraient être sous-évalués : ils prévoyaient 28 TWh en 2035, alors qu'il est possible que la consommation électrique des centres de données atteigne 50 TWh voire 70 TWh. Par comparaison, la consommation électrique actuelle de l'Ile-de-France est légèrement supérieure à 60 TWh⁶¹. Le récent forum organisé par l'IEA sur l'IA et l'Énergie a mis en évidence la

⁶¹ Bilan électrique francilien 2022 | RTE.

dé-corrélation entre le rythme, voire la capacité, d'investissement des opérateurs d'électricité et l'augmentation

considérable des besoins de l'industrie des centres de données due à l'IA⁶².

Les préconisations du CESE

Ces préconisations viennent en complément de l'avis *Impact de l'intelligence artificielle : risques et*

opportunités pour l'environnement adopté par le CESE le 24 septembre 2024.

→ Implantation des centres de données

PRÉCONISATION #5

En complément des préconisations émises dans son avis *Impacts de l'intelligence artificielle : risques et opportunités pour l'environnement*, le CESE confirme la nécessité de réaliser des études d'impact environnementales et sociales avant toute implantation d'un nouveau centre de données. Le Conseil recommande en particulier, de :

- **systematiser et publier toutes les enquêtes sur la consommation réelle d'eau pour refroidir les centres de données notamment les eaux fraîches prélevées en milieu naturel ou de qualité potable, tout comme les possibles rejets d'eau chauffée en mer ;**
- **informer les publics sur la consommation électrique réelle des centres de données et de réaliser des projections régulières dans l'hypothèse où l'utilisation des systèmes d'IA s'intensifierait à l'avenir ;**
- **réaliser une étude complète de l'impact environnemental des centres de données sur leur cycle de vie, en amont de leur implantation et en anticipant les conflits d'usage qui sont susceptibles d'émerger à l'échelle locale.**

→ Planification du développement de l'IA

Comme l'a rappelé le représentant du Shift Project lors de son audition, ne pas planifier le numérique, c'est laisser ce secteur croître sans respecter les limites

planétaires, préempter les ressources naturelles et énergétiques, et créer de multiples conflits d'usages.

C'est pourquoi, le système numérique français notamment en raison du développement des systèmes d'IA et de nouveaux centres de données, ne peut plus être un impensé de l'aménagement du territoire, ni être analysé uniquement sous l'angle de la performance économique à court terme. Une planification spatiale et énergétique s'avère donc nécessaire aussi bien à l'échelle de notre pays qu'au niveau européen.

De même, les émissions de gaz à effet de serre (GES) du numérique prennent une part croissante dans l'impact environnemental global du pays, et devraient quintupler d'ici 2050⁶³ si aucune politique publique n'est mise en place. L'objectif de neutralité carbone à cet horizon nécessite d'inclure le secteur du numérique dans la stratégie de décarbonation et de sobriété énergétique.

63 The Shift Project, Des réseaux sobres pour des usages connectés résilients, mars 2024.

PRÉCONISATION #6

Le CESE demande à l'État de planifier le développement de l'IA dans le respect des limites planétaires en :

- fixant une trajectoire de décarbonation spécifique à l'IA dans la prochaine stratégie nationale bas carbone : il faut conditionner les logiques industrielles et commerciales de ce secteur à la compatibilité de son développement à la trajectoire de décarbonation globale ;
- orientant l'innovation des usages vers une dynamique vertueuse via les leviers d'action suivants :
- choisir de manière démocratique les usages du numérique en fonction de leur utilité sociale et de leur impact environnemental ;
- travailler sur l'utilisation des ressources (eau, chaleur) pour limiter leur perte « fatale » ;
- utiliser l'IA avec une énergie décarbonée ;
- valoriser les bonnes pratiques d'usage aux niveaux individuels et collectifs qui répondent à des besoins réels. par exemple, baisser la résolution d'une vidéo permet de requalifier le besoin et de réduire l'envergure des actions prises par les concepteurs ;
- généraliser l'obligation d'émission d'un bilan environnemental de l'IA à l'échelle internationale ;
- intégrer l'éco-conception dans la formation initiale et continue des développeurs d'IA.

→ IA frugale

De plus en plus d'utilisateurs d'IA, à commencer par la puissance publique, ont des attentes concernant son impact environnemental. Pour y répondre et réduire globalement cet impact dans toutes les organisations, il est primordial de disposer d'un référentiel pour outiller les acteurs à l'évaluation de ces impacts.

Un groupe de travail a été lancé en janvier 2024, piloté par l'Ecolab du Commissariat général au développement durable (CGDD), en partenariat avec l'AFNOR et dans le cadre de la Stratégie nationale pour l'IA (SNIA)⁶⁴. Il a établi un référentiel⁶⁵ qui contient une méthodologie pour évaluer l'impact environnemental de l'IA et recenser toutes les bonnes pratiques qui peuvent être mises en œuvre par les organisations pour réduire cet impact. Les entreprises peuvent ainsi communiquer sur le caractère frugal de leurs systèmes d'IA, et se différencier en se basant sur un cadre de référence reconnu et partagé. Ce référentiel doit devenir un point de repère pour les entreprises pour la valorisation de leur savoir-faire en IA frugale et pour la maîtrise de l'impact environnemental de leur solution d'IA.

PRÉCONISATION #7

Le CESE soutient le référentiel élaboré par l'AFNOR, appelle à ce que son application soit labellisée « IA frugale » et généralisée dans les meilleurs délais afin de promouvoir un service frugal d'IA pour lequel :

- **la nécessité de recourir à un système d'IA plutôt qu'à une autre solution moins consommatrice pour répondre au même objectif a été démontrée ;**
- **de bonnes pratiques sont adoptées par le producteur, le fournisseur et le client pour diminuer les impacts environnementaux du service utilisant un algorithme d'IA ;**
- **les usages et les besoins visent à rester dans les limites planétaires, et ont été préalablement questionnés.**

⁶⁴ Initiative du ministère de la Transition écologique et de la cohésion des territoires, l'AFNOR Spec 2314 énonce des méthodologies de calcul et des bonnes pratiques pour mesurer et réduire l'impact environnemental de l'IA, et pour communiquer avec des allégations justes et vérifiables.

⁶⁵ Le référentiel général pour l'IA frugale est le résultat de 6 mois de travail avec 150 contributeurs, issus des entreprises, de la recherche, des associations et des administrations. Ce référentiel constitue une première au niveau international pour offrir à tous les acteurs des services numériques un référentiel, d'application volontaire, accessible et compréhensible par tous.

Recommandation de citoyennes et citoyens

L'IA offre des bénéfices comme des innovations rapides et des outils d'assistance, mais son impact environnemental est préoccupant : elle consomme énormément de ressources, notamment à cause des centres de données (Data Center) nécessaires à son fonctionnement. Pour limiter cet impact, le CESE recommande plus de transparence envers les utilisateurs sur l'impact environnemental de l'IA, de planifier une utilisation responsable de l'énergie, de surveiller l'empreinte écologique des systèmes d'IA et de favoriser des infrastructures durables. L'objectif est de concilier avancées technologiques et respect de la planète.

C. L'impératif d'autonomie stratégique européenne pour l'IA

Le développement de l'IA s'inscrit dans un contexte géopolitique global où la chaîne de valeur de cette technologie (de la production de semi-conducteurs à la conception de systèmes d'IA) est largement dominée par les États-Unis et la Chine.

Une comparaison des puissances technologiques entre États est complexe, compte tenu du manque de transparence sur les ressources informatiques dont disposent certains pays ou certaines entreprises, et de la surenchère financière et technologique permanente que se livrent les concepteurs d'infrastructures et modèles d'IA. Cependant, les ordres de grandeur suivants permettent d'évaluer l'écart qui se creuse avec l'Europe :

- les États-Unis disposent de plus de chercheurs en IA que l'Europe, et ces derniers sont mieux valorisés dans leurs compétences et leurs conditions de travail;
- une seule entreprise américaine, *OpenAI*, aura accès en 2025, à un centre de données de 100 000 semi-conducteurs (puces GPU) dont les performances de calcul seront cinq fois

supérieures à celles déployées dans Jupiter, le supercalculateur européen le plus puissant de notre continent, qui dispose lui de « seulement » 24 000 semi-conducteurs⁶⁶ ;

- sur les 11 modèles avancés d'IA générative développés par des entreprises européennes (françaises, britanniques, finlandaise et allemande) depuis 2017, tous ont utilisé des processeurs développés par des Américains : Nvidia, AMD, Google⁶⁷.

Cela a des conséquences directes sur la capacité de l'Europe et de la France à imposer un modèle de déploiement de l'IA respectueux des limites planétaires, de nos principes démocratiques, de nos libertés individuelles et collectives, et de nos droits fondamentaux.

Ainsi, le fait que des acteurs économiques privés, essentiellement non-européens, détiennent de façon opaque les données, les algorithmes et les infrastructures d'organismes et de services publics pose un réel problème de fond.

Prenant conscience de ces risques, l'Union européenne appelle, depuis plusieurs années, à une autonomie stratégique vis-à-vis de certaines

66 Euractiv, *Malgré ses efforts, le retard de l'Europe en matière d'IA se creuse, novembre 2024.*

67 Le Grand Continent, *Les États-Unis devancent l'Union et la Chine en matière d'IA, avril 2024.*

technologies critiques, dont l'IA.

Pour autant, la problématique de l'autonomie stratégique ne se limite pas à ses dimensions économiques et concurrentielles pour les entreprises nationales et européennes. Elle doit avant tout préserver la capacité collective de garantir des intérêts publics majeurs dans des domaines tels que la défense, l'éducation, la santé, etc. ainsi que la détention des données stratégiques que les administrations publiques détiennent, le tout dans le respect des limites planétaires.

La valorisation des GAFAM et des géants dits « de la tech » impliqués dans le développement massif de l'IA dépasse les 10 000 milliards de dollars. Dans ce contexte, il serait illusoire de laisser penser que la France, et même l'Europe, auraient à la fois vocation et capacité à combler nécessairement un tel retard ou *gap* comparatif.

Force est de constater que les microprocesseurs nécessitent des savoir-faire technologiques et des investissements colossaux, que l'on ne trouve pas en Europe aujourd'hui sinon les machines de gravure optique d'ASML (NL) pour le dessin des puces. La réalisation de cartes et supercalculateurs à partir de ces briques, ainsi que la maîtrise des algorithmes (explicabilité, frugalité), paraissent plus abordables.

Il convient donc d'être conscients que la marge d'action des individus

et des régulateurs, y compris institutionnels (États, organisations internationales), est limitée face aux multinationales hyperpuissantes qui développent l'IA. Dans ce contexte dissymétrique, la plus-value de la France et de l'Europe ne saurait être essentiellement financière (le « combien »), mais se doit d'être qualitative et de résider, précisément, dans la préservation des intérêts publics – notamment via la maîtrise des données stratégiques propres (le « comment »), en tant que condition première au développement de l'IA.

La mise en œuvre de celle-ci se heurte à des difficultés pratiques, que le CESE a invité, dès 2022⁶⁸, à résoudre rapidement pour ne pas laisser l'Europe être davantage distancée.

1. Contexte géopolitique global

L'IA est une innovation duale : ses usages sont à la fois civils et militaires. Cette double nature explique pourquoi les États - en particulier les États-Unis, la Chine, et la Russie - se livrent, à l'échelle planétaire, la bataille de l'accès à la puissance technologique. Dans son rapport sur les risques géopolitiques de l'IA, Benjamin Pajot, chercheur à l'IFRI⁶⁹, explique que « *la course à l'IA a été régulièrement comparée à une nouvelle course aux armements* ». Cette perception a été accentuée non seulement par certaines déclarations

⁶⁸ CESE, *Ambition et leviers pour une autonomie stratégique de l'Union européenne dans le domaine économique*, avis 2022-013 rapporté par Catherine Lion et Olivier Mugnier au nom de la Commission Affaires européennes et internationales, septembre 2022. *Ambition et leviers pour une autonomie stratégique de l'Union européenne dans le domaine économique* | Le Conseil économique social et environnemental.

⁶⁹ Benjamin Pajot, *Les risques de l'IA, Enjeux discursifs d'une technologie stratégique*, IFRI, juin 2024.

volontairement tonitruantes, comme celles de Vladimir Poutine pour qui « *celui qui deviendra leader dans l'IA sera le maître du monde* » ou celles d'Elon Musk, pour qui « *l'IA est bien plus dangereuse que l'arme nucléaire* »⁷⁰, mais également par la rivalité sino-américaine, matérialisée par la politique de la Chine affirmant depuis 2017 son ambition de devenir leader international de l'IA en 2030, et la volonté des États-Unis de maintenir leur hégémonie technologique par tous les moyens. Cette compétition entre les deux grandes puissances a été renforcée avec la vague de l'IA générative. Pékin et Washington se sont ainsi engagés dans un bras de fer stratégique consistant notamment à priver le camp adverse des ressources indispensables au développement de l'IA, et à mettre en œuvre toutes les armes économiques et politiques à leur disposition pour favoriser leurs intérêts nationaux (préférence nationale dans la commande publique, investissements massifs dans des technologies souveraines, réglementation pro-innovation, assouplissement de la politique concurrentielle, dérogation aux règles du commerce international par l'imposition de tarifs douaniers, etc.).

Dans ce contexte, l'Europe tente de proposer une troisième voie, motivée par la défense de nos droits fondamentaux et libertés individuelles, le respect de notre feuille de route environnementale et climatique et l'impératif d'autonomie stratégique. Cette dernière, à laquelle il était peu fait référence dans l'Union européenne avant la pandémie de Covid-19 et l'invasion de l'Ukraine par la Russie, est désormais un objectif à part entière. Les rapports Letta⁷¹ et Draghi⁷²

témoignent du changement d'état d'esprit à l'œuvre en Europe : cette dernière semble considérer qu'elle pouvait toujours jouer un rôle prépondérant sur la scène internationale et imagine des politiques qu'elle pourrait déployer en vue d'accroître sa compétitivité économique. Indirectement, cela pourrait aussi l'aider à préserver, en partie, ses intérêts sociaux, économiques et politiques. Ces rapports appellent ainsi à prioriser les transitions écologiques et numériques, au travers de mesures économiques fortes. Le développement d'une IA européenne éthique et frugale peut également être une manière de se démarquer des concurrents à l'international et devenir un facteur de compétitivité. Dans le même sens, le déploiement d'un modèle de régulation, qui se veut garant des droits fondamentaux et des libertés individuelles peut être un vecteur de rayonnement à l'international.

Il est encore trop tôt pour analyser l'impact que pourrait avoir le second mandat de Donald Trump à la tête des États-Unis, épaulé par Elon Musk en tant que ministre de l'« Efficacité gouvernementale ». Cependant, les déclarations de ces derniers pendant la campagne électorale font état d'une volonté de déréguler l'IA aux États-Unis, de mener une politique commerciale agressive au travers du recours aux rétorsions commerciales, et de renforcer le protectionnisme américain.

2. La place de la France et de l'Europe dans la course à l'IA

Il est possible de qualifier l'importance d'un État dans une technologie, et en particulier dans l'IA, en fonction de

70 « *After Putin's AI comments, Elon Musk imagines World War III* », the Moscow AI Times, septembre 2017.

71 *Much More Than a Market*, Report by Enrico Letta, 2024.

72 *The future of European competitiveness*: Report by Mario Draghi, 2024.

plusieurs critères⁷³ :

- les **capacités financières** de ses entreprises : l'argent représente davantage l'ambition qu'une véritable puissance, mais permet d'acquérir les autres critères de la puissance (capacités de calcul, talents, données) et d'assurer une certaine résilience ;
- la capacité des entreprises **à toucher des utilisateurs** : certaines entreprises technologiques touchent, chaque mois, près de la moitié de la population mondiale⁷⁴ ;
- la capacité à **collecter et constituer d'immenses corpus de données** : ces mêmes entreprises, en touchant des utilisateurs, accèdent ainsi à leurs données personnelles ;
- la présence plus ou moins **dominante sur la chaîne de valeur de l'IA**. Cette dernière repose sur quatre couches fondamentales, qui englobent à la fois des éléments matériels et des couches logicielles : les puces, l'infrastructure (centres de données, réseaux telecoms), modèles d'IA et applications. La valeur économique réside principalement dans les puces et l'infrastructure. Des acteurs majeurs (tels qu'Amazon, Google, Microsoft et dans une moindre mesure, Nvidia) se distinguent par une présence à différents niveaux de la chaîne de valeur qui leur confère un avantage stratégique indéniable dans la capture de valeur⁷⁵ ;
- la faculté à **concevoir des algorithmes** pour tirer parti des données ;
- **l'accès à des puissances de calcul**, et donc à des puces et des sources d'énergie ainsi qu'à des moyens de refroidissement suffisants (débit d'eau) ;
- la capacité à **former et retenir des talents** dans l'IA : chercheurs, développeurs, concepteurs de modèles et services d'IA ;
- la capacité à **mettre la technologie au service de ses valeurs**. L'Europe, en forgeant l'une des premières régulations d'ampleur sur l'IA, tente ainsi de créer un effet géopolitique bien au-delà de ses frontières, le fameux « effet Bruxelles » (*Brussels Effect*), avec l'espoir de promouvoir ses normes et sa vision pour une IA au service de l'intérêt général, à l'échelle internationale⁷⁶ ;
- la capacité à créer des dynamiques avec le monde de la recherche sur des bases plus coopératives et durables que les appels à projets, et à valoriser

73 Audition de Laure de Roucy-Rochegonde, directrice du Centre géopolitique des technologies de l'IFRI, par la Commission temporaire IA, le 26 septembre 2024.

74 Francetv Info, De Facebook à Meta, cinq chiffres qui montrent comment le réseau de Mark Zuckerberg est devenu un géant économique, février 2024.

75 France Digitale, Des puces aux applications : l'Europe peut-elle être une puissance de l'IA générative ?, avril 2024.

76 Le RGPD (la réglementation européenne sur les données personnelles) avait déjà produit cet effet en 2016, en influençant une évolution de la réglementation sur les données personnelles en dehors de l'UE, notamment aux Etats-Unis (Californie, Colorado).

les compétences et les parcours des chercheurs.

Prenant acte de ce que certains considèrent comme un “retard technologique”, l’UE a fait du numérique l’une de ses deux priorités, aux côtés de la transition environnementale, et a mis en œuvre une stratégie politique et économique dans le but de faire émerger des champions technologiques européens d’envergure mondiale, dans certaines innovations de rupture dont l’IA.

En pratique, cette stratégie a conduit à :

- créer des fonds d’investissements publics pour financer directement ou au travers de fonds privés des entreprises européennes développant des technologies de rupture, de l’amorçage à la phase de croissance ;
- financer des projets de recherche fondamentale et appliquée, publique et privée ;
- déployer des appels à projets créant des collaborations entre États membres et entre plusieurs types de structures (laboratoires de recherche, startups, PME, grands groupes, organisations publiques) ;
- prévoir, dans les réglementations européennes, des mesures dérogatoires temporaires pour les structures innovantes en phase

de croissance (les “bacs à sable réglementaires”) ;

- concevoir et mettre à disposition des grands projets d’infrastructures stratégiques (tel que le supercalculateur Jupiter⁷⁷ donnant accès aux laboratoires de recherche et aux entreprises à de la puissance de calcul, ou des infrastructures de cloud européen, etc.) ;
- soutenir le développement des modèles et systèmes d’IA en source ouverte (“open source”). Il s’agit de technologies d’IA où le code source est librement disponible pour que quiconque puisse l’utiliser, le modifier et le distribuer⁷⁸ ;
- recourir à des partenariats stratégiques avec des acteurs extra-européens, lorsque certaines briques technologiques n’étaient pas réalisables en Europe, par exemple sur le cloud ou l’accès à la puissance de calcul.

Ces exemples mettent en lumière la volonté de l’Europe de concurrencer les principaux acteurs mondiaux et, d’une certaine manière, gagner en autonomie dans l’IA. Ils démontrent ainsi que l’Europe ne renonce pas : comme l’a rappelé en audition Yann Lechelle,⁷⁹ « *la bataille de la souveraineté technologique n’est jamais perdue, et doit se mener à l’échelle européenne* ». Pour le CESE⁸⁰,

77 Installé en Allemagne (2024) et qui sera suivi d’un deuxième supercalculateur *Exascale* (Jules Verne) installé en France en 2025 et sur lequel des parties quantiques seront ajoutées.

78 Le Sommet de l’IA organisé par la France en février 2025 pourrait être l’occasion d’appeler à la création d’une fondation dont le but serait de créer des briques d’IA open source et des bases de données dont pourraient se saisir tous les États et entreprises qui le souhaiteraient. Cette initiative s’inscrit dans le prolongement des objectifs de l’ONU de démocratiser l’IA pour tous les États qui n’y ont pas accès. Elle devrait toutefois définir au préalable ce qui est entendu par “open source”, dans un contexte où certains acteurs de l’IA mettent en avant des solutions open source qui ne permettent pas un véritable accès libre à leurs technologies.

79 Audition de Yann Lechelle, co-fondateur de la startup Probabl.ai, par la Commission temporaire IA, le 26 septembre.

80 CESE, *Ambition et leviers pour une autonomie stratégique de l’Union européenne dans le domaine économique*, avis 2022-013 rapporté par Catherine Lion et Olivier Mugnier au nom de la Commission Affaires européennes et internationales, septembre 2022. *Ambition et leviers pour une autonomie stratégique de l’Union européenne dans le domaine économique* | Le Conseil économique social et environnemental.

l'option stratégique de l'UE doit faire preuve d'ambition et de réalisme. Cependant, chercher une totale indépendance technologique de l'Europe dans l'IA serait un leurre. En revanche, l'Europe doit accepter de compléter son offre technologique par des collaborations, partenariats ou financements extra-européens,

afin de permettre l'émergence à moyen terme de solutions européennes, et toujours garantir le respect de notre socle de valeurs (droits fondamentaux et libertés individuelles) et nos réglementations européennes (sur l'IA, le droit d'auteur, etc.).

Encadré 1 : Zoom sur la place de la France dans l'IA

Même si le marché est hyper concentré autour de quelques acteurs américains ou asiatiques ultradominants sur la chaîne de valeur de l'IA, des alternatives françaises émergent sur certaines couches de cette chaîne de valeur, au niveau de la conception ou la fabrication de puces ou de cartes électroniques (avec des entreprises comme STMicroelectronics, Flex.ai, Kalray ou Sipearl), de la conception de centres de données et d'infrastructures de cloud (Scaleway, OvHCloud, Outscale, Orange Cloud ou encore Qarnot), de la production de modèles d'IA (Mistral, Hugging Face, Bloom, Light-on, Photoroom, LinguaCustodia ou encore Quicktext), ou de la programmation de logiciels d'IA⁸¹.

Les préconisations du CESE

La réglementation européenne en matière de commande publique date de 2014, et a été marginalement modifiée depuis, alors même qu'en dix ans, le contexte géopolitique, économique, environnemental, social et technologique s'est profondément transformé. À titre d'exemple, certains États comme les États-Unis ou la Chine dérogent ouvertement aux règles de l'Organisation mondiale du commerce, en imposant des sanctions commerciales ou en privilégiant l'échelon national dans la commande publique et les contrats d'innovation des grandes entreprises.

De plus, les grands groupes français bénéficiant d'aides à la

recherche et à l'innovation doivent également favoriser les solutions développées par les entreprises européennes s'engageant dans un numérique sobre et vertueux, que ce soit dans le cadre des appels d'offres européens et des contrats de partage voire de sous-traitance de l'innovation et à la recherche dans ces domaines spécifiques dits « stratégiques ».

⁸¹ France Digitale, Mapping des startups de l'IA, Mars 2023.

PRÉCONISATION #8

Le CESE appelle à la révision des règles de commande publique portant sur la fourniture de biens et services numériques, en :

- **intégrant des critères environnementaux et sociaux pour limiter l'empreinte environnementale des systèmes d'IA et ne pas aggraver ses conséquences sociales (transformation des emplois, formation à l'outil, etc.) ;**
- **instaurant une préférence communautaire pour choisir prioritairement des acteurs européens mieux à même de préserver l'autonomie stratégique des systèmes d'IA (protection des données, transparence des algorithmes, cybersécurité, etc.).**

Aux États-Unis, la politique de soutien à l'innovation de rupture est déployée au travers, notamment, des *Advanced Research Projects Agency (ARPA)*, à savoir des agences gouvernementales qui ont pour mission de modifier le modèle conventionnel de financement de la recherche et de soutenir, au travers d'investissements et de commandes publiques, des projets d'innovation de

rupture dans des domaines stratégiques (défense, énergie, biomédical, etc.).

En Europe, le *European Innovation Council (EIC)* joue ce rôle, dans une moindre mesure : doté d'un budget bien inférieur (1,4 milliards d'euros, vs. plus de 7 milliards d'euros pour les ARPA américaines⁸²), il finance des entreprises plus matures (alors que les entreprises technologiques ont besoin de soutien à l'amorçage), et souffre d'une approche bureaucratique (les délais de traitement des candidatures étant d'environ 240 jours, les taux de succès étant inférieurs à 8 %).

PRÉCONISATION #9

Le CESE appelle les États membres à renforcer le financement de la recherche vers l'innovation de rupture, dans le prochain cadre budgétaire pluriannuel de l'Union européenne qui débutera en 2028. Au-delà des financements publics, il est nécessaire de mobiliser les investisseurs privés, notamment institutionnels, pour qu'ils prennent leur part dans le financement des innovations européennes, malgré le risque associé.

En outre, l'un des principaux problèmes rencontrés par les entreprises européennes pour devenir des entreprises d'envergure internationale

82 Econpol, Bocconi Università, Toulouse School of Economics, EU Innovation Policy - How to Escape the Middle Technology Trap, 2024.

est la fragmentation du marché européen, notamment d'un point de vue réglementaire. Dans leurs rapports, Enrico Letta et Mario Draghi appellent à une harmonisation européenne du droit des affaires, au travers de la création d'une nouvelle forme sociale pour les entreprises, appelée le "28^{ème} régime", que ces dernières pourraient adopter sur la base du volontariat afin de bénéficier de règles juridiques communes. Le contenu de ce 28^{ème} régime ainsi que les modalités d'accès restent encore à définir ; le Commissaire européen à la Démocratie, à la Justice et à l'État de droit, Michael McGrath, en a fait l'une des priorités de son mandat.

PRÉCONISATION #10

Le CESE appelle les institutions européennes à harmoniser le droit des affaires dans l'UE. Cela ne devra pas aboutir à abaisser les niveaux existants de protection sociale, de droits sociaux, de droits des travailleurs et de droit environnemental. Cette harmonisation ne serait acceptable que dans un objectif de transition juste.

Dans son avis rendu en juin 2024⁸³, l'Autorité de la concurrence constate que l'IA - et en particulier l'IA générative - présente des risques de pratiques anticoncurrentielles élevées, compte tenu des

importantes barrières à l'entrée liées à la concentration d'un petit nombre d'acteurs dominants sur l'ensemble de la chaîne de valeur, de la nécessité de disposer de larges volumes de données, de la rareté des compétences disponibles, ou encore du besoin de financement important.

Toutefois, la preuve de tels comportements anti-concurrentiels peut être difficile à apporter : certains abus peuvent émerger d'une accumulation de pratiques de marché apparemment légitimes prises individuellement, mais devenant anti-concurrentielles lorsqu'elles sont appliquées de manière systématique par des entreprises dominant leur marché ou profitant de la dépendance de leurs partenaires. Si, à ce jour, la preuve de pratiques de marché ouvertement abusives n'a pas été rapportée, l'existence d'entreprises dominant le marché et disposant d'une structure verticalement intégrée n'est pas sans risque.

83 Autorité de la concurrence, AVIS 24-A-05 DU 28 juin 2024 relatif au fonctionnement concurrentiel du secteur de l'intelligence artificielle générative.

PRÉCONISATION #11

Le CESE appelle le Bureau de l'IA et l'autorité nationale compétente en France, qui sera désignée prochainement, à mobiliser les outils du droit de la concurrence, notamment le *Digital Markets Act*, pour contrôler et sanctionner les comportements abusifs ou dominants, et permettre l'émergence ainsi que l'expansion d'opérateurs de taille plus modeste notamment européens.

Le rapport de la Commission de l'IA de mars 2024⁸⁴ rappelle ainsi que « *Les talents constituent la principale ressource de la recherche publique comme des entreprises du domaine, dont ils représentent environ un tiers des coûts. Or, la France est en forte concurrence avec les États-Unis. (...) De nombreux jeunes français diplômés, notamment des écoles d'ingénieurs, s'expatrient pour y travailler. Dans les entreprises américaines, les services spécialisés en IA comptent ainsi de nombreux Français, soulignant certes la qualité de la formation française, mais aussi le manque d'attractivité de notre pays* ».

Enfin, la puissance technologique d'un État dépend de la capacité à former et retenir des talents, à savoir toute personne possédant des compétences rares qui peut contribuer au développement de l'IA que ce soit dans sa dimension technique, scientifique ou commerciale.

84 Commission pour l'Intelligence artificielle, IA : notre ambition pour la France, mars 2024.

PRÉCONISATION #12

Le CESE recommande aux pouvoirs publics et aux entreprises de former, retenir et attirer en France les talents avec la plus grande diversité de profils, pour favoriser l'innovation nationale et limiter les risques de biais, encourus lorsque les équipes affichent des profils trop similaires en :

- proposant une formation initiale et continue de haut niveau ;
- valorisant le statut des chercheurs notamment des doctorants et post-doctorants ;
- favorisant la mixité et la diversité des équipes ;
- réactualisant les repères de salaires existant dans les branches, au regard des niveaux élevés de qualifications nécessaires ;
- soutenant les innovations prenant en compte les limites environnementales et les impacts sociaux des solutions développées.

Recommandation des citoyennes et citoyens

L'Europe doit avoir un rôle crucial dans la course à l'IA, confrontée à la domination des États-Unis et de la Chine. L'IA offre des opportunités pour renforcer l'autonomie stratégique européenne et développer des solutions technologiques alignées sur les valeurs éthiques et environnementales du continent. Cependant, l'Europe est freinée par des infrastructures fragmentées, un sous-investissement dans la recherche et un retard dans la mise en œuvre de projets ambitieux. Le CESE préconise de renforcer la coopération entre les pays européens, d'investir massivement dans des technologies propres et durables, et de garantir une régulation harmonisée pour promouvoir une IA compétitive, éthique et respectueuse des droits sociaux et environnementaux.

II - Les impacts de l'IA dans la société

Dans cette deuxième partie, le CESE analyse les effets de l'IA dans la société, en particulier :

- les impacts de l'IA sur l'emploi et le marché du travail ;
- les enjeux d'intégration de l'IA dans l'entreprise ;
- le recours à l'IA dans les services publics,
- la frontière créée par l'éloignement numérique ;
- le déploiement de l'IA dans l'éducation ;
- l'intégration de l'IA dans la santé ;
- et la capacité de l'IA à générer de l'accessibilité, au travers de l'exemple du handicap.

A. Création, suppression, transformation : l'emploi et le marché du travail à l'ère de l'IA

Le lancement de ChatGPT, le 30 novembre 2022, a mis en lumière les progrès spectaculaires réalisés dans le domaine de l'IA. Cet outil, basé sur la technologie des grands modèles de langage, est un exemple parmi d'autres de l'IA générative qui a démontré une véritable capacité à automatiser des tâches diverses, telles que la rédaction, la création d'images ou la programmation. En raison de l'ampleur de son impact et de sa capacité à transformer des secteurs entiers, l'IA est désormais perçue comme une technologie générique comparable au moteur à combustion interne, l'électricité, l'ordinateur et l'Internet (les réseaux). Ces dernières années, les investissements privés pour financer la recherche et le développement de l'IA se sont multipliés, et les États, toujours plus soucieux de la productivité et de l'attractivité économique, prennent pleinement part à cette course à l'innovation, tout en tentant d'anticiper les risques associés à la pénétration de l'IA dans nos sociétés. La rapidité de son développement entraîne une prise de conscience à grande échelle

de la capacité de l'IA à transformer nos vies, nos économies et le monde du travail notamment, mais ses conséquences potentielles restent, à ce stade, très incertaines.

Dans cette partie, le CESE décrit les impacts quantitatifs et qualitatifs de l'IA sur l'emploi en s'appuyant en partie sur les conclusions de l'analyse de controverse menée par la commission travail Emploi du CESE.

1. Une évaluation quantitative des emplois touchés par l'IA

1.1 LA DIFFICILE ÉVALUATION DES CRÉATIONS ET SUPPRESSIONS D'EMPLOIS PAR L'IA

L'IA pourrait toucher toutes les activités et toutes les professions, et ainsi transformer l'ensemble du marché du travail. Selon le Forum économique mondial, 50 % des employeurs à travers le monde prévoient que l'IA favorisera la croissance de l'emploi dans les cinq prochaines années, contre 25% estimant

l'inverse⁸⁵. Comme l'a expliqué la sociologue Dominique Méda lors de son audition, de multiples études ont été menées ces dernières années afin de tenter de quantifier les évolutions associées à la pénétration de l'IA sur le marché du travail, et celles-ci aboutissent à des résultats extrêmement hétérogènes. À titre d'exemple, une analyse menée par Carl Benedikt Frey et Michael Osborne de l'université d'Oxford et publiée en 2013⁸⁶ concluait que l'automatisation des tâches, associée notamment au développement de l'IA, pourrait faire disparaître 47 % des emplois aux États-Unis durant les deux décennies à suivre. Cette prédiction ne s'est pas réalisée. Une autre étude menée conjointement par trois scientifiques du Département de recherche du Bureau international du Travail à Genève⁸⁷ projette que le potentiel de remplacement des tâches par l'IA est largement inférieur à celui de l'augmentation de l'activité, avec des estimations suggérant que seuls 2,3 % des emplois pourraient être entièrement automatisés à court terme, contre 13 % susceptibles d'être stimulés par l'IA.

En clair, au cours de la dernière décennie, les analyses ont eu tendance à mettre l'accent tantôt sur la substitution de la main d'œuvre - la « perspective des prophètes de

malheur »⁸⁸ - tantôt sur les gains d'efficacité et les nouveaux types de travail découlant de la mise en œuvre de l'IA - la « perspective de l'optimiste »⁸⁹. À l'évidence, il est extrêmement complexe d'anticiper les impacts de l'IA sur le marché du travail, surtout que ces prévisions dépendent des politiques publiques et d'entreprises mises en œuvre. Dans un récent rapport, publié en septembre 2023, le Fonds Monétaire International (FMI) résume la situation ainsi : « *l'évaluation de l'impact de l'IA sur l'emploi est complexe en raison de son évolution rapide, de l'incertitude quant à son intégration dans les processus de production et de l'évolution des perceptions sociétales. Compte tenu de la rapidité des progrès et de l'évolution des capacités des technologies basées sur l'IA, prévoir quels processus de production intégreront l'IA et quelles tâches humaines seront remplacées ou améliorées demeure incertain* »⁹⁰.

85 Forum économique mondial. Future of Jobs Report. 2023.

86 Frey et Osborne. The future of employment : how susceptible are jobs to computerisation ?, 2013.

87 Gmyrek et al. Generative AI and Jobs : a global analysis of potential effects on job, quantity and quality, Organisation mondiale du travail, 2023.

88 Audition de Bilel Benbouzid, maître de conférences spécialiste de la sociologie de l'action publique et de la sociologie digitale à l'université Gustave Eiffel, par la Commission temporaire IA, le 17 octobre 2024.

89 Frank et al. Toward understanding the impact of artificial intelligence on labor, 2019.

90 Cazzaniga et al. Gen-AI, Artificial Intelligence and the Future of Work, Fonds monétaire international, 2024.

1.2 UNE CERTITUDE : L'IA VA TRANSFORMER LE MARCHÉ DU TRAVAIL

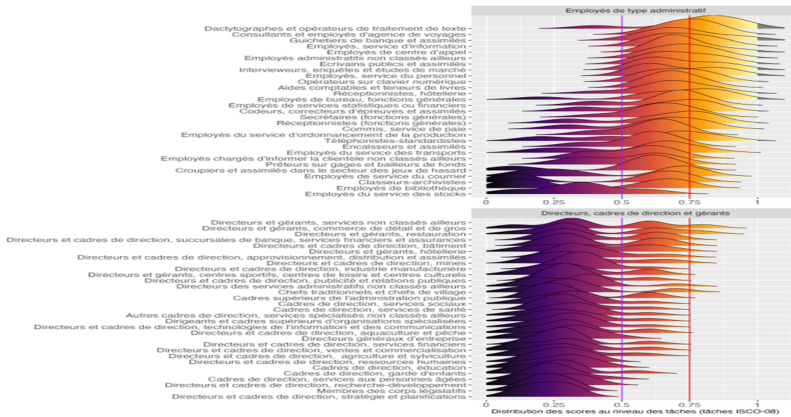
Certains emplois sont particulièrement automatisables, et vulnérables face à l'IA.

Malgré toutes ces incertitudes, la communauté scientifique s'accorde à dire que certains emplois sont plus facilement substituables par l'IA que d'autres. Contrairement aux précédentes vagues de développement technique qui ont affecté le marché de l'emploi, l'IA, puisqu'elle consiste en un enchaînement de choix

logiques et qu'elle est donc capable de prendre des décisions rationnelles, touche directement les métiers considérés comme intellectuels.

Plus en détail, de nombreuses analyses se concentrent sur l'exposition des professions à l'IA générative. Le FMI⁹¹ estime ainsi que les deux secteurs les plus exposés à l'IA sont « les services financiers et marchés » et « l'assurance et la gestion de l'épargne retraite ». Entre 32 % et 37 % des tâches seraient automatisables dans ces domaines.

GRAPHIQUE 2



Source : Gmyrek et al. op. cit.

Dans les deux schémas de l'OIT de cette partie, l'unité mesurée en abscisse est le niveau d'exposition à l'automatisation potentielle par l'IA générative ayant des capacités similaires à GPT4. L'étude considère que l'exposition est « moyenne » pour des scores de 0,5 à 0,75, et « forte » pour des scores

supérieurs à 0,75.

De son côté, la dernière étude de l'Organisation internationale du travail (OIT)⁹² à ce sujet part de la combinaison entre intelligence artificielle et jugement humain, et identifie les tâches – parmi celles des 436 professions normalisées au niveau international – qui pourraient

91 Ibid.

92 Gmyrek et al. op. cit.

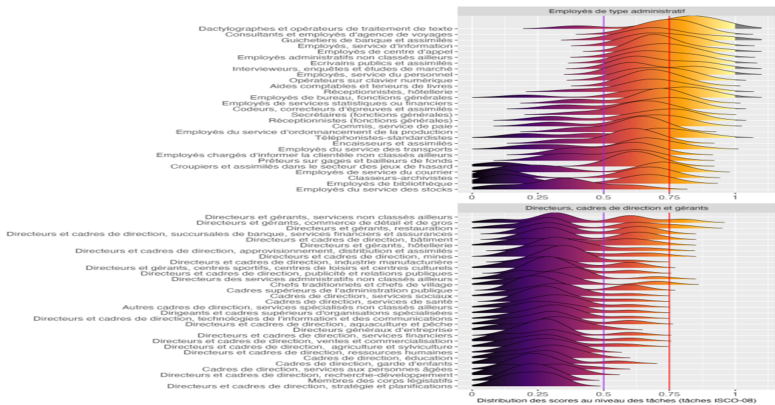
être affectées dans les années à venir par des technologies dotées de capacités similaires à celles de GPT-4⁹³. Ainsi, les emplois dont le degré de potentiel d'exposition présente une moyenne élevée et un faible écart-type⁹⁴ sont considérés à fort potentiel d'automatisation, car la majorité de leurs tâches typiques présentent des niveaux d'exposition élevés. Au global, les emplois occupés par les employés de type

administratif sont particulièrement automatisables, comme le suggère la figure ci-dessous extraite de l'étude l'OIT.

Logiquement, les emplois disposant d'un fort niveau d'exposition sont ceux les plus facilement substituables par l'IA.

D'autres métiers risquent fortement d'être transformés.

GRAPHIQUE 3



Source : Gmyrek et al. op. cit.

L'étude de l'OIT établit une autre catégorie d'emplois qui pourraient être particulièrement affectés par le développement de l'IA générative : les emplois à fort potentiel de transformation. Ces emplois comprennent bon nombre de tâches difficiles à automatiser, mais en contiennent aussi qui peuvent l'être facilement. Ceux-ci se situent donc à l'autre extrémité du spectre :

leur degré d'exposition moyen à l'automatisation est faible, mais avec un écart-type élevé. Les emplois occupés par les directeurs, les cadres de direction et les gérants répondent particulièrement à ces critères.

Pour ces emplois à fort potentiel de transformation, l'IA générative pourrait supprimer certaines des tâches, mais l'intervention humaine

93 À l'heure actuel, GPT4 est le dernier modèle de langage multimodal développé par la société Open IA et accessible au grand public. De nouveaux modèles, encore bien plus puissants, sont attendus prochainement.

94 Cet indicateur mesure la dispersion des valeurs.

resterait nécessaire pour la performance globale du poste. Il convient cependant d'anticiper que la multiplication des emplois à fort potentiel de transformation pourrait également se traduire par une réduction du nombre d'emploi total, l'automatisation partielle permettant à moins de travailleurs de produire une même quantité de travail. En clair, l'automatisation de certaines tâches pourrait induire des gains d'efficacité, ce qui pourrait affecter en profondeur l'activité professionnelle des salariés et la structure du marché du travail au sens large.

Certains groupes sociaux sont spécifiquement touchés

S'agissant des tâches considérées comme les plus facilement automatisables, différents groupes peuvent être identifiés comme particulièrement vulnérables face aux futures évolutions du marché de l'emploi impulsées par la pénétration

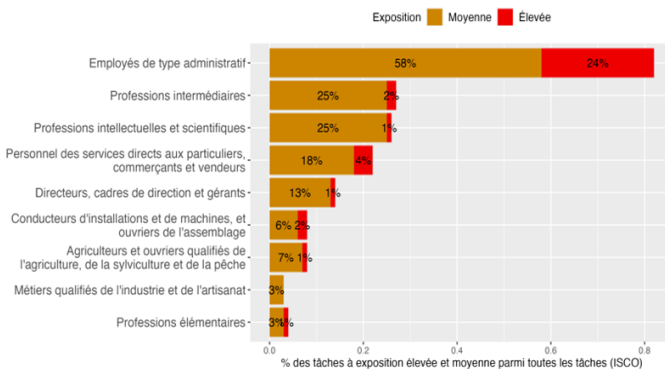
de l'IA. Les paragraphes qui suivent présentent certains de ces groupes sociaux, de manière non hiérarchisée et non exhaustive.

Les employés administratifs

Comme évoqué précédemment, l'étude de l'OIT révèle que les employés administratifs constituent le groupe professionnel le plus exposé à l'automatisation. Ainsi, la figure ci-dessous indique que 24 % des tâches exercées par ce groupe professionnel présentent un niveau d'« exposition élevée » à l'automatisation, tandis que 58 % des tâches sont considérées comme « moyennement exposées ». À l'inverse, les autres groupes professionnels sont moins affectés, avec seulement 1 à 4 % des tâches présentant un « potentiel d'automatisation élevé ». Les tâches « moyennement exposées » dans ces groupes ne dépassent pas 25 %.

GRAPHIQUE 4

Figure 1. Tâches moyennement et fortement exposées à l'IA générative, par groupe professionnel



Note : Grands groupes de professions (codes à un chiffre) Classification internationale type des professions (CITP-08). Niveaux d'exposition à l'automatisation potentielle par l'IA générative ayant des capacités similaires à GPT-4 sur une échelle de 0 à 1. "Exposition moyenne" pour les scores de 0,5 à 0,75 et "exposition forte" pour les scores supérieurs à 0,75.

Source : Gmyrek et al. op. cit.

• Les femmes

Des études récentes convergent pour dire que les femmes risquent davantage de voir leur emploi affecté à cause de l'intelligence artificielle, que les hommes. En effet, dans une étude publiée en juillet 2023, le cabinet de conseil McKinsey estime que les femmes ont 1,5 fois plus de risques de devoir changer de postes à cause de l'intelligence artificielle que les hommes. Selon *Pew Research*, 21 % des travailleuses risquent de perdre l'emploi à cause de l'IA, contre « seulement » 17 % des travailleurs. Enfin, la société Revelio Labs, spécialisée dans

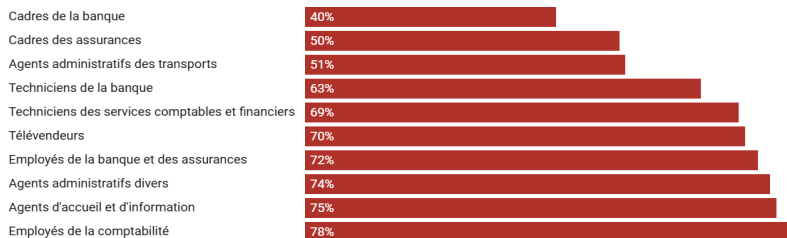
l'analyse de données dans les ressources humaines, estime que les quinze métiers les plus exposés à l'IA aux États-Unis sont en moyenne occupés à 71 % par des femmes⁹⁵, les emplois de bureaux, concernant principalement des femmes, étant les plus impactés.

Dans une récente note⁹⁶ réalisée pour l'INSEE, Antonin Bergeaud détaille les dix métiers les plus exposés à l'IA en France. En accord avec les différentes analyses mentionnées précédemment, la figure ci-dessous démontre que ces professions sont largement féminisées.

GRAPHIQUE 5 :

Les dix métiers les plus exposés à l'IA sont très féminisés

Parts de femmes dans chaque profession



La proportion de femmes par métier est estimée à partir de chiffres de la Sécurité Sociale de 2015

Source : AGEFI, op. cit.

• Les jeunes diplômés

Dans ce même document⁹⁷, Antonin Bergeaud pointe un autre groupe social qui pourrait particulièrement être pénalisé par la pénétration de l'IA dans notre système productif :

les jeunes diplômés. Alors que la prédiction peut paraître contre-intuitive, il estime habituellement que les jeunes générations sont généralement plus à l'aise que les anciennes sur les nouvelles technologies. Il s'agit là d'un

⁹⁵ Johann Corric, Les femmes, les jeunes et les financiers semblent les plus exposés à l'IA, AGEFI, 2024.

⁹⁶ Antonin Bergeaud, Exposition à l'intelligence artificielle générative et emploi : une application à la classification socio-professionnelle française, INSEE, 2024.

⁹⁷ *Ibid.*

préjugé identifié de longue date qui empêche la société de se mobiliser sur l'accompagnement et l'éducation des jeunes au numérique. Le chercheur met en exergue une explication plutôt évidente avant de monter en compétence, les stagiaires, alternants ou autres débutants commencent généralement par des tâches simples et répétitives. Ces tâches pourraient commencer à être remplacées par l'intelligence artificielle, et pourraient bien l'être encore davantage à l'avenir. Or, sans ce passage par des tâches répétitives ou l'acquisition de compétences et connaissances approfondies, les jeunes professionnels se retrouvent esclaves d'une technologie qu'ils ne peuvent maîtriser, non par incompétences numériques, mais par un manque de compétences métier.

- **Certains métiers sont « isolés » du développement de l'IA**

Selon Dominique Méda⁹⁸, certains métiers pourraient être spécifiquement « isolés » du développement de l'IA. Il s'agit notamment de tous les métiers du *care* – infirmière, aide-soignante, auxiliaire de vie ou encore aide à domicile – car ceux-ci, puisqu'ils visent à répondre aux besoins élémentaires des personnes qui ne sont pas autonomes, nécessitent des facultés humaines *a priori* irremplaçables. Notons pourtant que certaines sources contradictoires considèrent que ces métiers pourraient tout de même être touchés par le développement de l'IA, dans la continuité de l'émergence des chiens robots programmés pour prendre en charge certains soins médicaux. D'autres professions, dans le secteur du service notamment, impliquent presque obligatoirement un lien entre humains,

à l'image du métier de bibliothécaire. Certaines professions manuelles et artisanales, à l'instar du métier de cordonnier, devraient connaître un second souffle avec l'urgence écologique et être relativement peu influencées, ou du moins pas directement, par le développement de l'IA. Le domaine de la construction et celui de l'hôtellerie devraient également être particulièrement « isolés » de la pénétration de l'IA dans le marché du travail.

⁹⁸ Audition de Dominique Méda, sociologue et philosophe du travail, professeure des universités, membre de l'Institut de recherche interdisciplinaire en sciences sociales (IRISSO) de Paris Dauphine, par la Commission temporaire IA, le 24 octobre 2024.

ENCADRÉ 2 : ZOOM : FAIRE DE L'IA UN LEVIER POUR L'ÉGALITÉ

Les femmes ont progressivement été écartées des métiers du numérique, conduisant aujourd'hui à une surreprésentation des hommes parmi les équipes de conception des systèmes d'IA, ce qui contribue à produire des biais dans la conception des systèmes d'IA. Aujourd'hui, les stéréotypes jouent un rôle clé dans cette exclusion : dès l'enfance, les filles sont découragées d'aller vers les sciences et le numérique. Ce biais se traduit par une moindre orientation des filles vers les écoles d'ingénieurs et des performances scolaires influencées par la perception des exercices. Par ailleurs, le secteur numérique est marqué par une culture sexiste, avec 46 % des femmes témoignant de comportements discriminatoires, notamment dans les études supérieures.

Les biais dans les algorithmes affectent les carrières des femmes. En 2015, Amazon a écarté des CV féminins lors d'un recrutement automatisé, et les annonces d'emploi ciblent souvent les hommes pour des postes prestigieux. L'intelligence artificielle, en apprenant sur des données déséquilibrées, peut amplifier ces inégalités en favorisant des traits comportementaux masculins comme le *leadership*.

Cette exclusion des femmes du numérique représente un coût économique pour elles. D'abord l'IA peut aussi devenir un critère d'employabilité et affecter l'emploi des femmes. Par ailleurs, privées d'opportunités dans le secteur du numérique en forte croissance (+2,5 % par an), elles subissent davantage l'automatisation de leurs métiers, nécessitant une reconversion plus fréquente.

Plusieurs initiatives émergent pour une IA plus inclusive : le laboratoire de l'Égalité propose des bases de données représentatives et promeut la formation des femmes dans ce domaine. De même, des programmes comme ceux de *WomeninIA* ou l'UNESCO favorisent leur participation active.

L'IA peut être utilisée pour produire des statistiques sur les inégalités de représentation et pour détecter et corriger les biais dans les bases de données. L'intégration plus grande des femmes à la conception des IA et la sensibilisation des développeurs aux stéréotypes de genre sont essentielles pour produire des algorithmes moins biaisés. L'IA peut aussi réduire l'influence des biais humains dans les recrutements et les évolutions de carrière, en se basant sur des critères objectifs plutôt que des relations interpersonnelles souvent favorables aux hommes. Enfin, elle peut servir à identifier les inégalités (statistiques, visibilité accrue des femmes sur des plateformes comme Wikipédia) et à corriger les biais des bases de données.

Pour qu'elle soit un outil au service de l'égalité, il est essentiel de former les concepteurs d'algorithmes à la lutte contre les biais sexistes et d'intégrer davantage de femmes dans leur développement.

Toutefois l'IA peut constituer un levier pour l'égalité si la volonté politique derrière sa conception s'inscrit dans cette perspective. Comme l'indiquait Flora Vincent pendant son audition, produire une IA moins biaisée que l'humain est techniquement possible, à condition de vouloir « encoder l'égalité ».

2. Une évaluation qualitative des évolutions du travail liées à l'IA

L'impact de l'IA n'est pas seulement dans la transformation des emplois mais davantage sur le travail lui-même et l'organisation du travail. En effet, la valorisation du « hors IA » devrait reposer sur l'apport du travail humain, de précision, de savoir-faire, de contact, de créativité, de pilotage, qui nécessite une totale recomposition de la valeur ajoutée, de la hiérarchisation des emplois et de l'organisation du travail.

La partie qui suit recense, de manière large mais non exhaustive, une série d'enjeux et de conséquences sur les manières de travailler associée à l'intégration de l'IA dans notre modèle productif.

2.1 L'IA PEUT AVOIR DES CONSÉQUENCES POSITIVES POUR LES TRAVAILLEURS⁹⁹

Utilisée en synergie avec les compétences humaines, l'intelligence artificielle pourrait être bénéfique pour les travailleurs, en allant dans le sens d'un accroissement de leur bien-être, au travers notamment de :

→ **la facilitation des tâches complexes** : le recours à l'intelligence artificielle peut faciliter la réalisation de tâches complexes (analyse de données, planification logistique, diagnostic

médical) et ainsi permettre aux travailleurs de gagner du temps et de l'énergie ;

→ **la délégation des activités à faible valeur ajoutée** : certaines activités à faible valeur ajoutée peuvent être déléguées aux systèmes d'intelligence artificielle, permettant aux travailleurs de se concentrer sur des tâches plus valorisées. L'IA peut ainsi accroître la productivité des organisations en minimisant ou en éliminant les tâches routinières ou répétitives ;

→ **la rationalisation des processus** : l'intelligence artificielle peut permettre de rationaliser certains processus ;

→ **la sécurisation de l'activité de travail** : l'intelligence artificielle peut sécuriser l'activité de travail par la vérification et la confirmation, à condition qu'elle soit au service des travailleurs qui doivent rester maîtres des prises de décision. Elle augmenterait alors la capacité des travailleurs à assumer la responsabilité tout en réduisant les risques d'erreurs, soulageant ainsi la charge mentale associée. Elle peut aussi renforcer la sécurité physique des travailleurs, et ainsi constituer un atout majeur pour le bien-être au travail ;

→ **l'optimisation de l'efficacité par l'IA** : l'intelligence artificielle (IA) et les algorithmes jouent déjà un rôle crucial

99 PEGA, Tout savoir sur la complexité de la main-d'œuvre moderne, 2022.

dans l'optimisation de l'efficacité, notamment dans la gestion de la main-d'œuvre. Les employeurs explorent ces applications dans des environnements variés tels que les entrepôts, la distribution et la livraison. Par ailleurs, l'IA est également utilisée pour la gestion des affectations des travailleurs. Il est impératif que cette utilisation se fasse de manière responsable, en respectant le libre choix des travailleurs et en évitant des charges de travail excessives ;

- **la diversification des savoir-faire** : l'intelligence artificielle peut développer les capacités d'action des travailleurs et stimuler l'actualisation et la diversification des savoir-faire. Dans ce cadre, l'intégration des systèmes d'intelligence artificielle dans les processus productifs pourrait conduire à la (re)valorisation de l'expertise humaine lorsque l'IA agit en complémentarité avec le travail humain.

2.2 L'IA PEUT TOUTEFOIS DÉGRADER L'EXPÉRIENCE DE TRAVAIL.

Une étude publiée en 2022¹⁰⁰ identifie sept dimensions de l'expérience de travail potentiellement affectées par les systèmes d'intelligence artificielle.

- **La fragilisation de la reconnaissance** : La valeur associée au travail, c'est-à-dire la reconnaissance matérielle et symbolique éprouvée par les travailleurs et la société au sens large pourrait évoluer, et dans

certain cas se dégrader. La reconnaissance est un moteur essentiel de l'engagement humain dans le travail, qu'il s'agisse de valorisation de la production, des revendications sociales et matérielles (salariales, statutaires) ou symboliques (respect, dignité, estime, mérite). La reconnaissance au travail engage l'identité sociale et le sentiment de fierté du travailleur. Elle prend différentes formes de valorisation ou de dévalorisation spécifiques à chaque métier ou génériques au sein des normes sociales. La reconnaissance peut être « existentielle » (liée à l'individu), orientée sur la « pratique », les « efforts consentis » ou les « résultats effectifs observables, mesurables et contrôlables ». L'introduction de l'intelligence artificielle dans les organisations transforme les manières de faire, les efforts à réaliser et la cause des résultats obtenus, ce qui peut modifier les modes de reconnaissance institués et créer de nouvelles luttes de reconnaissance. Le déploiement d'une IA dans une équipe de travail peut entraîner une dévaluation de la reconnaissance, affectant la singularité de l'individu, ses pratiques, ses efforts et les résultats qu'il obtient. Cela peut avoir des conséquences sur la motivation, l'engagement et la satisfaction au travail des individus.

- **Le sentiment de dépossession et de perte de compétence** : un risque de dévalorisation des savoirs et des savoir-faire existe,

¹⁰⁰ Ferguson et al., L'IA au travail : propositions pour outiller la confiance, Conférence Nationale sur les Applications Pratiques de l'IA, 2022.

pouvant mener à une perte d'estime de soi. Les savoir-faire contribuent à construire des identités professionnelles à partir desquelles s'articulent légitimité, positions et statuts. Avec l'introduction des systèmes d'intelligence artificielle, les savoir-faire évoluent, ce qui peut ralentir ou accélérer l'innovation de processus par la rationalisation des modes opératoires ou la mise en place de la flexibilité et de la polyvalence du travail. La perte de compétences et la régression des savoir-faire nécessaires à l'activité sont des risques redoutés, tout comme l'abaissement du niveau d'expertise nécessaire à l'exécution des tâches. La standardisation et l'harmonisation du travail induites par l'IA constituent donc des risques de déqualification et de perte de compétences. Les agents les plus novices sont particulièrement vulnérables, car ils n'ont pas encore acquis les habitudes du métier. Les pannes ou bugs de réseau, qui suppriment temporairement toute intervention de l'IA, confrontent le technicien à sa seule expertise et expérience. L'automatisation de la recherche documentaire pourrait également marginaliser le travail de mise à jour des connaissances, entraînant une perte de compétence.

- **Le désengagement relationnel :** l'introduction de systèmes d'intelligence artificielle peut contribuer à la désorganisation et à la fragilisation du travail collectif et des collectifs de travail. Cette situation peut déshumaniser les pratiques et appauvrir les interactions sociales, entraînant un risque de désengagement relationnel.
- **L'intensification du travail :** L'adoption

de l'IA peut intensifier le travail en augmentant les attentes de productivité et en accélérant les rythmes de travail. Les employés peuvent être soumis à une pression constante pour maintenir des niveaux de performance élevés, souvent sans pause adéquate. L'automatisation des tâches banales par l'IA peut paradoxalement rendre les emplois restants plus exigeants, nécessitant une vigilance accrue et un effort mental supplémentaire. Cette intensification peut entraîner une augmentation du stress, une fatigue plus rapide et une diminution de la satisfaction au travail, impactant négativement la santé mentale et physique des travailleurs.

- **Le contrôle et la surveillance :** les chevauchements des IA avec le travail humain peuvent parfois dissimuler des « *systèmes de surveillance et de contrôle* »¹⁰¹ de l'activité de travail. En dehors des obligations de surveillance inhérentes à l'employeur pour protéger les installations et les personnes, le droit de surveillance est « *soumis à un contrôle de proportionnalité* » et doit être « *justifiable par les intérêts légitimes de l'employeur* ». Cependant, les IA modernes « *mobilisent des algorithmes apprenant à partir de données collectées* » qui peuvent concerner directement les travailleurs et être exploités pour évaluer leur productivité. Les outils de « *gestion algorithmique* » peuvent s'appuyer sur divers dispositifs de mesure de l'activité, assurant ainsi la micro-gestion du travail avec un contrôle étroit des temps et des performances. Les systèmes de planification automatisés transfèrent, par exemple, davantage de pouvoir des travailleurs aux gestionnaires.

¹⁰¹ Ferguson et al., L'IA au travail : propositions pour outiller la confiance, Conférence Nationale sur les Applications Pratiques de l'IA, 2022.

- **La réduction de l'espace des pratiques et perte d'autonomie** : les travailleurs peuvent ressentir une dégradation de leur flexibilité cognitive face aux systèmes automatisés. Le déploiement de l'IA dans les organisations de travail risque de réduire l'« espace des pratiques au profit de process », diminuant ainsi la capacité d'agir des individus, qui peuvent avoir l'impression d'être réduits à suivre des instructions sans aucune finalité. Cependant, cette domination ne découle pas uniquement de formes actives, mais aussi de la passivité des travailleurs dans leur interaction avec le système. L'« excès de confiance », l'« effet de contentement » (se contenter d'une solution relativement satisfaisante obtenue sans effort) et l'« excès de prudence »¹⁰² peuvent désengager les travailleurs, réduire leur expertise et augmenter leur dépendance vis-à-vis du système mais risquent aussi d'entraîner une baisse généralisée de capacité à faire preuve de créativité ou à résoudre des problèmes. L'automatisation de certaines tâches peut diminuer la sensibilité sociale associée à l'expérience de travail et développer un sentiment de déclassement.
- **La déresponsabilisation** : la gestion d'une partie du travail par l'IA peut créer une barrière entre le travailleur et l'organisation, ainsi qu'entre le travailleur et la direction, semblant dégager

l'entreprise de ses responsabilités. Cela supprime les moyens par lesquels les travailleurs comprennent et exigent que ces responsabilités soient assumées¹⁰³. Le risque de déresponsabilisation des travailleurs ou de délégation passive à l'IA est élevé, ce qui peut entraîner une perte de sensibilité et de compétence dans l'exercice du travail. La faible maîtrise des sujets, le manque de compétences, la faible qualité relationnelle avec le prestataire de la solution technologique, et le manque d'implication des travailleurs peuvent conduire à un effet « boîte noire », remettant en question le sentiment de responsabilité au travail. Il est crucial de comprendre et d'auditer les tâches réalisées par l'intelligence artificielle pour éviter de s'en remettre aveuglément à ces outils.

102 J. M. Hoc, From Human-Machine Interaction to Human-Machine Cooperation, 2000.

103 Joyce et al. Trade union responses to platform work: An evolving tension between mainstream and grassroots approaches, 2021.

Au-delà de la santé au travail, les effets négatifs relatés ici peuvent dégrader le rôle du travail dans notre société, et en

particulier dans ses dimensions sociales, politiques et identitaires du travail.

ENCADRÉ 3 : ZOOM SUR LES TRAVAILLEURS DU « CLIC »

Il est important de noter que les « travailleurs du clic » jouent un rôle crucial dans les processus. Ils interviennent dans les phases de préparation, de validation, d'amélioration des données (nettoyage, correction des erreurs). Ils étiquettent, labellisent les images, les vidéos pour entraîner les algorithmes supervisés. Ils annotent les textes pour classifier les sentiments, les émotions, pour la traduction automatique. Ils ont un rôle dans le test de performance des modèles, par exemple en testant sur des entrées spécifiques et bien maîtrisées, et en évaluant les sorties générées. Ils peuvent également vérifier pour leurs donneurs d'ordre les résultats obtenus dans un but d'amélioration des données d'entrée. La création de données synthétiques peut leur être confiée. Ce sont « les petites mains » de l'IA, aujourd'hui aussi indispensables qu'invisibles. Ils vivent de façon précaire sans législation protectrice, exploités par les grands groupes de l'IA, souvent dans les pays en voie de développement (Inde, Philippines, Afrique de l'Est, Madagascar, Venezuela)¹⁰⁴.

Les préconisations du CESE

L'exercice ci-dessus vise à établir un panorama des évolutions potentielles du monde du travail.

• Principes préalables à l'intégration de l'IA dans le monde du travail

Faire prévaloir le principe de prudence doit donc être une nécessité : les études d'impact antérieures à l'adoption de système d'intelligence artificielle par les entreprises doivent être robustes, les travailleurs et les instances représentatives du personnel doivent être associés avant que ces technologies soient adoptées et ils doivent être accompagnés dans leur utilisation si elles le sont.

Des gains en efficacité associés au développement de l'IA sont promis par

leurs promoteurs. C'est un intérêt central de l'intelligence artificielle Dans le système économique que nous connaissons, l'IA risque d'être un outil utilisé pour produire davantage, en risquant d'accélérer le rythme de consommation des ressources disponibles. Il s'agit donc, avant tout, d'interroger les finalités de l'IA, pour la façonner de manière à ce qu'elle nous aide à produire mieux, en adéquation avec nos besoins réels, dans le cadre d'un réel progrès social.

¹⁰⁴ L'ADN, L'IA a besoin de millions de travailleurs payés quelques centimes, octobre 2023.

PRÉCONISATION #13

Le CESE préconise de la prudence avant d'adopter un système d'IA (réalisation d'études d'impacts, association des salariés ou de leurs représentants), et d'en faire une base pour le dialogue social sur l'IA au sein des instances représentatives du personnel.

nouveau paysage technologique, en faisant évoluer leurs compétences pour non seulement s'adapter, mais aussi prospérer dans ce nouveau monde économique qui va se transformer par l'IA.

Le dialogue social joue un rôle crucial dans le déploiement de l'IA en entreprise.

- **L'IA au service des travailleurs**

L'intégration de l'IA dans le milieu de travail peut paraître inévitable et constitue les entreprises (qu'elles soient publiques, privées à but lucratif ou non lucratif), leurs dirigeants et leurs salariés un défi majeur pour leur pérennité. L'investissement dans l'évolution des compétences à travers la formation professionnelle est crucial pour éviter les suppressions d'emplois et pour transformer l'IA en un catalyseur de croissance économique. En mettant l'accent sur la formation continue, les partenariats efficaces, la valorisation des compétences humaines, un dialogue social inclusif, de bonnes conditions de travail et une gestion du changement efficace, nous pouvons assurer une transition réfléchie vers un avenir où l'IA et les travailleurs peuvent coexister en synergie, créant ainsi un environnement de travail plus productif et épanouissant. Il est impératif que les travailleurs et les dirigeants d'entreprises soient équipés pour naviguer dans ce

PRÉCONISATION #14

Le CESE recommande aux organisations syndicales et patronales de définir et négocier dans un accord national interprofessionnel (ANI), sur le déploiement de l'IA dans les entreprises et la manière de façonner l'IA pour produire mieux, en adéquation avec les besoins identifiés.

L'accord devra permettre aux branches professionnelles et aux entreprises de négocier sur les modalités prédéfinies par l'ANI et en particulier :

- d'impliquer les travailleurs dans la conception des systèmes d'IA pour s'assurer qu'ils répondent aux besoins réels et qu'ils créent des opportunités d'amélioration ;**
- d'intégrer les enjeux de mixité et de diversité depuis la phase de conception jusque dans la finalité des usages ;**
- d'assurer un dialogue social continu pour aborder les préoccupations et trouver des solutions adaptées.**

La transparence dans l'utilisation de l'IA ainsi qu'une communication claire des objectifs et des impacts pourront également être définis. Cette nécessité de dialogue social doit aussi s'appliquer dans les trois versants de la fonction publique.

De façon générale lors de l'introduction de l'IA les effets sur les conditions de travail doivent être soigneusement surveillés et adaptés notamment dans le document unique d'évaluation des risques professionnels (DUERP) pour que tous les algorithmes d'IA développés et programmés dans l'entreprise respectent bien une démarche partagée de prévention des risques professionnels et les conditions de travail, la sécurité, la santé et le bien-être des salariés.

PRÉCONISATION #15

Dans le cadre du déploiement de l'IA, le CESE appelle au respect des conditions de travail, et notamment à :

- **utiliser l'IA pour améliorer la sécurité au travail en identifiant et atténuant les risques potentiels ;**
- **prendre en compte la santé et le bien-être des travailleurs pour éviter le surmenage et le stress ;**
- **maintenir l'autonomie des travailleurs et leur capacité à prendre des décisions éclairées, malgré l'intégration des systèmes d'IA ;**
- **intégrer la prévention des risques professionnels liés à l'IA dans le DUERP et dans le suivi en Commission santé, sécurité et conditions de travail (CSSCT).**

- **Formation professionnelle**

L'IA redéfinit les dynamiques du travail à une vitesse sans précédent. L'automatisation des tâches autrefois réalisées par des humains menace la pérennité des emplois traditionnels, suscitant des craintes légitimes concernant les suppressions d'emplois. Cependant, une des clés pour limiter les risques associés au développement de l'IA dans le monde du travail réside dans l'investissement massif dans la formation professionnelle.

L'évolution des compétences des travailleurs est essentielle pour les préparer à interagir efficacement avec ces nouvelles technologies et pour garantir leur employabilité à long terme, d'autant plus si nous considérons que l'intégration de l'IA dans le milieu de travail est inévitable et constitue un défi majeur pour la pérennité des emplois traditionnels. Mettre l'accent sur la formation continue, les partenariats efficaces et la valorisation des compétences humaines, est essentiel pour anticiper et limiter les risques associés à cette rupture technologique.

Pour mettre en œuvre ces formations, il est indispensable d'accompagner les salariés en apportant des réponses à leurs interrogations légitimes. La peur de l'inconnu et la perte d'emploi peuvent générer des résistances à l'adoption de l'IA, freinant ainsi les initiatives de formation.

La gestion du changement est essentielle pour assurer une transition en douceur vers l'utilisation de l'IA.

PRÉCONISATION #16

Le CESE recommande d'anticiper la gestion du changement au cœur du suivi managérial, en :

- instruisant une démarche participative et consultative pour informer les travailleurs, en toute transparence, des avantages et des inconvénients de l'IA ;
- planifiant le changement avec des étapes claires et des objectifs mesurables ;
- favorisant, à travers le dialogue social et la définition d'un calendrier de déploiement, une intégration progressive de l'IA dans les processus de travail, permettant aux salariés de s'habituer aux nouvelles technologies sans pression excessive ;
- engageant les parties prenantes dès le début et tout au long du processus ;
- formant de manière adéquate ;
- assurant le support continu pour aider les employés à s'adapter ;
- communiquant de manière transparente et fréquente sur les raisons du changement, les bénéfices attendus et les progrès réalisés ;
- mesurant les résultats de l'introduction de l'IA de manière objective. Des indicateurs de performance peuvent être définis pour évaluer l'impact de l'IA sur l'efficacité opérationnelle, la satisfaction des employés, et la performance organisationnelle. Une évaluation régulière permet d'identifier les succès et les domaines nécessitant des améliorations ;
- engageant les dirigeants dans le suivi des progrès, et l'évaluation des résultats ;
- ajustant la stratégie en fonction des retours d'expérience. Il est important de rester flexible et de s'adapter aux évolutions technologiques et aux besoins de l'entreprise.

Par ailleurs, l'IA fait courir différents risques pour les travailleurs de :

- **déqualification** : la vitesse d'évolution de l'IA dépasse souvent la capacité des travailleurs à s'adapter, risquant de rendre obsolètes leurs compétences actuelles ;
- **disparités de compétences** : l'adoption de l'IA peut accentuer les disparités entre les travailleurs qualifiés et non qualifiés, renforçant les inégalités sur le marché du travail ;
- **résistance au changement** : la peur de l'inconnu et de la perte d'emploi peut générer une résistance à l'adoption de l'IA, freinant ainsi les initiatives de formation ;
- **conflits de rationalité** entre précision / efficacité, et entre subordination/responsabilisation ;
- **responsabilité d'une décision prise avec le support de l'IA au regard des vides juridiques existants** (qui a fait l'erreur ?).

Pour accompagner au mieux les travailleurs et pour éviter les suppressions d'emploi, il est impératif de se concentrer sur l'évolution des compétences à travers une formation professionnelle adaptée. Elle doit permettre aux travailleurs de faire évoluer leurs compétences afin qu'ils puissent non seulement s'adapter aux changements, mais aussi en tirer parti pour se repositionner et saisir de nouvelles opportunités professionnelles. Le succès de l'introduction de l'IA dans une organisation dépend en grande partie de la formation et de l'accompagnement des employés. L'investissement de l'entreprise dans des programmes de formation pour développer les compétences techniques et numériques des employés est un prérequis au déploiement de l'IA.

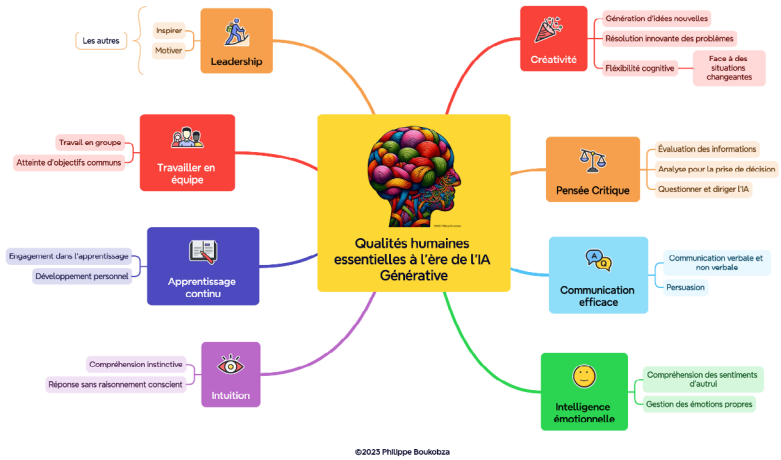
PRÉCONISATION #17

Le CESE appelle à intégrer la formation à l'IA dans les entreprises (publiques, privées à but lucratif ou non lucratif), pour doter les travailleurs en connaissances et compétences nécessaires pour utiliser et coexister avec les technologies de l'IA, en :

- proposant systématiquement un plan de développement des compétences des salariés, concomitamment au déploiement d'un système d'IA en entreprise ;**
- développant des programmes de formation continue en collaboration avec les gouvernements, les entreprises et les institutions éducatives, incluant des sessions sur les fondamentaux de l'IA, l'apprentissage machine, la gestion et la protection des données, et les aspects de sécurité et d'éthique liés à l'utilisation de l'IA, favorisant ainsi une application responsable des technologies ;**
- incitant lors des négociations autour de la Gestion prévisionnelle des emplois et des compétences (GPEC) et de la Gestion des emplois et des parcours professionnels (GEPP) à la prise en compte de cet outil.**

Le CESE suggère que le congé de formation obligatoire pour tous les membres élus titulaires et suppléants des Comités sociaux et économiques (CSE), Comités sociaux d'administration (CSA) et Comités sociaux territoriaux (CST), notamment ceux de la Commission santé, sécurité et conditions de travail (CSSCT) soit étendu, afin d'intégrer les connaissances et compétences nouvelles sur l'IA dans toutes les entreprises d'au moins 300 salariés.

FIGURE 2



Source : Philippe Boukobza

PRÉCONISATION #18

Le CESE appelle les entreprises à valoriser la complémentarité entre les compétences humaines et les systèmes d'IA, en :

- **mettant en place des programmes de mentorat où les experts en IA peuvent guider les autres travailleurs dans l'acquisition de nouvelles compétences ;**
- **promouvant des rôles où l'expertise humaine est complémentaire à l'IA, tels que les analyses complexes nécessitant un jugement humain ;**
- **formant les travailleurs au déploiement d'une application responsable des technologies.**

Recommandation de citoyennes et citoyens

L'impact de l'IA sur l'emploi. D'un côté, elle peut améliorer l'efficacité des entreprises, créer de nouveaux emplois dans la technologie et faciliter les tâches répétitives. Mais de l'autre, elle risque de remplacer certains emplois, notamment ceux qui sont peu qualifiés, et de créer des inégalités sur le marché du travail. Le CESE recommande de mettre en place des formations pour préparer les travailleurs aux nouveaux métiers, et de favoriser l'intégration de l'IA de manière à ne pas laisser de côté les travailleurs les plus vulnérables.

B. L'intégration de l'IA en entreprise

1. Les entreprises ne sont pas toutes en capacité d'intégrer l'IA

1.1 L'ADOPTION DE L'IA DIFFÈRE BEAUCOUP SELON LES ENTREPRISES

Seulement 13 % des TPE-PME ont recours à des solutions d'intelligence artificielle, selon le baromètre France Num 2024¹⁰⁵. Toutefois, ce chiffre est en augmentation depuis l'an dernier (en 2023, il était de 5 %). Les taux d'utilisation les plus importants sont constatés dans les secteurs du numérique (40 %) et des services spécialisés et techniques (bureaux d'étude, ingénierie, architectes, avocats, etc.) soit 22 %. En revanche, ils sont les plus faibles dans l'industrie agro-

alimentaire (6 %). En pratique, les usages sont dominés par l'intelligence artificielle générative (10 %), suivie des *chatbots* et assistants (5 %). Les solutions permettant l'automatisation de tâches, l'analyse des données ou de documents (chacune à 3 %), ou le contrôle qualité et la détection des anomalies (1 %) sont moins répandues chez les TPE-PME. Cependant, les organisations professionnelles représentatives de secteurs d'activités composés principalement de TPE se mobilisent pour sensibiliser et informer leurs membres sur les opportunités, comme sur les risques, liés à l'IA.

Dans les startups¹⁰⁶, l'intégration de l'IA est plus engagée. 67 % d'entre elles utilisent l'IA dans leurs produits,

105 Baromètre France Num 2024 : perception et usages du numérique par les TPE et PME.

106 Baromètre France Digitale - EY sur la performance économique et sociale des startups et fonds de capital-risque, 2024

43 % dans leurs outils marketing et commerciaux, 37 % dans leurs outils de communication, 30 % pour le service client et service après-vente, et 13% dans leurs outils de ressources humaines. Pour les startups, l'IA est vue comme un moyen d'optimiser les tâches (44 %), une opportunité de gagner en compétitivité face à des concurrents (27 %), une possibilité de créer de nouveaux services (22 %) et, pour 6 % d'entre elles, une solution pour compenser une réduction d'effectifs.

Il convient de noter que l'IA touche toutes les entreprises, les entreprises publiques, les entreprises privées à but lucratif mais aussi celles à but non-lucratif ou à lucrativité limitée. Les inégalités devant l'IA en fonction des ressources de ces structures sont également prégnantes et son intégration doit faire l'objet d'une attention tout aussi stricte.

Cela passera nécessairement par une transparence renforcée comme recommandé par les projets SeCoIA Deal¹⁰⁷ et Dial-IA¹⁰⁸.

1.2 L'ADOPTION DE L'IA EN ENTREPRISE ÉVOLUE DEPUIS 10 ANS

Théo Alves Da Costa¹⁰⁹ rappelle qu'il y a dix ans, les grandes entreprises et les startups ont pris conscience de l'opportunité que pouvait présenter l'IA, et ont commencé à intégrer des systèmes d'IA sous des formats de tests expérimentaux, appelés des *Proof of concept* (POC). Les grands groupes ont

ainsi développé des laboratoires d'IA, des incubateurs pour tester en interne des solutions d'IA sur certains produits, services ou outils. Ils ont commencé à recruter des développeurs, des chefs de projets IA. Mais en pratique, le retour sur ces expérimentations a été assez mitigé, pour plusieurs raisons, notamment : manque de coordination globale dans la stratégie, peu de réflexion sur les usages attendus de la technologie pour les utilisateurs finaux (les clients, les collaborateurs), manque de données disponibles en interne pour véritablement déployer l'IA.

Depuis cinq ans, une certaine rationalité méthodologique est déployée dans les grands groupes pour mieux intégrer l'IA. Une attention particulière est désormais portée sur l'utilité concrète de l'IA (ses usages, son impact potentiel sur les utilisateurs finaux) avant de prêter attention à la technologie. En interne, le déploiement de l'IA est désormais réparti dans plusieurs métiers : les métiers technologiques (développeurs tech, fonctions supports type maintenance cyber, direction des systèmes d'information), mais aussi les métiers produits, relations clients, vente, services internes (RH, administratif, financier, juridique, etc.). Désormais, les grands groupes sont moins dans des stratégies d'expérimentation, qu'à la recherche de solutions fiables, ultra-innovantes, et pouvant être déployées sur-mesure dans l'entreprise.

¹⁰⁷ SeCoIA Deal (SErvenir la COnfiance dans l'Intelligence Artificielle par le Dialogue).

¹⁰⁸ Dial-IA (Dialoguer sur l'IA).

¹⁰⁹ Audition de Théo Alves Da Costa, directeur de l'unité IA & transition écologique d'Ekimetrics et co-fondateur de l'ONG Data for Good, par la Commission temporaire IA, le 26 septembre 2024.

Depuis deux ans, l'arrivée de ChatGPT contribue à démocratiser l'IA auprès des TPE-PME, permettant notamment la diffusion de l'IA générative dans les fonctions supports (service client, services juridiques, etc.). Toutefois, les mêmes erreurs tendent à être reproduites : beaucoup d'organisations, appâtées par la technologie, ont voulu la déployer sans réfléchir aux usages, contribuant à une certaine déception face à la technologie. L'intégration de l'IA doit être préparée, réfléchi, et intégrée dans un processus global de digitalisation pour être véritablement bénéfique aux entreprises. Nous y reviendrons un peu plus tard.

1.3 QUELS SONT LES BÉNÉFICES ATTENDUS DE L'IA POUR LES ENTREPRISES ?

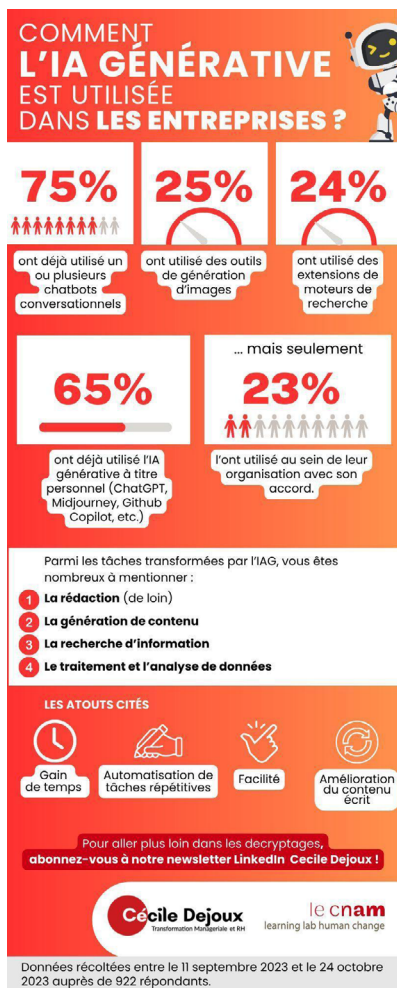
Il existe de multiples exemples des bénéfices possibles d'une intégration de l'IA dans les entreprises, notamment : gain de productivité sur les tâches à faible valeur ajoutée, notamment grâce à l'automatisation de tâches routinières et répétitives. L'IA peut ainsi libérer du temps pour les salariés, leur permettant de se concentrer sur des tâches plus complexes

et créatives. Cela peut améliorer l'efficacité et la qualité du travail, tout en contribuant à une meilleure satisfaction des collaborateurs.

aide à la prise de décision, grâce au traitement et l'analyse de données en temps réel. Bien utilisée, dans le respect des salariés, l'IA peut conduire à une meilleure performance organisationnelle et à une compétitivité accrue.

amélioration de la relation client : les chatbots, par exemple, peuvent également gérer efficacement un grand volume de requêtes clients 24h/24, à condition qu'ils soient bien entraînés pour ne pas produire d'erreurs ou dégrader l'expérience client et que soit systématiquement prévue la possibilité de recours à un humain en cas d'inadéquation de la réponse fournie par l'IA. L'analyse de données peut aussi permettre une personnalisation accrue des offres, répondant plus précisément aux besoins des consommateurs. L'IA permet aux entreprises de créer des expériences client plus engageantes, personnalisées et efficaces, ce qui peut offrir un atout majeur dans le paysage commercial concurrentiel actuel.

FIGURE 3



Étude réalisée en 2023 par Cécile Dejoux^{*10}

Pour autant, au regard des bénéfices attendus/réels de l'automatisation et du numérique, on se rend compte que le report sur des tâches moins pénibles et plus qualifiantes

a été relativement faible, il faut donc observer une plus grande vigilance sur les reconversions et requalifications des métiers existants.

^{*10} Cécile Dejoux, données récoltées entre le 11/09/2023 et le 24/10/2023 auprès de 922 répondants.

Dans le secteur agricole, il y a de nombreux exemples d'utilisation de l'IA : l'Institut national de recherche pour l'agriculture, l'alimentation et l'environnement (INRAE), à Clermont-Ferrand, développe des robots agricoles autonomes qui utilisent des algorithmes pour s'adapter en temps réel aux variations environnementales (météo, croissance des plantes), améliorer l'efficacité et réduire les interventions humaines. Le groupe Zekat, avec son projet Ortikat, travaille sur des outils d'aide à la décision pour les engins agricoles, grâce à des boîtiers qui recueillent des données, analysent les performances et proposent des optimisations pour une meilleure efficacité des machines. Kermap, basé à Rennes, se concentre sur le traitement d'images satellites pour fournir des informations détaillées sur les cultures, la couverture des sols et les rotations, contribuant ainsi à des pratiques agricoles plus durables.

Certains considèrent toutefois que ce processus de numérisation de l'agriculture peut avoir pour effet d'éloigner les agriculteurs des savoir-faire paysans basés sur la connaissance empirique de la terre et des écosystèmes, pourtant nécessaires dans le cadre de la transition agroécologique. Notons également que les coûts associés à l'adoption d'outils numériques de plus en plus poussés sont tels qu'ils ne peuvent être absorbés que par les fermes disposant d'un chiffre d'affaires élevé, ce qui risque d'amplifier les inégalités entre les exploitations et de freiner les installations.

Dans son audition auprès du CESE¹¹¹, la SNCF a détaillé des exemples concrets d'utilisation de l'IA au sein du groupe :

- création d'outils de maintenance prédictive grâce à la détection de l'usure de certaines pièces (portes, climatisation, surveillance de liquides, échauffement, etc.) ou le suivi de l'approvisionnement en pièces (gestion des stocks, planification des opérations) permettant une réduction des ruptures de prise en charge, et rendant possible un objectif de zéro panne en 2030 ;
- amélioration du service, au travers, par exemple, d'une optimisation des plans de transport en gérant de multiples contraintes, d'une supervision et gestion des voies, par l'analyse vidéos, photos, ou de prédictions trafics ;
- renforcement de la sécurité, au travers de la vidéosurveillance sur les voies ;
- amélioration des process internes, en simplifiant l'accès à la documentation interne et la documentation technique (pour les développeurs) ;
- optimisation des coûts de production, notamment des factures d'énergie, grâce à la mise en œuvre d'outil d'éco-conduite ;
- amélioration de la relation client / entreprise, au travers du lancement d'une borne interactive pour informer les voyageurs en plusieurs langues et en temps réel.

Cependant, en pratique, il est très difficile pour les entreprises d'évaluer concrètement les gains de productivité (aussi bien le chiffre d'affaires supplémentaire généré ou le montant des ressources économisées) ou le pourcentage de productivité humaine grâce au recours à l'IA. En effet, la SNCF a expliqué que la technologie de

¹¹¹ Audition de Marlène Dolveck, Directrice générale de SNCF Gares & Connexions et Directrice Générale Adjointe du Groupe SNCF en charge de la Transformation et Julien Nicolas, Directeur Numérique du Groupe SNCF, par la Commission temporaire IA, le 3 octobre 2024.

l'IA et ses usages sont encore si nouveaux et susceptibles d'évolution que l'intégration et le déploiement de l'IA dans l'entreprise devait nécessairement être réalisée étapes par étapes, en utilisant des outils d'évaluation et de mesure des impacts de l'IA alternatifs aux facteurs clés de succès traditionnels, comme par exemple l'accès aux compétences, l'implication des parties prenantes (dirigeants, organisations de salariés, etc.), ou l'accès aux technologies, etc.

La SNCF appelle par ailleurs à une certaine vigilance quant à l'appréciation de l'impact de l'IA sur l'emploi dans les entreprises. En effet, les concepteurs de modèles ou de systèmes d'IA annoncent parfois des gains de temps ou des améliorations de taux de productivité humaine en pourcentage, mais ces chiffres sont établis sur des hypothèses décorrélées de la réalité des entreprises. Par exemple, un développeur peut économiser x% de temps de travail s'il utilise l'IA générative, ce pourcentage pouvant être multiplié en fonction de son volume horaire à l'année. Mais en pratique, un développeur dans une entreprise ne développe pas 8h/jour, il a d'autres missions (maintenance, chefferie de projet, etc.), si bien que le pourcentage annoncé par le concepteur du modèle ou système d'IA s'en trouve nécessairement réduit.

Il est donc nécessaire, pour évaluer les bénéfices et l'impact que peut avoir l'IA sur une entreprise, d'adapter cette évaluation à chaque

cas particulier, chaque métier ou chaque secteur. Fort de cette granularité dans l'évaluation, il est possible de tirer des conclusions chiffrées. Par exemple, la Direction des services marchands de la Direction générale des entreprises¹¹², a mesuré les apports de l'IA pour les professions libérales. Ainsi, les notaires disent gagner 1h30 par dossier en utilisant l'IA générative, et les médecins disent gagner l'équivalent d'1h30 par jour grâce à une automatisation des comptes-rendus de consultation, la génération automatique d'email pour adresser des patients à des confrères, etc.

2. Les conditions de succès d'une intégration de l'IA dans l'entreprise

L'intégration réussie de l'IA dans une entreprise ou une structure non marchande est conditionnée à plusieurs étapes obligatoires, identifiées par le CESE au travers des auditions menées par la Commission temporaire sur l'IA. Elle nécessite un travail impératif de concertation, de dialogue et de formation des travailleurs salariés ou des bénévoles. Le choix d'une entreprise ou d'une association d'adopter l'IA doit prendre en considération les préoccupations des salariés et les impacts sur leurs conditions de travail dans un objectif de progrès.

2.1 LA PRISE EN COMPTE DE L'IMPACT SOCIAL DE L'IA

L'intégration de l'IA doit être réalisée de manière éthique. Les entreprises

¹¹² Audition d'Aurélien Palix, Sous-directeur des réseaux et des usages du numérique de la Direction générale des entreprises (DGE) du Ministère de l'Economie et des finances, par la Commission temporaire IA, le 26 septembre 2024.

doivent veiller à ce que l'IA soit utilisée de manière transparente, respectant les normes éthiques et les réglementations en vigueur. Cela inclut la transparence dans l'utilisation des données ou les algorithmes déployés et le respect de la vie privée des clients et des salariés.

Cela suppose aussi que l'IA soit déployée avec l'adhésion des salariés. En pratique, cela signifie investir au plus tôt dans la formation et l'accompagnement des salariés pour les aider à s'adapter aux nouvelles technologies mais aussi assurer la participation des employés aux processus de prise de décisions vis-à-vis de l'intégration de l'IA dans l'entreprise, et les sensibiliser aux enjeux éthiques et de protection des données.

Un accompagnement adéquat, un dialogue social nourri et la co-construction de solutions RH avec les salariés favorisent une adoption harmonieuse de l'IA.

2.2. L'IDENTIFICATION DES OBJECTIFS DE L'ENTREPRISE ET DES BESOINS AUXQUELS VIENT RÉPONDRE L'IA

L'étape préalable indispensable à toute intégration de l'IA en entreprise consiste à évaluer précisément les besoins et les objectifs de l'entreprise pour lesquels l'IA peut apporter une valeur ajoutée significative. Cela peut inclure l'amélioration de l'efficacité opérationnelle, l'optimisation des processus, l'amélioration de l'expérience client, l'innovation de produits et services, l'amélioration de la conciliation des temps de vie personnelle et professionnelle, la réduction des vecteurs d'accidents du travail et de maladies professionnelles.

La nécessité de cette évaluation approfondie permet d'éviter les surinvestissements ou les déploiements hasardeux qui peuvent mettre en péril l'organisation voire l'entreprise dans son ensemble.

Une fois les besoins et les objectifs

définis, il convient en outre d'identifier les cas d'usage spécifiques de l'IA.

Cela peut inclure l'automatisation des tâches répétitives, l'analyse prédictive, la personnalisation des offres, l'amélioration de la gestion de la chaîne d'approvisionnement, etc. L'identification de ces cas d'usage pertinents permet de maximiser l'impact de l'IA sur les performances de l'entreprise, de définir avec précision les orientations stratégiques et de commencer à définir les impacts et les défis liés au déploiement.

2.3 LA PRÉPARATION DE LA MISE EN PRODUCTION DES SYSTÈMES D'IA

L'entreprise doit s'adapter en interne pour que le déploiement de l'IA soit utile.

La gestion des données : Dans certains cas le déploiement de l'IA sera efficace, si celle-ci est basée sur les données de l'entreprise - cela suppose donc que cette dernière collecte et structure sa donnée... - et si les données sont interopérables entre les différents services de l'entreprise, c'est-à-dire que les données sont écrites dans un seul et même langage, selon une seule et même structure, pour que tous les services de l'entreprise puissent se comprendre. En pratique, cela suppose donc que l'entreprise mette en œuvre une procédure de collecte des données complètes, à jour, de haute qualité ; traite les données (anonymisation, étiquetage, nettoyage, etc.) en vue de créer l'interopérabilité ; et déploie des processus de protection des données. La gestion rigoureuse des données est essentielle pour garantir la fiabilité des systèmes d'IA : cette étape n'est pas simple à réaliser et crée un obstacle important pour beaucoup d'entreprises. Tous les acteurs ont un intérêt à intégrer dans leur maquette ressources humaines (RH) une responsabilité « données et IA ». Ce rôle stratégique permet de définir et d'exécuter des stratégies globales

en matière de données et d'IA. Il permettra également aux entreprises de faciliter leur conformité vis-à-vis des réglementations en vigueur, à commencer par le règlement européen sur l'IA, et aux grands principes éthiques.

L'évaluation des capacités technologiques : l'IA va ajouter une surcouche technologique à une infrastructure technologique existante. Cela suppose d'évaluer et le cas échéant, de renforcer l'infrastructure informatique existante, les logiciels, et les compétences techniques des employés. Lorsque le renfort interne n'est pas possible, l'entreprise doit évaluer le coût d'opportunité d'avoir recours à des solutions externes, notamment au travers de la sous-traitance, et les risques associés (notamment en vue du respect des normes éthiques et des exigences de sécurité de l'entreprise).

Le risque cyber et la fuite des données : L'intégration de l'IA dans les entreprises expose à des risques accrus de cybersécurité, notamment par le biais de systèmes vulnérables aux attaques. Les données sensibles utilisées pour entraîner les modèles peuvent être compromises, entraînant des fuites massives aux conséquences financières et réputationnelles graves. Il est donc essentiel de renforcer la sécurité des infrastructures et d'adopter des pratiques rigoureuses pour protéger ces informations.

2.4 L'ÉVALUATION DES COÛTS

L'intégration de l'IA nécessite souvent des investissements lourds en termes d'accès à des solutions technologiques innovantes, d'adaptation et maintenance

d'infrastructures technologiques internes et externes, de formation des employés ou de mise en conformité réglementaire.

Ce coût technologique, humain et juridique doit pouvoir être absorbé par l'entreprise sans entamer sa rentabilité, et pouvoir ensuite être absorbé par le client final dans le prix du produit ou service proposé. C'est pourquoi l'IA n'est pas nécessairement accessible à toutes et tous, et ne doit être déployée que si un besoin de l'entreprise ne peut être résolu qu'au travers de cette technologie.

2.5 PRISE EN COMPTE DE L'IMPACT ENVIRONNEMENTAL

Il est également important de considérer les impacts environnementaux de l'IA. Une évaluation équilibrée des risques et des impacts permet de garantir une intégration responsable et durable de l'IA, notamment dans les politiques de RSE des entreprises. À l'ère de la crise écologique et du réchauffement climatique, ces études d'impact sont recommandées pour permettre, par exemple, le recours à des IA « frugales » ou « labellisées ».

2.6 LE CADRE RÉGLEMENTAIRE DE L'ENTREPRISE

Les entreprises doivent se mettre en conformité avec le règlement européen sur l'IA. De même, le règlement général de la protection des données personnelles doit garantir que la collecte et le traitement des données afin d'entraîner les systèmes d'IA respectent les principes fixés.

Les préconisations du CESE

- **Enjeux éthiques et écologiques dans la conception et la mise en place de l'IA**

Les enjeux éthiques et environnementaux soulevés par les systèmes d'IA, nécessitent de respecter les réglementations et de minimiser l'impact écologique des technologies numériques. Les exemples de « bonnes pratiques » déployées par le groupe SNCF¹¹³ peuvent être intéressants à dupliquer et à adapter aux différentes échelles d'application. Ainsi, la SNCF a créé une charte éthique pour l'utilisation de l'IA, qui définit les lignes directrices relatives à la manipulation et à la protection des

données, le respect du RGPD et la consommation énergétique des outils d'IA en fonction de leur utilité et finalité. La direction de l'éthique est impliquée dès le début de chaque projet lié à l'IA pour s'assurer que les pratiques respectent les normes éthiques et réglementaires. L'impact environnemental de l'IA est en outre pris en compte dans la formation initiale et continue notamment des développeurs en les formant à l'écoconception pour réduire la consommation énergétique des data centers et des algorithmes.

PRÉCONISATION #19

S'agissant du déploiement des systèmes d'IA dans les lieux de travail, le CESE recommande aux entreprises, dans le cadre du dialogue social, de s'engager à travers la rédaction de chartes éthiques pour :

- **informer en toute transparence et clarté les utilisateurs quant à l'utilisation des données, au respect des réglementations (RGPD, AI Act) et l'impact environnemental des outils d'IA ;**
- **intégrer des exemples de « bonnes pratiques » pour favoriser la sobriété des usages numériques ;**
- **réaliser une évaluation préalable des impacts des systèmes d'IA sur le quotidien, le travail, le changement de trajectoire de compétences et d'emploi qui doit associer les salariés et leurs représentants ;**
- **questionner l'intérêt d'un déploiement au regard des alternatives.**

¹¹³ Audition de Marlène Dolveck, Directrice générale de SNCF Gares & Connexions et Directrice Générale Adjointe du Groupe SNCF en charge de la Transformation et Julien Nicolas, Directeur Numérique du Groupe SNCF, par la Commission temporaire IA, le 3 octobre 2024.

PRÉCONISATION #20

Le CESE recommande aux entreprises et aux trois versants de la fonction publique qui déploient des solutions d'IA de le faire dans le respect des limites planétaires, et notamment à :

- mesurer l'impact environnemental de cette technologie sur leurs usages. Cette mesure est cependant rendue complexe par le manque d'informations transmises par les grands producteurs de modèles d'IA et d'infrastructures (cloud, puces, etc.) ;
- promouvoir l'efficacité énergétique et la réduction des déchets tout au long du cycle de vie des services d'IA ;
- adopter des pratiques de développement durable en favorisant la collaboration entre producteurs, fournisseurs et utilisateurs ;
- éduquer et sensibiliser tous les acteurs aux enjeux de frugalité de consommation des ressources et de durabilité des équipements. L'AFNOR¹, en partenariat avec de nombreux acteurs économiques et institutionnels, vient de publier un référentiel pour une IA frugale qui comporte des pistes de mise en œuvre de grilles d'auto-évaluation et de bonnes pratiques sur l'amélioration de l'impact environnemental de l'IA.

¹ AFNOR SPEC 2314. (2023). Référentiel général pour l'IA frugale.

PRÉCONISATION #21

Le CESE appelle le gouvernement français et son administration, ainsi que les collectivités territoriales à accompagner les TPE, PME, ETI, grandes entreprises et l'ensemble des acteurs de l'ESS qui souhaitent adopter l'IA, en tenant compte des inégalités de moyens et des besoins différents, pour favoriser le chemin vers le déploiement d'une IA plus souveraine.

ENCADRÉ 4 : ZOOM SUR UN OUTIL AU SERVICE DES ENTREPRISES : LE TABLEAU DE MATURITÉ DU DÉPLOIEMENT DE L'IA

La Cour des Comptes¹¹⁴ propose un tableau de maturité en matière d'IA qui, à partir de la définition de critères, propose d'établir une évaluation initiale, d'élaborer un tableau de maturité, qui doit être mis en œuvre de manière progressive avec un suivi et une mise à jour régulière. Les résultats du suivi du tableau de maturité doivent être communiqués de manière transparente aux parties prenantes internes et externes. En suivant ces étapes :

1. Définition des critères de maturité

La première étape consiste à définir les critères de maturité pertinents pour évaluer le déploiement de l'IA dans l'entreprise. Ces critères doivent inclure :

Stratégie et Gouvernance : Existence d'une stratégie claire pour l'IA et d'une gouvernance dédiée.

Données et Infrastructure : Qualité et disponibilité des données, ainsi que l'infrastructure technologique.

Compétences et Formation : Niveau de compétences des employés et programmes de formation en IA.

Cas d'Usage et Implémentation : Identification et mise en œuvre des cas d'usage de l'IA.

Éthique et Conformité : Respect des normes éthiques et des réglementations en vigueur.

Impact et Performance : Mesure des impacts économiques, sociaux et environnementaux de l'IA.

Exemple : Une entreprise de services financiers pourrait définir des critères de maturité tels que l'automatisation des processus de conformité, l'amélioration de la détection des fraudes, et l'optimisation de la relation client grâce à des chatbots.

2. Évaluation initiale

Une évaluation initiale doit être réalisée pour établir un point de départ. Cette évaluation permet de déterminer le niveau de maturité actuel de l'entreprise en matière d'IA. Elle peut être réalisée à l'aide de questionnaires, d'entretiens et d'analyses de documents internes.

Exemple : L'entreprise de services financiers réalise une évaluation initiale et découvre qu'elle est au niveau 1 (Initial) pour l'automatisation des processus de conformité mais au niveau 2 (Développement) pour la détection des fraudes.

¹¹⁴Cour des comptes, L'intelligence artificielle dans les politiques publiques : l'exemple du ministère de l'économie et des finances, 22 octobre 2024.

3. Élaboration du tableau de maturité

Sur la base des critères définis et de l'évaluation initiale, un tableau de maturité est élaboré. Ce tableau doit être structuré de manière à permettre une évaluation claire et objective des progrès réalisés. Il peut inclure des niveaux de maturité tels que :

Niveau 1 : Initial : L'entreprise commence à explorer l'IA, mais n'a pas encore de stratégie ou de gouvernance en place.

Niveau 2 : Développement : L'entreprise a défini une stratégie et commence à mettre en place des projets pilotes.

Niveau 3 : Intégration : L'IA est intégrée dans plusieurs processus métiers, avec une gouvernance et une infrastructure en place.

Niveau 4 : Optimisation : L'IA est optimisée et utilisée de manière efficace dans l'ensemble de l'entreprise.

Niveau 5 : Leadership : L'entreprise est un leader dans l'utilisation de l'IA, avec des pratiques exemplaires et une innovation continue.

Exemple : Le tableau de maturité de l'entreprise de services financiers montre qu'elle est au niveau 2 pour la détection des fraudes et au niveau 1 pour l'automatisation des processus de conformité.

4. Mise en œuvre du tableau de maturité

Le tableau de maturité doit être mis en œuvre de manière progressive. Les dirigeants doivent communiquer clairement les objectifs et les attentes à l'ensemble des employés. Des sessions de formation et des ateliers peuvent être organisés pour sensibiliser les équipes aux critères de maturité et aux étapes à suivre. Un dialogue social est essentiel pour assurer l'adhésion des employés et répondre à leurs préoccupations.

Exemple : L'entreprise de services financiers organise des ateliers pour former les employés à l'utilisation des nouveaux outils d'IA et les sensibiliser aux critères de maturité. Un comité de dialogue social est mis en place pour recueillir les retours des employés et ajuster les stratégies en conséquence.

5. Suivi et mise à jour

Le suivi régulier du tableau de maturité est essentiel pour mesurer les progrès réalisés et identifier les domaines nécessitant des améliorations. Des revues périodiques doivent être organisées pour évaluer les résultats et ajuster les stratégies en conséquence. Le tableau de maturité doit être mis à jour en fonction des évolutions technologiques et des besoins de l'entreprise. Une prise en compte des ressources humaines est cruciale pour adapter les formations et les accompagnements nécessaires.

Exemple : L'entreprise de services financiers organise des revues trimestrielles pour évaluer les progrès réalisés dans la détection des fraudes et l'automatisation des processus de conformité, et ajuste ses stratégies en conséquence. Les besoins en formation et en accompagnement des employés sont régulièrement réévalués.

6. Rapport et communication

Les résultats du suivi du tableau de maturité doivent être communiqués de manière transparente aux parties prenantes internes et externes. Des rapports réguliers peuvent être élaborés pour présenter les progrès réalisés, les défis rencontrés et les actions correctives mises en place. Une communication claire et ouverte favorise l'adhésion des employés et renforce la confiance des partenaires et des clients.

Exemple : L'entreprise de services financiers publie des rapports trimestriels pour informer les employés, les partenaires et les clients des progrès réalisés dans le déploiement de l'IA et des actions correctives mises en place. Des réunions régulières avec les représentants des employés sont organisées pour discuter des impacts de l'IA sur les conditions de travail et les perspectives d'évolution professionnelle.

En mettant en place un tableau de maturité et en suivant ces étapes, les dirigeants d'entreprises peuvent structurer et piloter efficacement le déploiement de l'IA, tout en assurant une évaluation continue de la maturité et des impacts de cette technologie. Une approche équilibrée et réfléchie, incluant un dialogue social et une prise en compte des ressources humaines, est essentielle pour maximiser les avantages tout en minimisant les risques.

Recommandation de citoyennes et citoyens

L'intégration de l'IA en entreprise peut apporter des avantages comme l'amélioration de la productivité et la réduction des coûts. Cependant, cette intégration peut aussi entraîner des disparités selon les profils des entreprises, des tensions sociales, avec des employés inquiets de la perte de leur emploi ou de la dégradation de leurs conditions de travail. Le CESE recommande que les entreprises accompagnent cette transition avec des formations adaptées et une gestion responsable, afin de préserver les droits des travailleurs et de garantir une utilisation équitable de l'IA. L'objectif est de s'assurer que l'IA profite à la fois à l'économie et aux employés.

C. Les services publics et l'IA

Lorsqu'elle est utilisée dans le service public, l'IA doit répondre aux spécificités de l'action publique et à ses obligations particulières liées à l'égalité d'accès, à la transparence et à la redevabilité. Le service public obéit, en effet, aux lois de Rolland¹¹⁵ qui posent trois principes : continuité, égalité et adaptabilité (ou « mutabilité »).

Ces principes ont une incidence directe sur le déploiement de l'IA dans les services publics, présentée par Laure Lucchesi¹¹⁶ lors de son audition :

- le principe de continuité, qui impose que le service public soit effectivement et toujours fourni, questionne à la fois la dépendance de l'État aux fournisseurs d'IA (notamment pour éviter toute forme de rupture dans la fourniture et la maintenance des logiciels) et les enjeux de souveraineté (sur le stockage et le traitement de données sensibles) ;
- le principe d'égalité exige d'éviter toutes les formes de discriminations en garantissant l'accès de toutes et tous au service public, et oblige l'État à lutter contre l'illectronisme ;
- le principe d'adaptabilité impose au service public de s'adapter aux progrès techniques mais aussi de répondre aux attentes de la société : ce principe peut être interprété

comme une injonction à « mettre à jour le logiciel » de l'action publique et de s'interroger sur le bien-fondé des transformations envisagées.

1. L'IA est déjà intégrée dans les services publics

Un récent rapport de la Défenseure des droits¹¹⁷ souligne que « *les algorithmes et les systèmes d'IA sont déployés de plus en plus massivement dans les politiques publiques et les services publics. L'action publique comme les droits des usagers concernés s'en trouvent substantiellement affectés* ». En effet, les exemples ne manquent pas sur des usages de l'IA dans les services publics.

1.2 L'INTÉGRATION DE L'IA AU SEIN DE LA DIRECTION GÉNÉRALE DES FINANCES PUBLIQUES (DGFIP) : L'EXEMPLE D'UN DÉPLOIEMENT NON CONCERTÉ AVEC DES OBJECTIFS D'INTÉRÊT GÉNÉRAL FLOUS

Depuis 2017, l'administration fiscale s'est lancée dans un processus d'automatisation croissante, avec pour objectif de dématérialiser la totalité des démarches administratives. La DGFIP utilise l'IA pour optimiser ses processus, en particulier dans le cadre de la lutte contre la fraude fiscale. Afin de garantir une meilleure fiabilité

¹¹⁵ Les lois du service public telles que conçues par Louis Rolland dans les années 1930 sont au nombre de trois : le principe de continuité, de mutabilité et d'égalité du service public.

¹¹⁶ Audition de Laure Lucchesi, ancienne directrice d'Etalab et experte de la transformation digitale et data IA dans les secteurs public et privé, par la Commission temporaire IA, le 24 octobre 2024.

¹¹⁷ Rapport de la Défenseure des droits, Algorithmes, systèmes d'IA et services publics : quels droits pour les usagers ? Points de vigilance et recommandations, novembre 2024.

des bases de la fiscalité directe locale, la DGFIP recourt, dans le cadre du projet « Foncier innovant », aux technologies novatrices d'IA et de valorisation des données à partir des prises de vue aériennes de l'Institut national de l'information géographique et forestière (IGN). Sur le site [Impot.gouv.fr](http://impot.gouv.fr), ce projet est présenté de façon positive car « *il vise, en optimisant le processus de détection des constructions ou aménagements non déclarés, à permettre de lutter plus efficacement contre les anomalies déclaratives et ainsi mieux répondre aux souhaits d'équité et de justice fiscale des citoyens, par la juste imposition des biens. Il va également offrir la possibilité de faciliter la représentation sur le plan cadastral des bâtiments et des piscines, à partir des prises de vues aériennes* ». Des erreurs constatées ont été corrigées suite aux signalements qui ont dû être effectués par les agents.

Un autre point de vue a été présenté par Sandra Demarçq et Benjamin Gandouin, représentants de Solidaires Finances publiques, syndicat auteur de l'ouvrage *L'intelligence artificielle aux impôts : réflexions et actions syndicales*. Ces derniers ont expliqué que les outils d'IA présentaient des défauts notables, comme l'incapacité à distinguer une piscine d'une simple bache bleue, entraînant des erreurs d'imposition injustifiées, coûteuses pour l'État qui doit mobiliser ses ressources internes pour effectuer des contrôles sur le terrain et indemniser les contribuables victimes de telles erreurs. Cela illustre, pour Sandra Demarçq et Benjamin Gandouin, les limites de l'IA dans des tâches traditionnellement effectuées par des agents publics. Par-delà l'utilisation des

systèmes d'IA à des fins analytiques et/ou de vérification, on peut aussi observer que l'usage des données individuelles à l'appui de l'établissement et de la diffusion des actes d'imposition pose question et n'est pas, pour l'heure, exempté de dysfonctionnements. En témoigne l'envoi tout récent d'avis de taxe sur les logements vacants à des contribuables qui n'y étaient pas assujettis¹¹⁸. Ils mettent également en avant l'opacité entourant le déploiement des technologies d'IA dans l'administration publique. Selon eux, ni les agents ni les représentants du personnel n'ont été consultés sur la mise en œuvre de ces nouvelles technologies. Les projets, souvent externalisés à des cabinets de conseil, sont décidés sans réelle transparence sur les algorithmes utilisés.

D'autres exemples témoignent au contraire que l'IA peut être mise au service d'une meilleure efficacité et efficience du service public. À titre d'exemple, le projet LlaMandement¹¹⁹ développé par la DGFIP permet de générer des résumés neutres des amendements parlementaires, soulageant ainsi les agents publics qui doivent analyser et traiter un nombre croissant d'amendements lors de l'examen du Projet de loi de finances. Ce projet a été transféré à la Direction interministérielle du numérique (DINUM) qui a pour mission d'élaborer la stratégie numérique de l'État et de piloter sa mise en œuvre, afin qu'il soit partagé avec d'autres ministères, notamment les ministères sociaux, qui sont confrontés à des problématiques similaires dans le cadre du Projet de loi de financement de la sécurité sociale.

L'IA peut aussi être mise au service d'une meilleure qualité du service public. Ainsi,

¹¹⁸ Les Échos, *Impôts : des taxes sur les logements vacants envoyées par erreur*, 18 novembre 2024.

¹¹⁹ Le nom du projet est un jeu de mots entre « amendements » et « LLM », le modèle d'IA d'apprentissage automatique capable de comprendre et générer des textes.

« Albert » est un outil mis au service des agents des maisons France Services. Il s'agit d'une IA développée par l'équipe DataLab de la Direction interministérielle du numérique (DINUM) pour aider les agents de l'administration à mieux répondre aux demandes des Français. Cet outil d'IA permet de communiquer aux citoyens la bonne information sans demander aux agents d'être omniscients, dans un contexte où la réglementation est de plus en plus complexe, et évolue constamment.

L'IA est, en outre, utilisée dans les services publics territoriaux. Ainsi, dans le domaine environnemental, le projet « Aigles » développé par la Direction départementale des territoires (DDT) de l'Hérault contribue à une meilleure efficacité de la protection du littoral en automatisant la détection de nouveaux objets dans des milieux sensibles (caravanes, constructions illégales, décharges sauvages, etc.) à partir d'images satellites.

2. Le service public à l'épreuve de l'IA

En application des lois de Rolland, le déploiement des systèmes d'IA dans les services publics doit se faire au service de l'intérêt général. La Défenseure des droits appelle à promouvoir, « dans l'ensemble des administrations une véritable *« culture des droits », centrée sur les usagers. Cela suppose de placer le respect des droits des usagers au cœur des objectifs assignés aux managers publics comme aux agents de terrain et, corrélativement, de ne pas laisser l'administration se concentrer, notamment s'agissant du déploiement des systèmes algorithmiques, sur des objectifs prioritaires de réduction des coûts et de personnel et sur des indicateurs de rentabilité »*¹²⁰.

Les algorithmes peuvent offrir, dans certaines situations, une solution technologique pour mettre en œuvre des droits, par exemple pour traiter un nombre important de données. Cependant, dans tous les cas, l'IA doit s'adapter à l'humain et elle doit être employée au service des agents, pour les aider et non pour les remplacer ou les positionner dans une relation de subordination à la machine. En l'état de la science et de la technologie, mettre l'IA face à des usagers est aujourd'hui impossible, cela mettrait les citoyens et citoyennes face à des risques d'hallucinations ou d'arbitraire, qui s'opposeraient directement aux principes de qualité ou d'opposabilité du service public.

¹²⁰ Rapport de la Défenseure des droits, *Algorithmes, systèmes d'IA et services publics : quels droits pour les usagers ? Points de vigilance et recommandations*, novembre 2024.

Autre conséquence des lois de Rolland : quand elle aboutit à des décisions individuelles, l'action publique est soumise à des obligations particulières de transparence, de redevabilité et d'explicabilité. L'usage de l'IA doit ainsi se conformer au principe posé à l'article 15 de la Déclaration des Droits de l'Homme et du Citoyen de 1789 : « *La société a le droit de demander compte à tout agent public de son administration* ». En d'autres termes, même rendue par une IA, une décision publique individuelle doit pouvoir être expliquée. Le code des relations entre le public et l'administration oblige, en outre, à tenir l'inventaire des différents algorithmes utilisés par l'administration et à en publier les règles¹²¹. Ces obligations relèvent autant des droits des usagers que de l'administration elle-même, pour l'intelligibilité de sa propre action.

Or, les modèles de langage ne permettent pas, aujourd'hui, de retracer la manière dont un résultat a été produit, les raisons pour lesquelles telle décision a été prise par l'algorithme ou de détecter des réponses fausses. L'IA générative fournit des réponses vraisemblables et possibles, selon des probabilités mathématiques, n'empêchant ni les erreurs, ni les hallucinations.

¹²¹ Article L.312-1-3 du Code des relations entre le public et l'administration - Légifrance.

ENCADRÉ 5 : LE « CAS » PARCOURSUP

Parcoursup est un exemple prégnant de la problématique de l'explicabilité de la décision individuelle dans le cadre d'une IA utilisée dans un service public.

Le système d'inscription dans l'enseignement supérieur a toujours été complexe, et le recours aux technologies s'est imposé depuis plus de trente ans : à partir de 1987, les vœux des futurs étudiants sont enregistrés via le Minitel. En 2002, l'algorithme APB (Admission Post Bac) prend le relais. Depuis 2018, c'est via Parcoursup, la plateforme nationale de préinscription en première année de l'enseignement supérieur, que les candidats sont tenus de formuler des vœux d'admission dans des formations.

Or, Parcoursup reste critiqué pour son manque de transparence, de continuité et de justice. Un rapport du Sénat publié en 2023¹²² souligne ainsi l'opacité de la procédure d'attribution des formations et le manque d'explicabilité sur les sélections faites via les algorithmes de Parcoursup.

Saisi d'une question prioritaire de constitutionnalité à ce sujet, le Conseil constitutionnel, dans une décision du 3 avril 2020¹²³, a estimé que les établissements d'enseignement supérieur devaient publier « sous la forme d'un rapport, les critères en fonction desquels les candidatures ont été examinées et précisant, le cas échéant, dans quelle mesure des traitements algorithmiques ont été utilisés pour procéder à cet examen ». Ainsi, la diffusion a posteriori des critères et des modalités d'examen des candidatures et donc, le cas échéant, des critères utilisés dans le cadre du traitement algorithmique de ces candidatures s'impose aux établissements d'enseignement supérieur. Pour la juridiction suprême, l'article L. 612-3 du Code de l'éducation, qui prévoit « la nécessaire protection du secret des délibérations des équipes pédagogiques chargées de l'examen des candidatures » constitue une exception aux modalités larges du droit d'accès tel qu'il est prévu par le Code des relations entre le public et l'administration, mais limite la portée de cette exception, qui ne concerne, au final, que les tiers et encore en amont de la clôture de la procédure de présélection.

Il convient de noter qu'en tant que système d'IA déterminant l'accès, l'admission ou l'affectation dans les établissements d'enseignement supérieur et de formation professionnelle à tous les niveaux, Parcoursup est désormais classé par l'AI Act « système à haut risque », et devra se conformer à des obligations accrues en termes de transparence et d'explicabilité, d'ici le 2 août 2027.

¹²² Sénat, Rapport d'information n° 793 (2022-2023), Parcoursup : l'urgence à gagner la confiance des lycéens et des étudiants, 28 juin 2023.

¹²³ Décision du Conseil constitutionnel, 3 avril 2020, n° 2020-834 QPC.

3. L'IA ne peut remplacer l'humain dans les services publics

En 2024, la Défenseure des droits¹²⁴ rappelle « *qu'examiner les enjeux de l'action publique algorithmique ne doit pas conduire à occulter l'importance des problématiques spécifiques liées à la « seule » dématérialisation des services publics* ». Ce sujet était déjà au cœur du rapport de 2022 : « *La procédure dématérialisée devrait devenir une alternative aux autres modes de communication, au libre choix de l'utilisateur, et non un carcan imposé à tous* »¹²⁵.

En effet, la digitalisation, imposée sans alternative aux usagers du services publics, peut renforcer l'exclusion. Les freins que constituent le manque de compétences numériques et le manque d'accès aux infrastructures (matériel et connexion Internet fiable) peuvent être redoublés par l'intelligence artificielle. Les personnes qui ne maîtrisent pas internet ou qui n'ont pas les moyens de s'équiper, seront laissées pour compte comme l'ont déjà souligné les rapports du Défenseur des droits en 2019, 2022 et 2024 et l'avis du CESE *Droits sociaux : accès et effectivité*¹²⁶.

L'accueil physique est un lien essentiel avec les usagers. Dans une ère où la technologie occupe une place prépondérante, l'accueil physique et humain dans les services publics demeure indispensable. Cet accueil permet de créer un lien direct et personnel avec les usagers. Ce contact humain est crucial pour répondre aux besoins spécifiques et variés des citoyens. L'échange en

personne permet de mieux comprendre les demandes, de fournir des explications claires et de rassurer les usagers. Le Label Marianne¹²⁷, devenu Services Publics + depuis début 2024, met en lumière cette nécessité et valorise cet aspect en certifiant que les services publics offrent un accueil chaleureux, courtois et respectueux, garantissant ainsi une expérience positive et satisfaisante pour tous.

Le contact humain va au-delà de la simple interaction. Il incarne une écoute active, de l'empathie et une capacité à personnaliser les réponses en fonction de chaque situation. Les agents d'accueil, formés pour répondre avec compétence et bienveillance, jouent un rôle fondamental dans l'accompagnement des usagers. Grâce à l'humain, les services publics peuvent s'adapter plus facilement aux besoins spécifiques et assurer une dimension humaine essentielle que la technologie ne peut remplacer. Le Label Marianne représente un engagement fort des services publics à fournir une qualité d'accueil exemplaire. Il repose sur des critères stricts tels que la disponibilité, l'efficacité, la clarté des informations et le respect des usagers. Cette certification met en avant les efforts de la ville pour offrir un accueil chaleureux et efficace, renforçant ainsi la confiance des citoyens envers leurs institutions. En mettant l'accent sur ces valeurs, les services publics renforcent leur mission de service au citoyen et contribuent à une société plus solidaire et compréhensive.

¹²⁴ Rapport de la Défenseure des droits, Algorithmes, systèmes d'IA et services publics : quels droits pour les usagers ? Points de vigilance et recommandations, novembre 2024, p. 4.

¹²⁵ Défenseure des droits, Dématérialisation des services publics : trois ans après, où en est-on ?, rapport, 2022, p. 59.

¹²⁶ CESE, *Droits sociaux : accès et effectivité*, avis 2024-021 rapporté par Isabelle Dorese et Catherine Pajares y Sanchez au nom de la Commission Affaires sociales et santé, novembre 2024.

¹²⁷ Direction interministérielle de la transformation publique, Label de qualité pour les services publics.

Les préconisations du CESE

Moderniser ou réformer le service public, notamment pour permettre son accessibilité et sa continuité au travers de l'intégration de l'IA, nécessite de prendre en considération plusieurs interrogations de fond. Pourquoi cette modernisation est-elle essentielle : améliorer la qualité des services ? réduire les coûts ? réduire les effectifs d'agents publics ? En interrogeant les finalités de la transformation, cela permet de prioriser les projets les plus transformants en tenant compte des critères éthiques et des principes inhérents au service public.

PRÉCONISATION #22

Le CESE appelle à déployer des systèmes d'IA à condition :

- **qu'ils améliorent la qualité du service public et les conditions de travail des agents qui y concourent ;**
- **que soient véritablement appliquées les obligations de transparence, de redevabilité et d'explicabilité, lorsque les systèmes d'IA sont déployés pour exécuter le droit (par exemple, pour calculer le montant d'un impôt), qu'ils s'avèrent incontournable (à l'instar de Parcoursup, plateforme indispensable pour accéder à l'enseignement supérieur) ou qu'ils aboutissent à des décisions individuelles produisant des effets juridiques, économiques ou sociaux qui affectent les personnes.**

PRÉCONISATION #23

Face à une digitalisation croissante des services publics, le CESE recommande de garantir leur accès et le droit au non-numérique, en :

- > proposant systématiquement un accompagnement et un accueil humain en alternative au service dématérialisé pour garantir l'accès et la liberté de choix de tous les citoyens, comme cela se déploie au travers des espaces France Service.**
- > inscrivant dans la loi le principe d'un droit au non-numérique, et en garantissant l'égalité d'accès et la liberté de choix pour toutes et tous.**

Recommandation de citoyennes et citoyens

L'IA peut rendre les services publics plus efficaces et accessibles, en automatisant certaines tâches et en simplifiant les démarches. Mais elle ne doit pas remplacer l'humain, surtout dans des domaines sensibles comme les services sociaux. Le CESE recommande de miser sur une digitalisation équilibrée, où la technologie améliore les services publics en prenant en compte les besoins des citoyens et citoyennes.

D. La frontière créée par l'éloignement numérique

L'IA peut constituer un levier au développement des territoires car elle est susceptible d'offrir des opportunités à l'économie réelle et au pilotage des politiques territoriales tout comme celui de la mise en œuvre des Smart City. De plus, l'IA peut aussi être un outil d'attractivité et de réponses aux spécificités géographiques ou climatiques notamment dans les territoires éloignés en particulier dans les Outre-mer. Ainsi, elle permettra une accélération des développements durable et résiliente comme, par exemple, dans la gestion des risques naturels, la préservation de la biodiversité ou l'ouverture à l'internationale.

Cependant, pour répondre à ces objectifs, il faut relever les défis de l'éloignement numérique.

L'éloignement numérique recoupe plusieurs notions complémentaires, souvent utilisées pour qualifier les inégalités face au numérique. Il fait ainsi référence à la fois :

- à la fracture numérique, à savoir la difficulté matérielle à avoir accès à des outils informatiques (smartphone, ordinateur, réseaux télécoms) ;
- à l'illectronisme, à savoir l'absence de maîtrise des compétences de base du numérique ;
- et inclut la difficulté de ne pas tirer pleinement profit des usages numériques.

L'IA joue un rôle ambivalent vis-à-vis de l'éloignement au numérique :

elle contribue à réduire ce dernier en simplifiant l'accessibilité de certaines compétences numériques de base, mais elle peut aussi creuser les inégalités - notamment sociales - si des individus ne peuvent tirer profit des bénéfices de l'IA. Avant d'identifier l'impact de l'IA sur l'inclusion et les inégalités, il est essentiel de qualifier d'abord la situation de l'éloignement au numérique en France en 2024.

L'éloignement numérique est une des fractures supplémentaires après celles d'ordre économique, territorial et social qui rongent notre pays.

Certaines personnes choisissent délibérément de ne pas recourir au numérique, pour des raisons philosophiques, écologiques, ou par méfiance vis-à-vis des technologies et de leurs implications (surveillance, perte d'intimité, etc.). Les développements suivants illustrent les problématiques liées à l'éloignement au numérique subi.

1. Éloignement au numérique : où en est-on en France en 2024 ?

Quelques chiffres pour comprendre la tendance¹²⁸ :

- 31,5 % de la population française se considère « éloignée du numérique », soit 16 millions de personnes.
- 8,8 % de la population de 18 ans et plus est aujourd'hui non-internaute, soit 4,5 millions de personnes
- 11,5 millions d'internautes manquent d'aisance dans les

¹²⁸ Crédoc, La société numérique française: définir et mesurer l'éloignement numérique, avril 2023.

usages avancés du numérique, soit 22,9 % de la population française ; 25 % des Françaises et Français ne savent pas s'informer en ligne

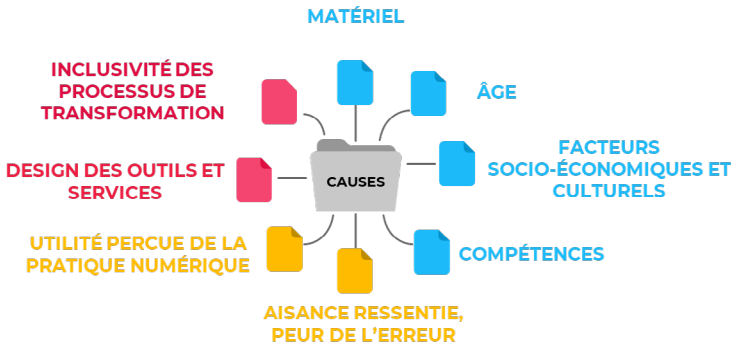
- 18 millions de Françaises et Français, soit 35,4 % de la population, ne se sentent pas à l'aise avec un ordinateur.

Cette inégalité face aux compétences numériques est en augmentation : si

18,3% des Françaises et des Français manquaient d'aisance dans les usages en 2019, ils sont 22,9 % en 2022 à se trouver dans cette situation¹²⁹.

Les inégalités face au numérique, ont plusieurs causes, regroupées dans le schéma suivant :

FIGURE 4 : LES INÉGALITÉS FACE AU NUMÉRIQUE



Source : Julhiet Sterwen, 2024

L'accès à l'équipement en smartphones ou tablettes reste vecteur d'inégalités, même si les taux d'équipements ont fortement progressé depuis dix ans. L'évolution technologique accélérée ainsi que l'obsolescence matérielle et logicielle accentuent ce phénomène.

L'âge est un facteur important, « les personnes âgées de 70 ans ou plus comptant une part de non-internautes nettement supérieure à la moyenne ». Cependant, « contrairement aux idées reçues, les plus jeunes ne sont

pas exempts d'une part significative d'éloignés du numérique (une personne sur cinq parmi les moins de 25 ans) ».

Le niveau de diplôme est également « associé à d'importants écarts dans l'éloignement : les personnes non diplômées sont environ trois fois moins souvent internautes que les personnes ayant un niveau d'études équivalent au BEPC. Elles sont 1,5 fois plus souvent éloignées du numérique »¹³⁰. L'illettrisme entraîne l'illectronisme.

129 Audition de Philippe Archias par la Commission temporaire sur l'IA, et ANCT, CREDOC, Université Rennes 2 CREAD-M@rsouin, La société numérique française : définir et mesurer l'éloignement numérique, 2023, infographie Ouest France La société numérique française: définir et mesurer l'éloignement numérique.

130 Marianne Bléhaut, Jérôme Clerget, Matthieu Serreau, Pascal Plantard, La société numérique française : définir et mesurer l'éloignement numérique, Crédoc avril 2023.

La mauvaise couverture des zones blanches de téléphonie mobile, le non-déploiement des réseaux de fibre et/ou de haut débit, particulièrement en milieu rural, découragent les habitants de ces zones d'utiliser le numérique.

Parce que l'absence de maîtrise des outils numériques (telle que la faculté de savoir écrire et améliorer un prompt) ou d'accès à du matériel fonctionnel risque de limiter l'accès aux services essentiels (santé, éducation, emploi), le lien de causalité entre inégalité numérique et inégalité sociale est renforcé.

Le phénomène d'éloignement au numérique s'assimile ainsi à une double peine : ce sont les populations les plus vulnérables (origine sociale, niveaux de formation, maîtrise de la langue, manque de capital culturel, manque de littératie numérique) qui ont les usages les plus limités et qui peinent le plus à accéder aux ressources numériques, alors même que ces ressources pourraient leur permettre d'accéder à leurs droits et à des services essentiels (santé, éducation, bancarisation, etc.), à un emploi, ou à des ressources pour des projets citoyens ou entrepreneuriaux.

C'est pourquoi l'inclusion numérique renvoie aussi à l'enjeu de lutte contre les inégalités sociales au sens large. Pour lutter contre ces inégalités, une politique d'inclusion numérique suppose d'agir sur l'accès à des équipements de qualité suffisante, l'accès à une connexion suffisante et l'accès à des compétences techniques mais aussi cognitives (gestion de l'information, esprit critique, créativité), sociales (communiquer avec autrui dans un contexte numérique) et de littératie numérique (utilisation confiante et

critique d'une gamme complète de technologies numériques pour l'information, la communication et la résolution de problèmes de base dans tous les aspects de la vie).

Au vu des évolutions technologiques rapides, il est nécessaire **de préparer la population à des usages diversifiés** (tous les aspects de la vie professionnelle ou privée sont concernés), **évolutifs** (pour s'adapter aux mutations permanentes des technologies) **et éclairés** (pour ne pas tomber dans des situations de dépendance ou d'enfermement, ni être victime d'arnaques, de cyberharcèlement).

2. La formation pour accompagner les publics éloignés du numérique

Selon Philippe Archias, sociologue du numérique, aussi bien la technologie de l'IA que ses usages sont si nouveaux, si mouvants, que nous avons encore peu de recul sur le rôle que pourrait avoir l'IA pour résorber ou atténuer les inégalités. Il appelle cependant au volontarisme car pour lui, « *l'IA, c'est aussi ce que l'on en fera* ».

2.1 FORMER À L'IA EST UN PRÉREQUIS POUR RÉDUIRE LES INÉGALITÉS FACE AU NUMÉRIQUE

Dans un rapport de 2024, l'UNESCO juge « indispensable de donner aux individus les connaissances, les aptitudes, les attitudes et les valeurs dont ils ont besoin non seulement pour se protéger des risques, mais aussi pour tirer pleinement parti des opportunités offertes par l'IA »¹³¹.

La formation à l'IA recoupe plusieurs dimensions : il s'agit à la fois de former à l'utilisation des outils (l'emploi de l'IA générative, la rédaction de prompt, la production de contenus), à la création d'outils (élaboration d'algorithmes, rédactions de code, etc.) mais également à la détection des limites de l'IA (erreurs, hallucinations ou biais dans les résultats produits par l'IA).

Un des enjeux est de rendre accessible la formation à l'IA. Beaucoup d'initiatives existent déjà.

La formation à l'IA peut s'inscrire dans des programmes de politiques publiques déjà engagés dans le cadre de la lutte contre les inégalités numériques.

Ainsi, le gouvernement français a développé depuis 2018 le programme Société numérique au travers de de l'Agence nationale pour la Cohésion des territoires, pour développer des solutions d'accompagnement aux usages du numérique au sein des collectivités et acteurs de proximité sur les territoires. Ce programme a permis de recruter et former 4 000 « conseillers numériques » chargés d'accompagner au numérique les personnes sur l'ensemble du territoire au plus près de leur domicile, dans les

maisons « France Services ». Ce chiffre envisageait d'être réduit à 1 500 dans le cadre du Projet de loi de finances pour 2025.

En 2021 a aussi été élaboré le programme « Aidants Connect » chargé de former les travailleurs sociaux, agents d'accueil de mairie ou autres aidants à l'accompagnement numérique à l'information, l'orientation et l'appui aux démarches des usagers, pour qu'ils accompagnent ensuite les personnes éloignées du numérique vers l'autonomie, tout en préservant leurs droits et leur confidentialité. Ces médiateurs travaillent dans des structures très diversifiées : au sein de grands acteurs à l'échelle nationale (Emmaüs Connect, Simplon, PIMMS, Mon Assistant Numérique), ou, plus localement, dans un nombre important de petites structures, essentiellement associatives, aux statuts très divers (EPN, PIMMS, régies de quartier, tiers lieux, fablabs, associations locales)¹³².

Philippe Archias pointe toutefois deux limites au recours aux travailleurs sociaux pour former au numérique : d'une part, l'acquisition et la mise à jour régulière des compétences, et d'autre part, la charge supplémentaire que représente cette tâche face à l'ensemble des autres actions qui leur incombent, à moyen constant et temps de travail équivalent, poussant ces aidants à souvent « faire à la place de » plutôt qu'« avec » les publics cibles.

À noter que les médiathèques et les mairies apparaissent comme « les lieux d'accompagnement numérique les plus connus » par les Françaises et les Français, d'après l'édition 2024 du Baromètre numérique¹³³.

¹³¹ UNESCO, L'autonomisation des utilisateurs grâce aux réponses apportées par l'éducation aux médias et à l'information à l'évolution de l'intelligence artificielle générative (IAG), 2024.

¹³² ANCT, Plan national pour la formation des aidants et médiateurs numériques, mars 2021.

¹³³ *Ibid.*

Cependant, ce baromètre pointe qu'en 2024, l'intelligence artificielle n'apparaît pas en tant que telle dans les « fondamentaux de la culture numérique », « l'usage du numérique raisonné et responsable » ainsi que le fait de « sensibiliser aux enjeux de sobriété numérique et à l'impact des usagers et des infrastructures sur l'empreinte carbone » étant jugés prioritaires. Aussi, pour favoriser l'appropriation de l'intelligence artificielle par l'ensemble de la population, **des « Cafés IA » ont été déployés sur tout le territoire français** sous l'égide du Conseil national du Numérique, à la suite des recommandations de la Commission IA en mars 2024¹³⁴. Il s'agit ici de fédérer, soutenir, promouvoir et rassembler les initiatives de débat mais aussi de partage de connaissances sur l'IA. Depuis 2024, une centaine de cafés IA ont déjà été organisés par différents acteurs sur l'ensemble du territoire.

Les entreprises peuvent aussi jouer un rôle dans le cadre de la formation au numérique et à l'IA. Les publics éloignés du numérique sont souvent des travailleurs, et le cadre de travail pourrait assumer le rôle de médiateur du numérique. Par exemple, au sein du groupe SNCF¹³⁵ est développée une stratégie de transformation numérique intégrant l'IA. L'un des facteurs clés de succès de cette stratégie est la prise en compte de l'humain avec un accompagnement renforcé. Ainsi, le groupe a mis en place des programmes d'acculturation

à ChatGPT (groupes de travail, conférences), pour familiariser les collaborateurs avec les concepts de l'IA et du numérique, en utilisant des modules adaptés à différents niveaux de compétences. La SNCF a mis en place des formations pour plus de 10 000 collaborateurs afin de les aider à comprendre et à utiliser l'IA. Les publics formés vont bien au-delà des seuls collaborateurs dont le métier est directement lié à la fonction numérique.

Pour utiliser l'IA comme levier d'inclusion, d'autres initiatives se multiplient dans la société civile. Ainsi en réponse au besoin de formation, l'association Konexio fondée en 2016 s'est fixé l'objectif de former les plus vulnérables (jeunes décrocheurs, personnes en recherche d'emploi, en reconversion professionnelle, allocataires du RSA, réfugiés, primo-arrivants, résidents des quartiers prioritaires...) aux compétences numériques pour faciliter leur inclusion socioprofessionnelle et inclure chacun dans la digitalisation de la société.

Le programme Kesk'IA lancé en 2023 par la start-up « Evolukid » forme des jeunes à l'IA sans aucun prérequis de départ dans 15 territoires, notamment Lens, Rouen, Nantes, Trappes, Grigny, Meaux et Marseille¹³⁶. À Trappes, des jeunes de 16 à 24 ans ainsi formés ont pu développer une application appelée « C du propre », facilitant le tri des déchets : lorsqu'un citoyen prendra en photo un déchet ou un

134 Rapport « IA : notre ambition pour la France », Commission de l'IA, mars 2024.

135 Audition de Marlène Dolveck, Directrice générale de SNCF Gares & Connexions et Directrice Générale Adjointe du Groupe SNCF en charge de la Transformation et Julien Nicolas, Directeur Numérique du Groupe SNCF, par la Commission temporaire IA, le 3 octobre 2024.

136 Kesk'IA, un projet innovant autour de l'intelligence artificielle.

dépôt de déchet sauvage, l'intelligence artificielle permettra de localiser l'endroit exact et d'avertir la mairie.

Les formations (initiales, continues) peuvent aussi utiliser l'IA pour augmenter l'accessibilité des programmes de formation et d'emplois qui paraissaient jusqu'alors inaccessibles. Selon Jean Guo, fondatrice de l'association Konexio¹³⁷, l'IA peut justement permettre de personnaliser les parcours d'apprentissage, de renforcer l'accessibilité et de suivre les résultats. L'IA augmente l'engagement des apprenants de 47 % en personnalisant leur parcours, elle réduit le taux d'abandon des cours de 32 % grâce à des interventions adaptées et les performances scolaires s'améliorent de 23 % avec l'utilisation d'outils d'IA comme les tuteurs virtuels¹³⁸.

Selon Saïd Hammouche, entrepreneur social, président fondateur de la fondation Mozaïk et conseiller au CESE¹³⁹, les outils d'IA offrent une opportunité de

rattraper un manque de compétences, notamment de savoir-être (*soft skills*) ou de culture générale, qui ne sont pas nécessairement enseignées dans le cadre de l'enseignement traditionnel. Pour lui, « l'IA offre la possibilité d'accéder à des informations, à s'entraîner à tenir des discussions ou argumentations avec des interlocuteurs virtuels, contribuant à réduire l'impréparation et la marginalisation des publics qui n'ont pas l'habitude de ce type d'exercice. L'IA permet ainsi de compléter les compétences validées par le système académique classique, à condition, bien sûr, que les individus sachent se servir de l'IA ».

On pourrait rajouter à ces propos la condition que les individus conservent et exercent leur esprit critique face aux informations et aux réponses fournies par l'IA, ce qui impose sans doute une évolution des programmes scolaires en ce sens.

Les préconisations du CESE

Au même titre que la lutte contre l'illettrisme, la formation au numérique et à l'IA est un enjeu démocratique qui concerne chaque citoyenne et citoyen et peut s'adresser à toutes et tous, par exemple dans les médiathèques et les mairies, bien identifiés par les Françaises et les Français comme lieux d'accompagnement au numérique ou les « Cafés IA » déployés sur le territoire.

Comme le CESE l'a préconisé dans son avis *Mieux connecter les Outre-mer*, l'IA suppose de relever le défi de la formation notamment « dans les Conventions d'objectif conclues par France compétences et les Opérateurs de compétences (OPCO), des actions spécifiques concernant la formation au numérique en Outre-mer »¹⁴⁰.

¹³⁷ Audition de Jean Guo, Directrice et co-fondatrice de Konexio, par la Commission temporaire IA, le 3 octobre 2024.

¹³⁸ The Potential Impact of Artificial Intelligence on Education: Opportunities and Challenges | SpringerLink, mars 2024.

¹³⁹ Audition de Saïd Hammouche, entrepreneur social, président fondateur de la fondation Mozaïk et conseiller au CESE, par la Commission temporaire IA, le 3 octobre 2024.

¹⁴⁰ CESE, *Mieux connecter les Outre-mer*, avis 2024-016, rapporté par Danielle Dubrac et Pierre Marie-Joseph au nom de la délégation aux Outre-mer, octobre 2024. Mieux connecter les Outre-mer | Le Conseil économique social et environnemental.

Les pouvoirs publics, les entreprises, l'école, l'enseignement professionnel et supérieur doivent collaborer pour élaborer des programmes de formation continue axés sur les technologies de l'IA. C'est aussi un enjeu de formation professionnelle continue dont la responsabilité se partage entre des organismes privés et publics, qui ont à s'inscrire dans une logique de complémentarité plutôt que de concurrence pour une meilleure adaptation aux besoins des différents publics.

Recommandation de citoyennes et citoyens

L'accès à l'IA reste compliqué pour ceux qui ne maîtrisent pas le numérique, ce qui aggrave les inégalités. Pour y remédier il faut renforcer les formations pour tous les âges et s'appuyer sur des collaborations entre écoles, associations et entreprises. L'objectif est de donner à chacun les moyens de comprendre et d'utiliser ces technologies en fonction de ses besoins.

PRÉCONISATION #24

Le CESE appelle à soutenir la formation au numérique et à l'IA de la population à tous les âges de la vie, en :

- créant une complémentarité entre les différents acteurs de la société civile pour assurer cette obligation de formation ;
- intégrant l'IA, ses usages et les risques qu'elle peut présenter au plan national de formation des médiateurs numériques, des travailleurs sociaux et des agents publics en charge de l'accueil, de l'information, de l'orientation et de l'appui aux démarches des usagers.

E. L'IA dans l'éducation

Déjà largement utilisée par les élèves eux-mêmes, l'IA constitue une nouvelle réalité pour l'éducation mais cela pose plusieurs questions : comment le système éducatif doit s'adapter à l'arrivée des IA ? quels sont les risques des impacts de l'IA sur l'éducation ?

1. L'intégration de l'IA dans l'éducation aujourd'hui

L'usage de l'IA générative est déjà massivement développé chez les élèves et les étudiants.

Un sondage réalisé auprès de lycéens scolarisés en région Nouvelle-Aquitaine¹⁴¹ révèle que plus de 90 % des élèves de seconde avaient déjà utilisé l'IA générative pour les aider à faire leurs devoirs. Les jeunes utilisateurs y auraient recours pour la révision de leurs cours en utilisant par exemple Gamma pour la préparation d'exposés ou de présentations PowerPoint, Quizlet pour la création de fiches de révision ou Photomath pour la résolution d'exercices de maths à partir de la photo du problème. Parmi les outils d'IA générative, ChatGPT s'impose comme le modèle dominant, affichant une prévalence d'utilisation de 75 % des sondés, en particulier pour réaliser des synthèses, des traductions, voire la création de textes.

Le recours à l'IA varie avec l'âge : une enquête récente¹⁴² indique que près de 70 % des 18-24 ans l'utilisent personnellement contre 47% des 25-34 ans et seulement 22% des 35 ans et plus. Les pratiques évoluent très vite : en un an, le nombre d'utilisateurs aurait crû de 60 %.

Les avantages du recours à l'IA mais aussi les points de vigilance dans les usages sont identifiés par les élèves eux-mêmes. L'IA est utilisée dans l'aide à la rédaction ou les recherches mais les élèves plébiscitent aussi la facilité à interagir avec l'outil, sans crainte de ses réactions. Ainsi lorsqu'ils sont interrogés¹⁴³ sur leur perception des outils d'IA générative, les élèves relèvent « *la patience infinie de ces systèmes, leur bienveillance, et le fait qu'ils se mettent à leur portée, sans jugement de valeur* ».

¹⁴¹ Équipe Flowers de l'Inria, juin 2024. Pierre-Yves Oudeyer, *IA générative, société et éducation : En quoi l'IA générative représente-t-elle un enjeu dans la formation des citoyens ?*, septembre 2024 cité dans Sénat, Rapport IA et éducation : accompagner les développements en cours, 2024.

¹⁴² Ifop pour Talan, « Baromètre 2024 Les Français et les IA génératives », avril 2024 cité dans Sénat, Rapport IA et éducation : accompagner les développements en cours, 2024.

¹⁴³ Sénat, Rapport IA et éducation : accompagner les développements en cours, 2024.

Encadré 6 : Parole aux collégiens et lycéens sur leurs usages de l'IA : extraits choisis¹⁴⁴

Carlos D : « Moi j'utilise ChatGPT mais surtout Perplexity. Je crois que c'est plus complexe, mais ça nous donne des informations plus concrètes. C'est plutôt faire préparer des exposés et lorsque j'ai besoin d'une information que peut-être je ne pourrais pas trouver dans Google. »

Juliette D. : « C'est pratique pour la recherche et pour améliorer nos textes. Mais c'est vrai qu'à force d'utiliser l'intelligence artificielle, on se demande même comment on faisait avant, et pourtant ça a à peine quelques mois, c'est dire comment ça a très vite pris place dans la vie de n'importe quel élève. Par exemple, si on a une dissertation et que je pose ma problématique en demandant à ChatGPT de me développer cette dissertation sous la forme thèse-antithèse, il fera ça très bien. Vous pouvez même demander le nombre de lignes. »

Juliette D. « L'IA n'est pas toujours précise. Il faut savoir poser les bonnes questions. »

Carlos T. « Elle peut rendre les élèves dépendants et diminuer la valeur du travail individuel. »

Juliette D. : « Le problème vient de la façon dont on utilise l'IA, qui peut être malveillante. »

Juliette D. : « En fait, il serait utile d'avoir des cours sur l'usage éthique et efficace de l'IA ».

En revanche l'usage de l'IA par les enseignants en est encore au stade de l'expérimentation. Si plus de la moitié des enseignants disent connaître l'IA, son usage est encore peu développé au-delà d'initiatives personnelles.

Dans un sondage¹⁴⁵, 56 % des enseignants de l'enseignement supérieur disaient connaître les outils d'IA sans jamais les utiliser et 9 % ne les connaissaient pas. 8 % utilisent régulièrement l'IA générative et 27 % occasionnellement.

Au sein de l'Éducation nationale, le recours des enseignants à l'IA résulte encore d'initiatives individuelles ou d'expérimentations à petite échelle. En effet, selon les informations collectées, les outils institutionnels d'IA demeurent globalement méconnus et les usages avancés sont rares. L'intégration des outils d'IA dans la pédagogie reste largement dépendante des goûts et intérêts personnels de chacun.

Ainsi, par exemple dans les Alpes de Haute-Provence, une professeure de lettres a mené un travail intégrant l'IA au collège avec des élèves de 4ème.

¹⁴⁴ Cahiers pédagogiques, Comment on faisait avant ? IA et pédagogie, mai 2024.

¹⁴⁵ Institut Sphinx pour le logiciel Compilatio en novembre 2023, cité dans Sénat, Rapport IA et éducation : accompagner les développements en cours, 2024.

Elle indique¹⁴⁶ : « il est très important d'inclure l'IA à l'école dans toutes ses dimensions, et tout d'abord en tant qu'outil. Il ne faut pas se leurrer sur le fait que de plus en plus d'élèves utilisent des IA pour faire leurs devoirs et effectuer des recherches. Plus vite on apprend à l'élève à utiliser plus intelligemment cet outil, plus vite on évitera ces écueils : leur apprendre qu'une IA fournit un « premier jet » à retravailler, à vérifier, à enrichir, à reformuler et à structurer. C'est une aide, non une finalité, tout comme une calculatrice en mathématiques. C'est donc bien à l'école d'éduquer les enfants à cet outil, en leur montrant ses atouts et ses limites ». Quelques sites proposant une IA sont devenus accessibles dans son collège (TalkAI par exemple), ce qui va lui permettre de refaire les mêmes expériences en projetant les réponses aux élèves, plutôt qu'en utilisant son téléphone personnel comme elle a dû le faire

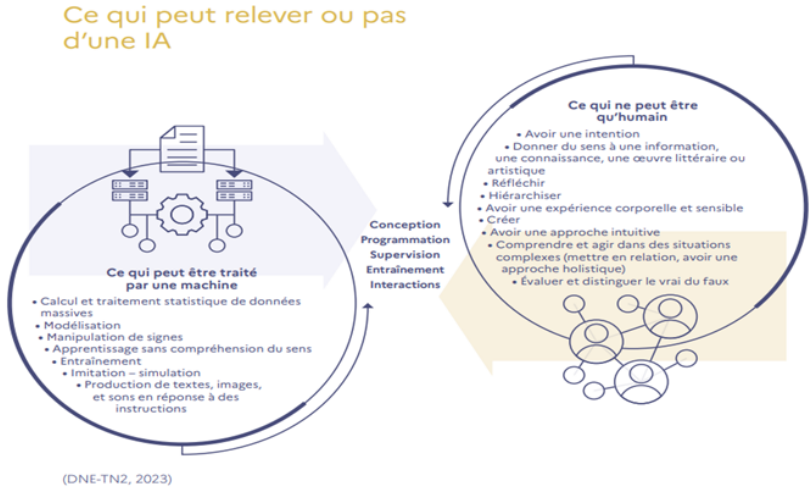
précédemment.

L'irruption de l'IA amène à s'interroger sur les compétences relevant de l'être humain et celles qui relèvent de la machine.

Delphine Le Serre indique que l'irruption de l'IA nous amène à « nous interroger sur ce qu'est un être humain par comparaison avec ces machines intelligentes. Qu'est-ce qui fonde notre humanité ? Et comment préserver ces fondements et développer les spécificités humaines ? ». Cela amène à un « tournant dans la mission de l'éducation » avec « le besoin d'une prise de conscience sur la nécessité de faire évoluer dès maintenant le modèle de l'éducation du XXI^e siècle ». Un schéma définit une première ligne de partage entre ce qui peut être traité par une machine et ce qui ne peut être qu'humain.

¹⁴⁶ Stéphanie Marquès : L'intelligence des élèves plus forte que l'intelligence artificielle !, 2024.

FIGURE 5



Sources : Ministère de l'Éducation nationale et de la recherche, IA et éducation, enjeux de la recherche et de la politique publique, janvier 2024

Delphine Le Serre met l'accent sur les compétences humaines qui différencient l'être humain de la machine : « *l'esprit critique, l'intuition, la perception sensorielle, l'intelligence socio-émotionnelle (gérer les émotions des autres en interaction avec les siennes), la créativité, l'imagination. C'est là que se situera la valeur ajoutée de l'être humain* ».

Pour faciliter la compréhension du monde, la différence entre ce que peut faire l'être humain et ce que peut faire la machine est à aborder dès le plus jeune âge, par exemple dans les ateliers philosophiques organisés dans certaines classes en maternelle qui ont pour objectif d'amener les enfants à développer un sens critique, à apprendre à se poser

des questions et à s'ouvrir au débat. La formation à l'esprit critique est cruciale et l'IA peut constituer un outil permettant de renforcer l'esprit critique lorsque des enseignants travaillent avec leurs élèves ou étudiants pour leur apprendre à poser des questions sous différentes formes et comparer les réponses, les vérifier avec une lecture critique pour repérer les erreurs. Philippe Meirieu¹⁴⁷, indique que « *ce travail, qui relève de la nécessaire formation à l'esprit critique, devrait être fait dans toutes les classes, de l'école primaire à l'université, et nul doute que les élèves et les étudiants en tireraient des bienfaits considérables : en termes de lucidité sur les productions de l'intelligence artificielle mais aussi*

147 Philippe Meirieu, Enseignement : qui a peur des robots conversationnels ?, 2024.

en termes d'exigence pour leurs propres écrits ».

Par ailleurs, l'OCDE¹⁴⁸ a évalué en 2023 les capacités de l'IA et de leurs implications pour le travail et l'éducation. L'augmentation des investissements dans la recherche et le développement de l'IA conduira à des progrès significatifs de l'IA à la fois en lecture et en mathématiques dans les années à venir et « l'IA pourrait surpasser une grande partie de la population en lecture et en mathématiques ».

À l'avenir, les travailleurs risquent d'être de plus en plus concurrencés par les machines dans ces domaines de compétences, l'irruption de l'IA nécessite donc de renforcer les compétences de base et de préparer les futurs travailleurs à l'utilisation de l'outil IA. L'OCDE souligne aussi l'importance d'un suivi périodique et systématique de l'évolution des capacités de l'IA et leur comparaison avec les compétences humaines¹⁴⁹.

Cependant, les apprentissages fondamentaux, les capacités réflexives et l'esprit critique seront toujours nécessaires pour que l'être humain, avec son expertise technique, garde le contrôle de la machine, avec un vocabulaire précis pour « prompter » correctement en amont et les connaissances adéquates pour vérifier en aval ce que la machine a produit. L'utilisation efficace et la compréhension des IA génératives, comme les modèles de langage avancés, requiert une maîtrise des apprentissages fondamentaux en langues, mathématiques et sciences.

Des outils d'IA sont mis à disposition par le ministère mais restent peu connus.

Un premier marché public innovant a été lancé en 2018 par le ministère de l'Éducation nationale, le Secrétariat général pour l'investissement et la Banque des Territoires pour développer avec les EdTech des solutions basées sur l'IA pour l'apprentissage du français et des mathématiques au cycle 2 (CP, CE1, CE2) et soutenir la souveraineté numérique. Ce marché a été suivi d'un deuxième marché public (« PIA2 ») pour développer des solutions basées sur l'IA pour le français, les mathématiques mais aussi les langues étrangères pour le cycle 3 (CM1-CM2-6ème). Les projets sont conçus grâce à une phase de recherche-développement menée en classe dans des écoles volontaires.

5 projets sont déployés pour le français et les mathématiques¹⁵⁰ :

- pour le français : LALILO (outil d'aide aux élèves pour l'apprentissage de la lecture utilisant la reconnaissance vocale) NAVI (outil d'assistance pédagogique pour les enseignants avec une plateforme en ligne fournissant des parcours personnalisés et adaptatifs) ;
- pour les mathématiques : ADAPTIV'MATH (outil d'assistance à l'enseignant permettant de mettre en place un parcours par groupe d'élèves, d'évaluer et analyser les résultats) MATHIA (outil d'aide à l'élève lui permettant d'interagir à l'oral avec l'IA pour découvrir des concepts mathématiques, être entraîné et encouragé, seul ou en groupe) SMART ENSEIGNO (outil pour l'enseignant dans sa classe proposant des activités

¹⁴⁸ OCDE, Intelligence Artificielle et les compétences de demain, 2023.

¹⁴⁹ Ministère de l'Éducation nationale et de la recherche, IA et éducation, enjeux de la recherche et de la politique publique, janvier 2024.

¹⁵⁰ P2IA : Partenariat d'innovation et d'intelligence artificielle.

d'apprentissage variées en mathématiques sur une plateforme avec tableau de bord).

Toutefois, la montée en puissance de ces outils reste à consolider si on se réfère au nombre de licences utilisées pour évaluer le nombre d'utilisateurs au printemps 2024 : 25 428 pour LALILO, 6431 pour MATHIA, 5822 pour ADAPTIV'MATH, 5044 pour NAVI, 2978 pour SMART ENSEIGNO¹⁵¹. Ces chiffres peuvent aussi refléter l'utilisation limitée d'outils clés en main, diffusés, voire imposés à l'ensemble des enseignants alors qu'il y aurait une demande de développer leur pouvoir d'agir et d'approfondir leur liberté pédagogique en partant de leurs besoins.

Les associations d'éducation complémentaire de l'enseignement public proposent également des parcours d'éducation aux médias et à l'information (Ligue de l'enseignement, CEMEA, etc.). Ces parcours visent moins une intégration de l'IA dans les pratiques pédagogiques qu'ils ne favorisent la capacité des élèves et des enseignants à se repérer dans les évolutions de l'IA.

2. L'IA au service de l'éducation : risques et opportunités

Les élèves relèvent plusieurs points de vigilance dans les usages de l'IA non anticipés et non encadrés avec des risques sur le déroulement du cursus.

2.1 UN OUTIL QUI PEUT ÊTRE UTILISÉ COMME « ASSISTANT PÉDAGOGIQUE DE L'ÉLÈVE » OU « COACH DE L'ÉTUDIANT », PERSONNALISANT L'APPRENTISSAGE

L'IA peut être utilisée comme « assistant pédagogique de l'élève » avec une utilisation « raisonnée » et « appropriée ». Par exemple une IA générative telle que ChatGPT peut fournir une assistance à l'écriture : corriger l'orthographe et la grammaire d'un texte pour des élèves débutants avec une instruction du type : « *Vérifie le texte suivant pour corriger l'orthographe, la grammaire et les erreurs de ponctuation.* » ; améliorer la clarté et la cohérence d'un texte rédigé par des élèves avancés avec une instruction du type « *Suggère moi des pistes d'améliorations pour retravailler ce paragraphe afin qu'il soit plus clair* » ; affiner la logique des arguments, restructurer des paragraphes pour des élèves experts avec une instruction du type « *Voilà un plan sommaire sur [tel sujet]. Affine mes premiers arguments et génère un brouillon* »¹⁵².

Un impact possible de l'IA sur l'éducation est la personnalisation des contenus selon Delphine Le Serre, professeure et dirigeante du fonds EdHu2050¹⁵³. Certaines universités aux États-Unis ou en Australie ont mis en place des agents conversationnels qui permettent à chaque étudiant d'avoir un « coach personnel artificiel » lui indiquant des révisions à faire en fonction de ses insuffisances sur tel sujet, proposant

151 Sénat, IA et avenir du service public, 2024.

152 L'intelligence artificielle dans l'éducation | Académie de Paris.

153 EdHu2050 - Le mouvement d'une Éducation centrée sur l'Humain.

des sources bibliographiques ad hoc et l'orientant vers les bons interlocuteurs.

L'irruption d'outils d'IA générative tels que ChatGPT donne accès aux élèves à un outil gratuit d'aide aux devoirs, immédiatement disponible. Cela contraste avec une source majeure des inégalités scolaires : la disparité préexistante d'accès aux ressources telles que l'aide des parents, des frères et sœurs, l'appel à des cours privés ou professeurs particuliers. Cette source d'inégalités en dehors de l'école est traditionnellement renforcée quand des devoirs à faire à la maison sont donnés aux élèves¹⁵⁴. En revanche, comme nous avons pu le constater pendant le confinement lié à la crise sanitaire du COVID, ces inégalités peuvent être accentuées si l'élève n'a pas accès chez lui à un outil numérique.

2.2. POUR LES ENSEIGNANTS, L'IA POURRAIT CONSTITUER UN OUTIL DE SOUTIEN DANS LEURS PRATIQUES HABITUELLES MAIS AUSSI UN MOYEN DE SE RECENTRER SUR LA RELATION PÉDAGOGIQUE

L'IA peut pourtant soutenir les enseignants dans leurs pratiques pédagogiques, qu'il s'agisse de la préparation des cours : générer des supports de cours, fournir des suggestions (« *J'ai dit à ChatGPT que l'apprenant souffrait de dyslexie, et finalement ChatGPT a recommandé du matériel d'apprentissage spécifique pour l'apprenant* »), traduire du matériel pédagogique dans différentes langues ou de l'évaluation des élèves : générer des tâches d'évaluation (des exercices, des quiz et des scénarios qui peuvent être utilisés en classe pour faciliter la pratique

et l'évaluation), noter les dissertations des apprenants¹⁵⁵.

Pour la correction des copies, considérée comme une des tâches les plus laborieuses et rébarbatives, des solutions techniques apparaissent. Néanmoins, il serait pertinent d'expérimenter l'utilité d'enlever au professeur ces tâches afin de vérifier les impacts dans la relation pédagogique. L'évaluation des résultats de l'apprentissage fait partie des systèmes d'IA identifiés « à haut risque » par l'AI Act donc avec des obligations de transparence et d'explicabilité à respecter d'ici le 2 août 2027.

La peur de la « triche » appellerait cependant à repenser les modes d'évaluation des élèves.

En juillet 2023, sur les réseaux sociaux, un nombre important d'élèves se sont vantés d'avoir obtenu une bonne note au grand oral en se contentant de livrer à l'oral un texte écrit par ChatGPT. Recourir à l'IA peut permettre de dépasser « l'angoisse de la page blanche » mais ce type d'usage peut aussi être vu comme une manière d'éviter le travail. Dans les médias, certains enseignants ont dénoncé une nouvelle forme de triche quand d'autres ont rappelé que ce n'est pas le candidat qui doit être mis en cause mais la légitimité même de cet exercice.

Au quotidien, les non-dits sur l'IA générative créent chez les enseignants une suspicion de fraude de la part des élèves alors que les élèves, eux, craignent d'être injustement accusés d'avoir triché. Ce non-dit nuit à la relation pédagogique et invite les professeurs à revoir leurs façons d'évaluer les apprentissages des élèves pour s'adapter à ce nouveau

¹⁵⁴ Mickael Bertrand, Dépasser la crainte de la triche ; Alexandra Maurice, La question n'est pas la bonne, peut-être ? ; Cahiers pédagogiques, Intelligence artificielle et pédagogie, mai 2024.

¹⁵⁵ Ministère de l'Éducation nationale et de la recherche, IA et éducation, enjeux de la recherche et de la politique publique, janvier 2024.

contexte intégrant l'IA.

Dans une tribune¹⁵⁶, Philippe Meirieu, professeur honoraire en sciences de l'éducation indique que « certains professeurs préfèrent s'emparer de l'outil pour travailler avec leurs élèves ou leurs étudiants. Ils leur apprennent à poser des questions sous différentes formes pour comparer les réponses ; ils confrontent ces réponses avec celles des manuels et des encyclopédies ; ils les aident à repérer les glissements sémantiques qui induisent des malentendus et compromettent l'objectivité du texte ; ils utilisent les propositions de ChatGPT comme des brouillons à compléter et à personnaliser, ou encore les font traduire sous d'autres formes textuelles, graphiques ou visuelles. Ces professeurs-là ne s'inquiètent guère que leurs élèves ou leurs étudiants leur remettent des textes issus de ChatGPT : ils les accompagnent suffisamment dans le processus d'écriture de textes pour voir ces derniers se construire progressivement sous leurs yeux. Et, si nécessaire, ils sont parfaitement capables, en une ou deux questions orales, de vérifier qu'ils en sont bien les auteurs ». L'arrivée de l'IA est perçue comme une occasion de repenser les épreuves d'évaluation en donnant une plus large part à l'invention et à l'expression personnelle.

Repenser l'évaluation permettrait de mettre l'accent sur le processus d'apprentissage et non sur le produit final, en valorisant les étapes de

recherche, de reformulation, de tâtonnement jusqu'au résultat individuel ou collectif. À l'ère de l'IA, l'évaluation des élèves pourrait donc ne plus prendre la forme d'examens de contrôle mais plutôt de situations dans lesquelles ils agissent et grâce auxquelles ils peuvent s'approprier des outils de manière constructive et responsable.

2.3 DES RISQUES À IDENTIFIER EN AMONT POUR LES ÉLÈVES ET LEURS PROFESSEURS

Le temps d'écran : l'émergence de l'IA dans le monde éducatif s'inscrit dans un contexte de vigilance accrue sur les effets de la surexposition aux écrans des enfants. La commission de travail missionnée pour évaluer l'impact de l'exposition des jeunes aux écrans, a remis son rapport¹⁵⁷ au Président de la République le 30 avril 2024 avec un guide pour le bon usage des écrans à la maison ou en classe. Face à « l'hyperconnexion » néfaste pour les enfants et la société en général, la commission souligne qu'il n'existe pas « d'addiction aux écrans » reconnue par la science mais que les écrans et réseaux sociaux sont des facteurs de risque supplémentaires s'il y a une vulnérabilité préexistante chez un enfant ou adolescent. Pour un usage des écrans adapté au développement de l'enfant, cette commission a formulé 29 préconisations avec une exposition aux écrans progressive selon l'âge avec une absence d'exposition aux écrans avant 3 ans, un usage déconseillé jusqu'à six ans,

¹⁵⁶ Philippe Meirieu, Le danger de ChatGPT n'est pas dans la fraude qu'il permet mais dans le rapport aux connaissances qu'il promeut, *Le Monde*, mars 2023.

¹⁵⁷ Proposition de loi, n° 406 visant à réguler l'usage des écrans en présence des enfants de moins de trois ans - 17e législature - Assemblée nationale.

ou limité, occasionnel, en privilégiant les contenus éducatifs avec un adulte et pas d'accès aux réseaux sociaux avant 15 ans. Une proposition de loi a été déposée le 15 octobre 2024 pour interdire l'exposition aux écrans des enfants de moins de 3 ans, qu'ils soient gardés par des assistantes maternelles ou en crèches.

S'agissant de l'IA spécifiquement, plusieurs limites et points de vigilance ont été relevés quant au recours à cette technologie dans l'éducation¹⁵⁸ :

- la possible fragilisation de la relation pédagogique ;
- la péremption des données utilisées par les modèles d'IA ;
- la présence de biais dans les résultats ;
- les erreurs factuelles et possibles « hallucinations » nécessitent quasi systématiquement une vérification des contenus par l'utilisateur, ce qui peut constituer un atout pour développer l'esprit critique et établir une distance avec les contenus générés ;
- le risque de produire et de diffuser des infox ou des informations tronquées, erronées ou non vérifiables ;
- la difficulté du référencement et des sources, de la paternité des contenus ou du droit d'auteur, voire de l'absence ou des erreurs sur les sources.

Plus largement, le contexte actuel de l'Éducation nationale et le manque de moyens dédiés à l'exercice du métier font peser sur l'IA le risque « *d'évacuer la question de leurs conditions matérielles d'exercice* » et que « *technologie et scientisme [soient] mis au service d'une dépossession du métier.* » Des

représentants du personnel¹⁵⁹ craignent en effet « *qu'à partir d'évaluations nationales standardisées (qui ne sont pas conçues par les enseignants), il n'y ait plus qu'à suivre des procédures normées (elles aussi définies par d'autres), au travers d'applications « mises à disposition » (par le Ministère et une EdTech). La liberté pédagogique n'est plus maintenue que comme une fiction : l'enseignant pourrait encore « choisir » en fonction de son expertise, d'utiliser ou pas l'application. Mais de quel choix parle-t-on, dans un contexte de dégradation des conditions de travail, de management de plus en plus injonctif et de suppressions de postes ?* »

En prenant encore un pas de recul, il apparaît que d'autres enjeux globaux relatifs à l'éducation découlent du déploiement de l'IA dans nos sociétés : elle transforme la nature des connaissances à transmettre et les manières de le faire ; interroge la place des savoirs au sens large ; ou encore pose la question de la confiance envers l'IA, mais aussi envers les autres êtres humains et nous-mêmes, compte-tenu de la possibilité de « vérifier » en permanence les informations, de manière encore plus précise et rapide qu'avec internet.

Étant donné le contexte dans lequel évolue l'Éducation nationale et le fait que l'usage massif est déjà développé chez les élèves, la question se pose aujourd'hui de la place à accorder à l'IA dans le système éducatif et surtout de la manière de le faire.

158 Ministère de l'Éducation nationale et de la recherche, IA et éducation, enjeux de la recherche et de la politique publique, janvier 2024.

159 IA génératives : pour un principe de précaution en éducation - SNES-FSU, décembre 2023.

3. Comment le système éducatif peut s'adapter à l'IA en partant des besoins et des choix de terrain

Au Québec, l'université Laval a travaillé sur cinq principes directeurs concernant l'intelligence artificielle dans l'enseignement et l'apprentissage, qui peuvent inspirer et guider des réflexions à venir en France :

- le premier principe rappelle la primauté de la relation humaine, à la fois au cœur de la relation pédagogique et condition essentielle à l'utilisation de l'IA ;
- le deuxième principe est celui «de l'ouverture, l'exploration et la formation» et vise principalement à encourager les postures d'ouverture à propos de la présence de l'IA ;
- le troisième vise quant à lui à encourager la prudence et une utilisation éclairée de l'IA qui s'appuie sur une posture réflexive et critique ;
- le quatrième principe vise l'intégrité et la transparence dans l'usage de l'IA ;
- enfin le cinquième rappelle que l'intégration de l'IA ne peut se faire que dans le cadre d'un travail collaboratif au sein d'équipes pédagogiques.

3.1 S'APPROPRIER LES ENJEUX DE L'IA AU SEIN D'UN PROJET D'ÉTABLISSEMENT PEUT CONSTITUER UNE PREMIÈRE ÉTAPE

Le lycée international Charles de Gaulle de Dijon¹⁶⁰ a inscrit dans le projet d'établissement l'objectif de « *Promouvoir le développement professionnel des personnels en s'appropriant les enjeux de l'intelligence artificielle* ». Pour désamorcer les craintes des personnels, il a fallu expliquer en conseil d'administration que l'objectif n'était pas d'en généraliser l'usage. La conduite d'un travail d'appropriation des enjeux au lycée a permis de donner à chacun la capacité de se forger une opinion en maîtrisant l'outil, son usage, ses ressources techniques, ses avantages et dangers. Une enquête a fait émerger deux types de besoins : un accompagnement par équipes disciplinaires (pour 59 % des enseignants ayant répondu) et des ateliers pratiques (52 %). Cela s'est traduit par deux actions concrètes :

La diffusion d'une lettre d'information aux membres de l'équipe éducative et aux élèves avec des contenus différents et adaptés à chacun des publics avec un objectif commun : apporter des outils, des pratiques et des mises en perspective sur les usages des IA génératives en éducation.

Ces apports sont ensuite développés lors d'ateliers d'accompagnement pour expérimenter les possibilités offertes par les IA génératives (découverte de l'outil ChatGPT, notions de base pour rédiger un

¹⁶⁰ Dominique Javel, Un défi collectif, dans les cahiers pédagogiques, IA et pédagogie, mai 2024.

prompt, création d'activités différenciées, élaboration de questionnaires à choix multiple...) avec deux ateliers distincts organisés pour les enseignants d'une et les élèves d'autre part en fonction des usages de chacun. La participation n'est pas encore massive mais l'intérêt et la curiosité sont présents.

Dans l'enseignement supérieur en France, certaines des grandes écoles¹⁶¹ ont mis en place des outils d'apprentissage adaptatif et des stratégies de formation à l'IA de l'ensemble des étudiants et des enseignants. Cependant, ces développements restent encore l'apanage de quelques établissements précurseurs disposant des moyens adéquats.

3.2 INTÉGRER L'IA DANS LA FORMATION INITIALE ET CONTINUE DES ENSEIGNANTES, ENSEIGNANTS ET L'ENSEMBLE DE LA COMMUNAUTÉ ÉDUCATIVE PARAÎT « URGENT »

Le programme national de formation (PNF) 2024-2025 des personnels du ministère de l'éducation, de la jeunesse et des sports ne mentionne l'IA que pour un colloque « IA et éducation »¹⁶².

Il existe cependant un cours en ligne¹⁶³ (« IA pour et par les enseignants »), développé et expérimenté auprès d'enseignants et chefs d'établissement dans le cadre du projet européen AI4T, réunissant cinq ministères de l'éducation de cinq pays dont la France pour élaborer les contenus. Ce programme de 3 heures vise à aider les enseignants « et la communauté éducative dans son ensemble » à comprendre les usages, technologies et limites de l'IA dans le contexte éducatif pour un usage éclairé.

Il doit permettre d'identifier le potentiel et les utilisations actuelles de l'IA en éducation, de comprendre ce qu'est réellement l'IA, d'expérimenter différentes techniques d'IA, d'identifier la place cruciale des données pour les outils éducatifs basés sur l'IA et d'appliquer une méthodologie pour analyser un outil pédagogique basé sur une IA. Sans imposer une méthode-type, ce cours a au contraire été conçu « pour aider les participants à se faire leur propre opinion sur les questions suivantes : À quoi sert l'IA dans le contexte de l'éducation ? Quel type d'IA pour quel objectif ? Quels sont les avantages et les défis ? Comment utiliser les ressources basées sur l'IA ? Quels impacts peut-on prévoir dans les pratiques d'enseignement et d'apprentissage ? ».

On constate en France une réelle demande de formation chez les enseignants. Pour être adaptée aux besoins, la formation doit être opérationnelle et s'inscrire dans une réflexion pédagogique plus large : pratiquer, rédiger des prompts, déceler les hallucinations, comprendre comment fonctionnent les grands modèles de langage, ainsi que les modèles de diffusion comme Midjourney ou Flux, programmer soi-même des assistants.

La formation doit également s'adresser à l'ensemble de la communauté éducative, notamment tous les adultes accueillant des jeunes dans différents espaces éducatifs (accueils collectifs de mineurs, colonies de vacances, clubs sportifs).

¹⁶¹ Sénat, Rapport IA et éducation : accompagner les développements en cours, 2024.

¹⁶² Le Programme national de formation 2024-2025 | Ministère de l'Éducation Nationale, de l'Enseignement supérieur et de la Recherche.

¹⁶³ Direction du numérique pour l'éducation.

3.3 ADAPTER ET RENFORCER LES COMPÉTENCES À ACQUÉRIR PAR L'ÊTRE HUMAIN À L'ÈRE DE L'IA

Il s'agit également de considérer que l'IA doit permettre à l'humain d'améliorer ses compétences et non pas seulement de les augmenter temporairement. À ce titre, l'augmentation temporaire agit comme une béquille, il s'agit de pratiques visant à répondre efficacement à une question, en déléguant par ailleurs le rapport à la connaissance et au savoir à la technologie. Dans ce cas de figure, aucun gain à long terme n'est obtenu et l'humain régresse sur le plan des connaissances et des compétences.

En revanche, dans une démarche d'amélioration, l'IA peut permettre à l'humain de renforcer ses compétences de long terme, sa capacité à comprendre et acquérir de nouvelles connaissances lui permettant d'enrichir son rapport au monde. Il s'agit donc de faire en sorte que l'usage de l'IA dans l'éducation soit centré sur le développement des compétences et connaissances, et non sur la capacité de l'élève à répondre correctement à une réponse avec l'aide de l'IA. L'usage de l'IA doit permettre à l'élève ou l'étudiante ou étudiant de répondre sans l'IA à l'issue du processus d'apprentissage.

Les préconisations du CESE

- **Évaluation des apprentissages**

L'IA est actuellement déjà utilisée massivement par les élèves et étudiants eux-mêmes. Son usage peut transformer le rapport au savoir mais aussi le mode d'évaluation des connaissances et compétences. Il importe de tenir compte du fait qu'il existe à disposition des élèves des robots conversationnels qui peuvent être gratuits, capables de collecter des données et de rédiger des réponses en un temps record. Pour cela, dans une logique de transparence, il doit être fait obligation à toutes et tous (et pas seulement aux élèves) de signaler tout contenu produit par l'IA. Cela nécessite aussi de valoriser davantage, dans le travail scolaire, la démarche que le résultat, soit toutes les étapes du processus et de l'effort d'apprentissage, y compris la capacité à travailler collectivement plutôt que les formes

d'individualisation systématiques.

PRÉCONISATION #25

Le CESE recommande de :

- **repenser les modes d'évaluation des apprentissages des élèves et étudiants en intégrant le fait que l'IA existe ;**
- **développer la capacité des élèves et étudiants à apprendre collectivement.**

- **Formation des enseignants**

Alors que les élèves utilisent déjà massivement l'IA, seuls 8 % des enseignants l'utilisent régulièrement et 27 % de manière occasionnelle. Il y

a un besoin de formation et la demande de formation est réelle chez les enseignantes et enseignants.

Cependant, s'il existe quelques formations embryonnaires, tel un MOOC de 3 heures, la formation du personnel éducatif reste à construire. Ainsi le programme national de formation 2024-2025 des personnels du ministère de l'éducation, de la jeunesse et des sports comprend un axe 6 « *Consolider les connaissances, les compétences et les usages du numérique* », l'intelligence artificielle n'y apparaît que très peu¹⁶⁴.

PRÉCONISATION #26

Le CESE préconise de concevoir et déployer une formation à l'IA initiale et continue prioritairement pour les enseignants mais aussi pour l'ensemble de la communauté éducative. Il appelle à s'appropriier les enjeux de l'IA au sein d'un projet d'établissement.

• Orientation et formation des élèves

Si les élèves utilisent déjà par eux-mêmes largement l'IA, avoir un usage raisonné et éthique suppose une « éducation à l'IA ». Être formé à l'IA pour pouvoir s'insérer au mieux sur le marché du travail et comprendre le contexte dans lequel on s'insère paraît indispensable. Actuellement, la plateforme Pix¹⁶⁵ permet de développer et tester par soi-même ses compétences numériques tout au long

de la vie. Une certification « Pix » de ses compétences peut être obtenue après passage d'un test adaptatif (sans niveau minimum mais avec des questions dont la difficulté augmente en fonction des réponses pour pousser le candidat au niveau maximum) permettant de mesurer son niveau sur les 16 compétences numériques sur une échelle allant du niveau « novice » à « expert ».

Le passage de la certification Pix est obligatoire pour les élèves de troisième mais cet examen peut aussi être réalisé tout au long de la vie. Au vu de l'essor de cette technologie, intégrer l'IA dans les compétences numériques à développer et à certifier en ligne de cette manière serait particulièrement pertinent pour les élèves mais serait aussi ouvert à toutes et tous avec une plateforme Pix prévue pour être utilisée tout au long de la vie.

¹⁶⁴ Le Programme national de formation 2024-2025 | Ministère de l'Éducation Nationale, de l'Enseignement supérieur et de la Recherche.

¹⁶⁵ Accueil | Pix.

PRÉCONISATION #27

Le CESE recommande de développer la formation des élèves à ce qu'est l'IA, à son fonctionnement, à ses capacités et à ses risques mais aussi des citoyennes et citoyens en déployant également un accompagnement humain via des formateurs.

Recommandation de citoyennes et citoyens

Pour les élèves, l'IA peut les aider dans leurs devoirs mais cela pose des problèmes de triche et de dépendance, et plus fondamentalement du rapport au savoir. Pour les enseignants, l'IA peut faciliter la préparation des cours, mais elle risque de réduire l'interaction humaine et peut introduire des biais. Le CESE recommande de former les enseignants et les élèves à une utilisation responsable de l'IA, et de repenser les méthodes d'évaluation pour encourager l'apprentissage plutôt que la simple performance.

F. Santé et intelligence artificielle

L'IA se déploie et peut apporter des améliorations dans plusieurs domaines de l'écosystème de santé : les soins (prévention, diagnostics, traitements), la prestation (prestataires), la gestion du dossier patient (organismes payeurs), la recherche médicale... Cependant, elle nécessite d'intervenir en amont pour limiter les biais induits par les algorithmes et les bases de données, ce qui rend encore plus impératif qu'elle soit utilisée de manière éthique et équitable pour bénéficier à toutes et tous.

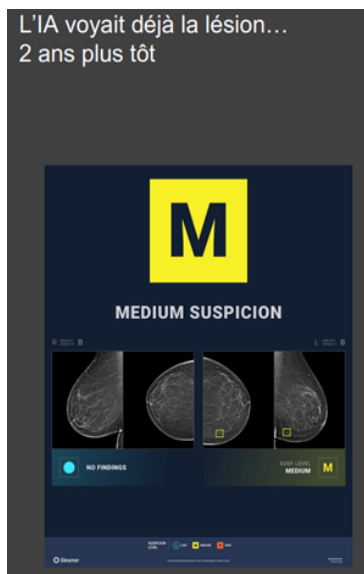
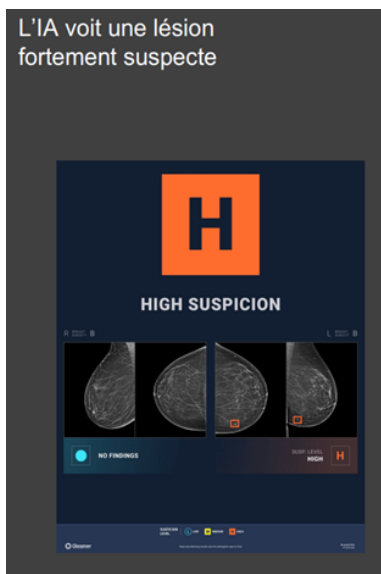
L'IA peut également améliorer et accélérer la détection des cancers du sein, le plus fréquent chez les femmes en France, et leur traitement. Elle permet d'intervenir dès la première lecture des mammographies, d'éviter les « faux négatifs », et de repérer de manière anticipée des lésions, comme le montre l'exemple suivant.

1. Amélioration de la qualité des soins, de l'accès aux soins et de la prévention, sans remplacer la relation soignant-patient

1.1 L'IA PEUT CONSTITUER UN APPUI UTILE POUR AMÉLIORER LE SOIN, DANS LE CADRE DU DIAGNOSTIC, DE LA PERSONNALISATION DE TRAITEMENTS OU D'OPÉRATIONS CHIRURGICALES.

À titre d'exemple, les diagnostics en matière de détection des lésions et fractures sont améliorés grâce au traitement des images par l'IA. Dans les services d'urgence, où près de 20 % des patients viennent pour des fractures, l'IA permet d'analyser les images, de produire un rapport automatique, de réduire le temps de lecture de 35 % et celui de passage aux urgences pour traumatologie de 1 heure¹⁶⁶. Ce gain de temps peut être mis au service des patients (information et explication) et des personnels en améliorant la qualité de vie et la santé au travail.

¹⁶⁶ Audition d'Albane Grandjean, externe en médecine et consultante de la plateforme Gleamer assistant les radiologues, par la Commission temporaire IA, le 10 octobre 2024.



Source : audition Albane Grandjean, Gleamer, 10 octobre 2024

L'IA peut permettre de développer une médecine prédictive¹⁶⁷ et d'adapter le traitement au plus près du patient. Une IA développée par l'entreprise OWKIN¹⁶⁸ permet sur la base d'une simple analyse de lame, d'identifier si un patient va répondre à l'immunothérapie et si le cancer est agressif ou pas. Il s'agit d'un marqueur prédictif de la capacité de traitement qui permet de personnaliser la médecine en traitant mieux et plus tôt le cancer. Un autre exemple est celui du laboratoire d'informatique médicale et d'ingénierie des connaissances en e-santé (LIMICS, unité Inserm 1142)

de l'APHP¹⁶⁹, qui a participé au projet européen Desiree, utilisant l'IA pour aider les cliniciens dans le traitement et le suivi des patientes atteintes de cancers du sein, maladies complexes, nécessitant souvent une adaptation des protocoles. L'enrichissement continu de la base de cas permet de faire évoluer les propositions faites par le système pour l'aide à la prise en charge thérapeutique des patientes. Enfin, selon le professeur Jean-Emmanuel Bibault¹⁷⁰, l'IA peut replacer le patient au centre du protocole et « remettre la qualité des soins au centre des

167 Selon Larousse, la médecine prédictive est une partie de la médecine qui s'attache à rechercher les risques génétiques que présente un individu d'être victime au cours de son existence d'une maladie. La médecine prédictive étudie les facteurs de risques d'apparition des maladies au stade le plus précoce possible, néonatal ou même prénatal. Elle est donc intimement liée à la génétique.

168 Audition du Professeur Benjamin Besse, oncologue et directeur de la recherche clinique de l'Institut Gustave Roussy, par la Commission temporaire IA, le 10 octobre 2024.

169 Inserm, Intelligence artificielle et santé. La science pour la santé.

170 Les Échos, Santé : comment l'IA va révolutionner la médecine, novembre 2024.

préoccupations ».

L'IA peut aussi être utile aux différents stades d'une opération chirurgicale : de la planification à la modélisation jusqu'à la réalisation. Des hôpitaux de l'AP-HP ont par exemple fait entrer au bloc opératoire « Connected Surgery », une plateforme d'assistance chirurgicale basée sur l'IA : disponible sur écran pendant une opération, la plateforme assiste l'équipe du bloc opératoire tout au long de la procédure, anticipe les risques et alerte en cas d'incident. Elle peut avertir les médecins de la proximité de certaines artères. Elle fournit une aide à la décision pendant l'opération en s'appuyant sur des données en temps réel telles que les signes vitaux du patient (pression artérielle, fréquence cardiaque et saturation en oxygène), les images radiographiques et les informations sur les médicaments administrés. L'analyse des actes par l'image après l'opération est également utile pour améliorer la prise en charge en cas de complications post-opératoires et pour améliorer la recherche et formation des futurs professionnels.

1.2 L'IA PEUT AMÉLIORER LA RECHERCHE CLINIQUE

Il existe de multiples exemples de l'usage de l'IA dans la recherche.

Le Professeur Besse, oncologue et directeur de la recherche clinique de l'Institut Gustave Roussy a présenté le projet développé par Lifen¹⁷¹ dans 800 établissements de santé, pour collecter et structurer la donnée clinique en vue de faciliter la recherche en oncologie.

Il souligne que le recours à l'IA apporte un gain de temps : sur une cohorte de

100 patients, il faut en moyenne 1 heures 30 par patient pour un médecin pour renseigner 200 champs obligatoires, alors que Lifen le fait en 1 minute et 30 secondes. Dans 7 % des cas seulement, il y a discordance entre l'homme et la machine, 8 fois sur 10 c'est l'humain qui s'est trompé.

L'IA permet aussi de créer des « jumeaux numériques » de patients suivis en essai clinique, pouvant être utilisés dans des essais virtuels lors de développement de nouveaux médicaments, pour déterminer l'efficacité potentielle d'une molécule en fonction des réactions du jumeau numérique.

1.3 L'ASSISTANCE DU MÉDECIN PAR L'IA DOIT ÊTRE COMPLÉTÉE PAR L'EXPERTISE MÉDICALE

L'expertise médicale reste fondamentale pour fournir des données de qualité et interpréter les résultats. S'il y a un avantage à mettre des outils experts dans les mains de praticiens généralistes, il y aura toujours besoin d'excellents médecins pour constituer le jeu de données qui va permettre à l'IA de détecter une lésion, et pour interpréter pour le patient l'éventuel diagnostic identifié par l'IA. De même, l'intégration de l'IA n'est envisageable que si la détection d'une lésion par cette technologie est également confiée à des médecins. Le respect de cette condition est indispensable pour interpréter le diagnostic identifié par l'IA auprès du patient et pour repérer les erreurs commises par cette technologie.

Le Professeur Pierre-Antoine Gourraud a scénarisé dans une tribune au Monde¹⁷² l'utilisation de l'IA en santé dans vingt

171 Lifen Documents - Gestion sécurisée de vos comptes rendus médicaux.

172 Professeur P.A Gourraud, L'intelligence artificielle en santé vue depuis l'an 2044, Le Monde, septembre 2024.

ans. Il met en avant le fait que « *plus l'IA devient accessible, plus l'expertise du spécialiste en IA médicale est indispensable. En effet, l'IA permet d'identifier des données jusqu'alors inconnues et pour lesquelles il n'est pas possible, en l'état actuel de nos connaissances, de conclure sur la causalité ou non pour expliquer les signes cliniques observés chez le patient* ».

Une fois les résultats produits par l'IA, et pour faire face au risque du «biais d'automatisation» (excès de confiance de l'humain envers les résultats produits par l'IA), la capacité d'interprétation du professionnel doit primer sur la base de ses compétences et de son esprit critique.

Si une décision est prise avec l'appui de l'IA, le professionnel doit pouvoir en informer son patient, expliquer pourquoi il suit les recommandations et comment il est arrivé à cette conclusion. Cela rend d'autant plus important le fait de concevoir des systèmes dont le fonctionnement est transparent, explicité et traçable. Le médecin doit aussi garder son autonomie et pouvoir être libre de refuser de suivre les recommandations de l'IA s'il pense qu'elle s'est trompée.

Pour Joris Galland¹⁷³, spécialiste en médecine interne : « *Votre smartphone sera votre futur stéthoscope* ». Il constate que, depuis 2020, cette prédiction se vérifie avec l'émergence de la m-santé (*m-health* pour mobile health) qui regroupe les objets portables connectés utilisés par les professionnels de santé (saturomètres, échographes portatifs, applications, etc.). L'essor croissant des appareils de santé connectés (smartphones, montres, vêtements, etc.) a permis « l'auto-mesure » par le patient en temps réel. Les montres

connectées peuvent ainsi enregistrer la fréquence cardiaque, reconnaître les arythmies, faire un électrocardiogramme, évaluer les efforts physiques (calories brûlées, nombre de pas, distance, etc.) Néanmoins, la masse d'informations médicales collectées ne peut être lue et interprétée par une seule personne en plus de ses consultations. Une IA est donc nécessaire pour interpréter les données, et alerter le médecin en cas d'anomalie. Le recueil de données physiologiques en temps réel, couplé à l'analyse des données par l'IA, participe à l'émergence d'une médecine prédictive et préventive qui détecte les pré-symptômes d'une maladie, le développement de la prévention étant également source d'économies.

1.4 L'USAGE DE L'IA EST DÉJÀ DÉVELOPPÉ CHEZ LES SOIGNANTS QUI SOUHAITENT CEPENDANT ACCÉDER À LA FORMATION

L'usage de l'IA semble déjà bien développé chez les soignants : plus d'1 soignant sur 2 (53 %) intègre l'IA dans sa pratique quotidienne d'après une étude récente effectuée auprès de 1 700 soignants par Pulse Life¹⁷⁴, parmi lesquels 52 % de médecins, 20 % d'infirmiers en pratique avancée, 8 % de pharmaciens, 3 % d'aides-soignants, 17 % issus d'autres professions (kinésithérapeutes, ostéopathes, dentistes, etc.).

La confiance dans l'aide au diagnostic se construit progressivement : 58,7% des professionnels déclarent avoir confiance, mais ce chiffre est légèrement supérieur chez les médecins (63,5 %).

Leurs compétences étant toujours nécessaires, les professionnels de santé ne se sentent pas menacés par l'IA : d'après la même étude, 66,6 % des

¹⁷³ Esanum, Internet des objets et m-santé, 2020.

¹⁷⁴ Baromètre IA en santé : alliée ou menace ? PulseLife, 2024.

soignants estiment que l'IA ne représente pas une menace pour leur profession (70,4 % pour les médecins). En revanche, 66% des soignants (60 % des médecins) expriment des inquiétudes quant à une possible dégradation du lien de confiance avec leurs patients.

Face à ce constat, un besoin de formation est clairement exprimé : si seulement 6,6 % des soignants interrogés ont reçu une formation spécifique sur l'IA, 78,2 % déclarent aspirer à recevoir une formation spécifique à l'IA (84 % des médecins). Les besoins de talents et de compétences numériques à traduire en termes de recrutements ont également été soulignés par les personnes auditionnées.

2. Les risques identifiés et les freins au développement de l'IA dans le domaine de la santé

2.1 LES RISQUES LIÉS AUX MODALITÉS DE DÉPLOIEMENT DE L'IA DANS LE SYSTÈME DE SOINS

L'essor de l'IA en santé entraîne une crainte liée à la déshumanisation du soin. Or, la relation humaine avec le médecin ne peut pas être réduite à un acte purement technique et remplacée par un outil. Dans le secteur de la santé mentale, l'usage de l'IA générative à des fins thérapeutiques se diffuse, des outils gratuits tels que « Psychologist » ou « Therapist », disponibles 24 heures sur 24 pouvant donner l'impression de parler à un vrai thérapeute. Toutefois, quand une personne évoque le sujet du suicide avec le *chatbot*, ces outils refusent d'évoquer le sujet et renvoient vers un professionnel de santé. Olivier Duris¹⁷⁵, psychologue clinicien et spécialiste des usages

numériques « *ne pense pas que ces outils suffiront à nous remplacer* ». Pour lui, qui utilise l'IA comme outil de médiation avec les enfants autistes, « *on peut faire des choses intelligentes avec l'IA. Mais ça n'a d'intérêt qu'à partir du moment où cela est pensé dans une relation avec un thérapeute, et jamais dans le but de le remplacer* ».

Le Professeur Besse a ainsi rappelé en conclusion de son audition que « *la médecine est d'abord une consultation* ». La relation soignant-soigné est fondamentale et ne peut être remplacée. La confiance établie en consultation est une composante essentielle du soin, au même titre que la thérapeutique proposée. L'IA peut en revanche assister le médecin qui restera toujours responsable de ses décisions.

Le risque accru de cyber-attaques est à tempérer par les possibilités d'améliorer la cybersécurité grâce à l'IA. Selon le professeur Besse, le risque de cyberattaques constitue le « *cauchemar* » des équipes soignantes, et le développement des usages de l'IA peut accroître les risques d'attaques, dans un contexte où sur les 581 incidents informatiques dans les hôpitaux et établissements médico-sociaux recensés en 2023 par l'Agence du Numérique en Santé, près de la moitié étaient des cyberattaques¹⁷⁶. Cependant, l'IA a aussi le potentiel de mieux détecter ces risques et de les réduire : détecter et bloquer les attaques en temps de réel, anticiper les comportements suspects, renforcer la sécurité des données, sécuriser le réseau ou prédire des failles.

Les inégalités territoriales d'accès aux soins pourraient s'accroître en

¹⁷⁵ Le Monde, Comment l'IA bouscule le milieu de la santé mentale : « Plutôt que de payer une nouvelle séance chez le psy, j'allais sur ChatGPT », 2024.

¹⁷⁶ Observatoire des incidents de sécurité des systèmes d'information pour les secteurs santé et médico-social 2023 | Agence du Numérique en Santé.

concentrant l'accès à l'IA dans les territoires les mieux dotés.

Comme l'a relevé un rapport de la Cour des comptes en mai 2024¹⁷⁷, les inégalités d'accès aux soins de premiers recours s'aggravent en France. Même si la télémédecine peut apporter une partie de réponse à ces inégalités, elle ne peut pas répondre aux inégalités de répartition des professionnels de santé sur le territoire. Par exemple, dans certains territoires, près d'un quart de patients n'ont pas de médecin traitant. Parallèlement l'investissement dans les équipements pour accéder à l'IA demeure coûteux pour les professionnels de santé et un risque de concentration du développement de l'IA sur un nombre limité de professionnels qui ont les moyens d'investir existe. Cette situation est source d'aggravation complémentaire des inégalités d'accès aux soins dans les territoires. Une « médecine à deux vitesses » risque donc de se développer entre territoires où les patients pourront accéder aux soins appuyés par l'IA et d'autres territoires dépourvus des progrès techniques. Or comme le notait Albane Grandjean : « *est-ce vraiment éthique de ne pas utiliser l'IA, si elle est plus efficace ?* ».

Dans la continuité du précédent paragraphe relatif au caractère inégalitaire de l'accès aux soins, il convient de préciser que les femmes pourraient être particulièrement défavorisées dans cette transformation du système de

santé qui advient avec l'intégration de l'IA. À titre d'exemple, depuis 2014 en Espagne, certains hôpitaux ont introduit l'intelligence artificielle dans leurs méthodes pour prédire la compatibilité des donneurs/ donneuses et des receveurs/ receveuses afin d'aider les médecins dans le cadre de greffes de foie. Si cette méthode semble très novatrice et pertinente, elle peut néanmoins comporter des biais en défaveur des femmes. En effet, deux ans plus tard, les médecins se sont rendu compte qu'aucun organe n'avait été attribué à une femme. L'apprentissage de l'algorithme avait été réalisé sur la base d'études qui n'avaient pris en compte que peu de femmes et n'avaient pas permis de constater la différence des biomarqueurs impliqués dans les greffes impliquant des hommes et des femmes. Plus largement, alors que les femmes vivent plus longtemps et donc plus souvent victimes de maladies chroniques, elles sont cependant moins bien soignées en raison des biais dans les diagnostics en fonction du genre, avec la perception de certaines maladies considérées comme « masculines », à l'image des maladies cardiovasculaires. Ainsi, l'introduction des algorithmes en santé, si ceux-ci apprennent sur des données déjà produites - qui tendent à invisibiliser les particularités de la santé des femmes -, on ne peut que supposer une amplification des discriminations à l'encontre des femmes dans ce domaine¹⁷⁸.

La médecine prédictive pourrait

¹⁷⁷ Accès aux soins : les inégalités territoriales s'aggravent | vie-publique.fr, mai 2024.

¹⁷⁸ Joséphine Lethbridge, « Oui, l'IA est sexiste - qu'est-ce qu'on peut faire ? », Les Glorieuses, 29 avril 2024 et M. Salle et C. Vidal, Femmes et santé, encore une affaire d'hommes ?, Égale à Égal (Belin, 2019), pp. 11-12.

entraîner un risque de surdiagnostic. Si la médecine prédictive peut prendre son essor avec l'IA, avons-nous réellement envie de tout prévoir ? de tout savoir ? Le surdiagnostic peut amener des personnes à être considérées comme malades. Comme l'ont montré plusieurs études, le bénéfice du dépistage du cancer de la prostate n'est pas toujours démontré au regard des inconvénients ou effets secondaires liés aux traitements. Il existe un risque de surdiagnostic et de surtraitement : le surdiagnostic varie de 30 % à 50 % des cancers de la prostate dépistés selon les études. En réalité, comme le montre l'Institut national du cancer, ces cancers auraient pu ne pas être traités sans que cela porte préjudice aux hommes. Les personnes concernées n'auront donc pas tiré bénéfice du dépistage quand il est organisé sous forme de dépistage collectif et non dans une logique de personnalisation de la médecine. En revanche, elles sont exposées à un surtraitement et aux effets indésirables et inconvénients potentiels liés à la prise en charge diagnostique et thérapeutique¹⁷⁹.

Les risques relatifs à la protection des données de santé sont particulièrement encadrés au niveau national et européen.

Particulièrement sensibles, les données de santé nécessitent une protection renforcée. Pour protéger la vie privée des individus tout en permettant leur utilisation pour améliorer les soins et la recherche, l'utilisation de ces données fait l'objet d'une réglementation stricte au plan national et européen : le RGPD de 2018 et le Règlement européen relatif à l'espace européen des données de santé (EHDS) de 2024. L'objectif de ce dernier texte est

de fluidifier le partage et la réutilisation de données à des fins de recherche et de politique publique avec un espace commun au sein duquel les personnes physiques pourront aisément contrôler leurs données de santé électroniques. Mon espace santé a été conçu et développé en ce sens. Il permettra aussi aux chercheurs, aux innovateurs et aux décideurs d'utiliser ces données de santé électroniques d'une manière fiable et sûre, qui préserve la confidentialité.

En France, les données sont pseudonymisées (à savoir anonymisées avec utilisation de pseudos pour les identifiants comme les noms, prénoms, emails, adresses, numéros de téléphone)¹⁸⁰ avant d'être transmises aux chercheurs et uniquement sur des projets autorisés par des comités éthiques. Elles peuvent être traitées sous réserve d'accomplir des formalités préalables auprès de la CNIL, sauf dans certains cas où le consentement explicite et préalable de la personne a été recueillie¹⁸¹. La complexité des formalités nécessite un accompagnement renforcé comme cela a pu être fait pour l'entreprise Lifem, spécialisée dans l'accès simplifié aux données de santé et ce durant plusieurs mois par la CNIL sur le plan juridique et technique pour mettre en œuvre les obligations liées à la protection des données¹⁸².

2.2 LES FREINS EXISTANTS AU DÉPLOIEMENT DE L'IA DANS LE SYSTÈME DE SOINS

Plusieurs personnes auditionnées ont insisté sur le manque de financements publics dédiés à l'IA, dans un contexte marqué par la crise de l'hôpital et la

179 CP prostate.

180 Botdesign, Livre blanc, Données de santé artificielles, avril 2024.

181 CNIL - Quelles formalités pour les traitements de données de santé, 2024.

182 CNIL, Accompagnement renforcé : la CNIL sélectionne 3 entreprises du numérique à fort potentiel, 2023.

pénurie de soignants. Par exemple, l'équipement en IA en soutien à l'imagerie médicale repose à ce stade sur les cabinets de radiologie, sans réelle compensation financière par l'Assurance Maladie.

La question du coût et du mode de financement est à aborder de manière globale avec les économies que l'investissement dans l'équipement en matière d'IA peut générer par la suite.

Le manque de numérisation a également été pointé par certains auditionnés : il existe encore des mammographies papier qui présentent le risque d'une perte de données. Cela s'accompagne d'un manque de littéracie numérique chez certains professionnels et d'un réel désir de formation.

Dans ce contexte, pour généraliser le partage fluide et sécurisé des données de santé entre professionnels et usagers afin de mieux prévenir, mieux soigner et mieux accompagner, une stratégie nationale a été lancée en 2018, le « Ségur du numérique en santé »¹⁸³. Son objectif¹⁸⁴ est de généraliser, au travers d'un dispositif innovant de financements à l'équipement et de financements à l'usage, le partage fluide et sécurisé des données de santé entre professionnels de santé et usagers ; et d'accélérer de façon significative la feuille de route du numérique en santé, en passant de 10 millions à 250 millions de documents médicaux échangés d'ici 2023 pour mieux prévenir et mieux soigner.

La question de l'interopérabilité des logiciels a été aussi mentionnée comme un frein : il n'y a pas de normes d'interopérabilité pour les logiciels médicaux et la faible portabilité génère la peur de perdre des données et de bloquer les professionnels dans l'utilisation d'un logiciel.

Au niveau de la réglementation, pour certains professionnels auditionnés par la commission IA, la réglementation relative à la protection des données « *reste un frein* ». Elle appellerait à être simplifiée lorsqu'il s'agit de dispositifs médicaux. L'ajout d'audits supplémentaires par rapport à l'existant en termes de réglementation des dispositifs médicaux est craint par certains professionnels car les audits nécessitent d'y consacrer du temps et des moyens financiers. Les exigences en termes de cybersécurité obéissent à des standards mondiaux mais de plus en plus de certificats locaux « hébergeurs de données de santé » apparaissent, multipliant les couches de réglementation.

Un avis du CESE sur le numérique en santé, à ce jour en cours d'élaboration par la Commission Affaires sociales et santé, viendra compléter cette partie et rappeler que le développement de l'IA en santé repose sur la collecte et l'analyse de données de santé de qualité et en nombre suffisant. Il cherchera à répondre aux questions stratégiques, économiques, sanitaires, sociales mais aussi

¹⁸³ Le Ségur du numérique en santé.

¹⁸⁴ Les chiffres clés du déploiement du Ségur du numérique en santé | Agence du Numérique en Santé.

éthiques que pose le numérique en santé

et donc les innovations produites par l'IA.”

Les préconisations du CESE

- **Formation des personnels soignant et administratif**

L'usage de l'IA est déjà bien intégré par les soignants, avec plus de 53% l'utilisant quotidiennement. La confiance dans l'IA pour le diagnostic est en croissance et, s'ils craignent une dégradation du lien de confiance avec leurs patients, les soignants ne se sentent pas véritablement menacés par l'IA.

Pour le personnel administratif, une formation IA peut également permettre un gain de temps par la réduction de tâches logistiques et ainsi améliorer la qualité de vie au travail.

PRÉCONISATION #28

Le CESE appelle à développer une formation à l'IA adaptée aux usages actuels et potentiels à destination des personnels soignants et administratifs.

- **Recherche médicale**

PRÉCONISATION #29

Le Cese recommande de veiller à ce que les programmes de recherche soient bien exempts de biais sexistes.

- **Équipement sur le territoire**

Alors que l'IA peut apporter des progrès en matière de diagnostic, de traitement et d'opération chirurgicale, le manque de financements dédiés a été relevé par plusieurs des personnes auditionnées dans le champ de la santé et dans un contexte financier dégradé pour la santé. La question du coût et du mode de financement est cependant à aborder de manière globale avec les économies que l'investissement dans l'équipement en matière d'IA peut générer par la suite, notamment grâce à la prévention.

PRÉCONISATION #30

Le CESE recommande d'assurer un financement dédié à l'IA et de le mutualiser entre l'ensemble des parties prenantes (publiques et privées) pour assurer une couverture équitable sur l'ensemble des territoires.

Face au manque de numérisation dans le domaine de la santé qui pose des risques de perte de données préjudiciable aux patients, une stratégie « Ségur du numérique en santé » a été lancée en 2018. Elle a pour objectif d'améliorer le partage sécurisé des données de santé entre patients et professionnels grâce à des financements innovants.

PRÉCONISATION #31

Le CESE préconise de poursuivre, dans un cadre éthique, le développement du numérique en santé engagé par le « Ségur du numérique » dans l'objectif d'une généralisation d'un partage fluide et sécurisé des données de santé entre les professionnels et les usagers par le déploiement d'infrastructures partagées à l'échelon national notamment publiques.

Recommandation de citoyennes et citoyens

L'IA dans la santé permet de poser des diagnostics plus rapides, plus précis et de personnaliser les traitements, ce qui améliore la qualité des soins. Cependant, elle présente des risques, comme le manque de formation des soignants et la déshumanisation des soins. Le CESE recommande de former les professionnels de santé à l'IA et de garantir un accès géographique et social équitable à ces technologies. Il insiste aussi sur la protection des données de santé et sur la nécessité de maintenir une relation humaine dans les soins.

G. L'IA est-elle génératrice d'accessibilité ? L'exemple du handicap

Selon Jérémie Boroy¹⁸⁵, il est essentiel de comprendre que le handicap résulte de la combinaison de deux facteurs : d'une part, une déficience personnelle (visuelle, auditive, moteur, intellectuelle, etc.), et d'autre part, un environnement qui n'est pas adapté et met la personne en situation de handicap. Cette dernière est donc directement liée à l'inadaptation de l'environnement face à la prise en compte de la déficience. L'enjeu est donc d'assurer, en toutes circonstances, que notre environnement soit adapté à toutes et tous, pour que les personnes qui ont une déficience ne soient pas en situation de handicap. En cas d'inaccessibilité de l'environnement, il faut mobiliser des outils de compensation. Si l'accessibilité est une démarche qui relève du collectif (État, collectivités, organisations, la société au sens large), la compensation repose sur des solutions individuelles. La priorité doit être de rendre l'environnement accessible, et à défaut, de permettre la mise en place de solutions de compensation.

L'intelligence artificielle doit répondre à ces mêmes injonctions : par nature, être accessible à toutes et tous, en particulier celles et ceux qui ont une déficience. À défaut, elle doit offrir des solutions de compensation.

1. L'IA permet de progresser en faveur de l'accessibilité

Pour chaque type de handicap, il existe des exemples pour illustrer l'accélération de l'accessibilité grâce à l'intelligence artificielle¹⁸⁶ :

→ la déficience auditive peut être

compensée par des solutions de reconnaissance vocale et de transcription à l'écrit de conversations. Il existe aussi des solutions de traduction automatique de contenus audios avec des avatars en langue des signes - cette solution étant, à ce jour, développée surtout pour des contenus écrits, et moins oraux.

- les troubles de l'élocution peuvent être compensés par des solutions qui visent à redonner la voix à des personnes ayant perdu la parole, au travers d'une transcription orale de contenus écrits, pouvant utiliser la voix préenregistrée de la personne qui aurait perdu la voix ;
- la déficience visuelle peut être compensée par une description orale d'images ou de contenus, ou par l'usage d'une écriture adaptative mettant en évidence de contenus importants dans différents formats ;
- les limitations de mobilité peuvent être compensées par l'identification d'obstacles sur un chemin, ou par l'usage d'exosquelettes ;
- les troubles « dys » (dyspraxie, dyslexie, etc.) peuvent être compensés par l'usage d'outils de prise de note automatique, de rédaction de comptes rendu ou synthèses, permettant à l'individu de se délester de tâches complexes, et de réaliser une tâche à la fois ;
- des technologies d'IA ont été mises au point pour faciliter les interactions avec des enfants souffrant de certains

¹⁸⁵ Audition de Jérémie Boroy, président du Conseil national consultatif des personnes handicapées (CNCPH) et conseiller du CESE, par la Commission temporaire IA, le 31 octobre 2024.

¹⁸⁶ Handicap et intelligence artificielle - CFHE, mai 2024.

troubles du spectre de l'autisme (TSA). En réduisant l'appréhension et la peur du jugement que ressent souvent l'enfant autiste à l'idée de parler avec un adulte, des applications d'IA peuvent en effet aider les enfants autistes à communiquer plus aisément. Le robot « Kaspar » a ainsi été développé par l'Université du Hertfordshire au Royaume-Uni, dans l'objectif d'interagir avec des enfants autistes pour les accompagner dans le développement de compétences sociales fondamentales, comme l'imitation ou l'acquisition de gestes du quotidien à travers le jeu. Il existe aussi le robot humanoïde « QRobot »¹⁸⁷, développé intégralement au Luxembourg par l'entreprise Luxai pour faciliter les interactions entre les thérapeutes et les enfants. Le robot doit, par une approche ludique, permettre que les enfants acquièrent des émotions, aptitudes sociales et capacités de communication¹⁸⁸.

L'IA permet ainsi de résoudre des problèmes d'accessibilité non résolus jusqu'alors, dans des environnements multiples, tels que l'école, l'enseignement supérieur, le travail, le monde médical ou plus largement, la société.

2. L'IA n'est véritablement accessible que si elle intègre la question de la diversité des publics et des usages dès la conception...

Si « l'IA est un boulevard vers l'accessibilité et l'autonomie », selon Jérémie Boroy, « c'est à condition de se poser en amont la question des publics et des usages ». En effet, si on laisse se développer des solutions d'intelligence artificielle sans intégrer l'enjeu d'accessibilité dès la conception de l'IA, on crée des fossés nouveaux qui n'existaient pas auparavant.

Thomas Fauvel¹⁸⁹, met en perspective la prise en compte de l'accessibilité dans l'IA avec l'accessibilité du numérique : « la notion d'accessibilité est souvent arrivée en retard dans l'informatique. L'enjeu est ici de ne pas reproduire les erreurs du passé, et de faire en sorte que l'IA soit transformante aussi pour les personnes en situation de handicap. L'IA doit être rendue accessible de matière native ».

En pratique, il est donc essentiel de s'assurer que tous les programmes de développement de solutions d'IA intègrent les critères et la vision établie par la Convention internationale des droits des personnes handicapées (CIDPH).

En outre, l'existence de solutions de compensation grâce à l'IA ne doit pas conduire à décharger de responsabilité la société, l'État, les collectivités ou les organisations dans leur obligation d'assurer l'accessibilité de l'environnement. L'IA peut compenser, peut apporter des solutions à des situations encore jamais résolues, mais ces solutions

¹⁸⁷ QTRobot, le compagnon des enfants autistes - Luxembourg.

¹⁸⁸ Rapport IA et Education, Délégation à la prospective, Sénat, 2024.

¹⁸⁹ Audition de Thomas Fauvel, ingénieur et pilote de la Délégation jeunesse du Conseil national consultatif des personnes handicapées (CNCPH), par la Commission temporaire IA, le 31 octobre 2024.

propulsées à l'IA ne correspondent pas toujours aux besoins de toutes et tous, ou ne sont pas souhaitées par celles et ceux qui vivent avec leur déficience depuis longtemps. L'IA n'est pas la solution unique.

Enfin, ces solutions techniques ne se substituent pas aux compétences et au soutien apporté par les êtres humains. En effet, pour la Professeure Enkelejda Kasneci¹⁹⁰ « *l'utilisation de grands modèles de langage doit être accompagnée de l'aide de professionnels tels que des orthophonistes, des éducateurs et d'autres spécialistes qui peuvent adapter la technologie aux besoins spécifiques des handicaps de l'apprenant* ».

3. ... et tient compte des situations hétérogènes des personnes en situation de handicap

L'adoption de l'IA pour les personnes en situation de handicap présente des défis, notamment financiers et sociaux. Les technologies, comme les prothèses intelligentes, sont coûteuses et nécessitent des mises à jour régulières. Cela peut exclure les personnes à faibles revenus, aggravant les inégalités. Le recours à l'IA pourrait aussi réduire les services humains, ce qui nuirait à certains. Les écarts sociaux risquent de se creuser si l'accès à ces technologies est limité. Les personnes handicapées, souvent confrontées au chômage et à la précarité, peinent à investir dans ces outils.

Il est crucial de rendre l'IA accessible financièrement et de simplifier les démarches. Les progrès technologiques dépassent les capacités des systèmes d'aide actuels. Une politique publique proactive et des partenariats sont nécessaires pour réduire ces inégalités, à inscrire dans une possible évolution de la loi handicap de 2005.

¹⁹⁰ Enkelejda Kasneci et al., 2023 dans *Intelligence artificielle et éducation : apports de la recherche et enjeux pour les politiques publiques : édition trilingue 2024*.

Les préconisations du CESE

PRÉCONISATION #32

Le CESE recommande d'intégrer systématiquement l'enjeu d'accessibilité dans les IA en :

- soutenant le développement de solutions de compensation et d'accessibilité fondées sur l'IA pour une meilleure inclusion dans la vie quotidienne des personnes en situation de handicap ;
- s'assurant que tous les programmes de développement de l'IA soient bien conditionnés au critère systématique de l'accessibilité des outils avec un critère transversal tel que « l'aménagement raisonnable » instauré dans le système juridique français par la directive européenne 2000/78 (qui le définit comme « l'ensemble des mesures appropriées visant à éliminer les diverses barrières qui entravent la pleine et effective participation des personnes handicapées à la vie professionnelle sur la base de l'égalité avec les autres travailleurs ») ;
- développant une politique publique volontariste et des partenariats innovants pour rendre ces technologies réellement accessibles et ainsi réduire les inégalités sociales qui en découlent. Une approche inclusive, éthique et adaptée permettra de garantir que ces technologies servent à réduire les inégalités, plutôt qu'à les renforcer. Cela pourrait nécessiter une évolution de la loi du 11 février 2005 pour « l'égalité des droits et des chances, la participation et la citoyenneté des personnes handicapées ».

📌 Recommandation de citoyennes et citoyens

L'IA peut vraiment améliorer la vie des personnes handicapées, en offrant des outils comme la reconnaissance vocale ou des interfaces adaptées. Mais beaucoup de technologies ne sont pas encore conçues pour être accessibles à toutes et tous, ce qui limite leur utilisation. Le CESE recommande de renforcer leur accessibilité pour qu'elles profitent à tout le monde.

Conclusion

Le déploiement de l'intelligence artificielle au service des êtres humains et de l'intérêt général nécessite une vigilance constante et un engagement collectif.

Comme l'affirme Daron Acemoğlu, prix Nobel d'économie : « *Nous avons toujours la possibilité de faire des choix qui sont essentiels... Le problème actuel est que nous renonçons à ce choix et laissons le débat sur l'ère future de l'IA à quelques individus. La société, les travailleurs, les syndicats doivent participer au débat sur l'IA.* »

Fidèle à son attachement à la démocratie, le CESE réaffirme que ce débat ne peut être laissé à une poignée d'acteurs, qu'ils soient économiques ou technologiques. Il en va de l'essence même des valeurs de la France et de l'Europe. Il s'agit d'assurer les conditions d'une réflexion inclusive et éclairée où chaque citoyen et citoyenne, et chaque organisation a son rôle à jouer.

C'est pourquoi le CESE appelle à une vigilance collective pour accompagner le déploiement de l'IA de manière responsable, éthique et durable. Au-delà des institutions, il invite chacun à se forger sa propre opinion, à s'informer, et à contribuer activement à ce débat crucial pour l'avenir de notre société.

Agir autrement pour l'innovation sociale et environnementale

Avec ou sans nous, l'IA se diffusera à l'échelle planétaire.

C'est une réalité que nous devons accepter. L'IA ouvre une révolution qui s'impose à nous à une vitesse exponentielle.

Et comme toutes les révolutions, elle s'accompagne de nombreux risques et de nombreuses limites très bien exposés dans cet avis : des risques pour la transition écologique mondiale pourtant vitale, pour l'emploi, pour nos droits fondamentaux et pour notre indépendance stratégique, celle de la France et celle de l'Europe.

Mais il ne tient qu'à nous de transformer ces risques en opportunités, ces limites en innovations.

Nous sommes intimement persuadés que le défi qui est le nôtre, c'est celui de faire de cette révolution une réelle opportunité pour réinventer notre modèle social qui peine à ce jour, et plus encore dans les années à venir, à se financer et à répondre aux défis croissants qui nous font face.

Cette innovation, créatrice de fortes valeurs ajoutées, doit nous permettre de bâtir une économie responsable, innovante et inclusive restant dans les limites de la planète.

Notre contrat social de demain devra répondre aux défis posés par l'IA, tout en exploitant ses opportunités. Et pour ce faire, nous devons nous saisir pleinement de ce sujet et de cette émergence, par une gouvernance proactive, par la mise en place de régulations adaptées, et par le partage de la

valeur créée par l'IA au profit de l'ensemble de nos concitoyens et de notre environnement. Pour être clair : confier l'IA aux seules mains des forces du «marché» relèverait d'une occasion ratée, pire de l'inconscience.

Une intelligence artificielle au service de notre projet social et environnemental, voilà ce que nous prônons, voilà ce que nous votons aujourd'hui.

Enfin, et j'en terminerais là, nous tenons à souligner le travail exceptionnel des deux rapporteurs, Marianne et Eric, de Thierry Cadart, Président de cette commission temporaire, ainsi que l'ensemble des conseillers qui a œuvré collectivement à cet avis remarquable et surtout aux 10 citoyens qui, dans un espace-temps très resserré, ont démontré toute la pertinence d'unir la société civile pour construire ensemble le monde de demain....

Agriculture

L'intelligence artificielle s'impose un peu plus chaque jour dans notre quotidien, avec ses opportunités mais aussi avec ses risques. Les représentants des organisations de la société civile organisée sont tout à fait dans leur rôle pour éclairer le législateur, les acteurs publics et les citoyens sur les enjeux de cette technologie désormais largement diffusée.

Avec cet avis nous avons voulu mettre en avant la nécessité de réinterroger nos modes de vie face à l'émergence de l'IA.

La technologie peut bouleverser nos façons d'apprendre, de produire, de travailler ou encore d'échanger. L'avis

appelle à la prudence face à d'éventuelles menaces sur nos libertés individuelles, sur nos conditions de travail, sur la préservation de l'environnement ; nous partageons la volonté des rapporteurs de proposer les indispensables régulations, condition pour bénéficier des progrès que peut apporter l'IA à chaque individu comme à la collectivité.

Le groupe de l'agriculture approuve l'état d'esprit de cet avis et estime nécessaire, comme le rappelle la conclusion, que la société civile continue à participer à une vigilance collective pour que soient préservées les valeurs qui fondent notre démocratie.

Les grands enjeux de l'IA pour le secteur agricole ont été exposés lors d'une audition thématique. La question du partage et de l'accessibilité des données fournies par le travail de l'agriculteur doit être prise en compte. Nous approuvons en ce sens, la préconisation 1 sur la transparence des algorithmes.

L'intelligence artificielle offre de réelles perspectives pour conduire les transitions écologiques et climatiques en apportant des garanties en termes de revenus, de baisses de charges, d'allègement de la pénibilité pour les agriculteurs et aussi, pour les consommateurs, des outils de traçabilité, de transparence, de connaissance des produits pour éclairer leurs choix. Il faut toutefois que les investissements technologiques soient financièrement accessibles à tous, qu'ils représentent véritablement un progrès et qu'ils soient fiables. À cet égard, la dynamique des investissements publics est essentielle. Plus nous gagnerons en compétence, plus nous maîtriserons la technologie et plus nous serons forts face à nos concurrents. Nous approuvons tout particulièrement les préconisations visant à renforcer l'autonomie stratégique européenne de l'IA.

Enfin nous sommes particulièrement

sensibles au diagnostic et aux préconisations relatives à l'éducation. L'IA doit rester un outil au service de l'homme : ce ne sera possible que si nous cultivons la valeur ajoutée de notre intelligence humaine grâce à notre esprit critique et notre créativité. Une des clés de cette réussite réside dans l'information et la formation, initiale et continue, de l'ensemble des utilisateurs et des acteurs tant sur le fonctionnement de l'IA que sur ses perspectives positives et ses risques. Le groupe a voté pour.

Alternatives sociales et écologiques

Cet avis repose sur un constat évident : l'IA est déjà présente, et transforme nos vies, nos économies, nos interactions. Notre responsabilité est donc qu'un cadre strict soit posé afin de modeler cette transformation pour qu'elle soit compatible avec les limites planétaires et socialement juste, qu'elle s'opère dans le respect des libertés individuelles, des droits fondamentaux et qu'elle préserve la démocratie.

Nous ne pouvons pas nous permettre de laisser la question du déploiement de l'IA entre les seules mains de quelques grands acteurs économiques. Il s'agit pour nous d'un sujet politique. Si nous ne nous saisissons pas collectivement de cet enjeu, l'IA deviendra un outil supplémentaire pour accroître les profits d'une élite économique, au détriment du reste de la population. Nous refusons une IA qui aliène davantage les travailleuses et les travailleurs, accentue les inégalités et aggrave la crise écologique. Comme il est rappelé dès l'introduction de cet avis : « le développement technique [et donc l'IA] n'est pertinent que s'il est au service des êtres humains et de l'intérêt général ».

Par ailleurs, l'avis insiste sur la nécessité de réflexion et de formation pour ne pas

subir l'IA. Un récent rapport de l'UNESCO, cité dans l'avis, précise : « il est indispensable de donner aux individus les connaissances, les aptitudes, les attitudes et les valeurs dont ils ont besoin non seulement pour se protéger des risques, mais aussi pour tirer pleinement parti des opportunités offertes par l'IA ». Plusieurs préconisations relatives à cet impératif de formation, pour l'ensemble des individus qui le souhaitent et à tous moments de la vie, sont présentes dans le texte. Nous les saluons, en particulier la préconisation 17 appelant à intégrer la formation à l'IA dans les entreprises. L'IA ne doit pas être seulement une affaire de spécialistes ou de décideurs, c'est une transformation globale qui exige l'implication éclairée de l'ensemble de la société.

Toutefois, le temps dont la commission temporaire disposait pour construire cet avis a été trop court. L'ampleur du sujet est immense, et certains items n'ont pas pu être explorés avec toute la profondeur qu'ils méritaient. Néanmoins, la discussion est restée ouverte, les divers points de vue ont été intégrés, et le texte a ainsi largement évolué au fil des débats.

L'avis présenté est le fruit d'un intense travail de négociation et de compromis, fidèle à l'esprit du CESE. Nous pensons qu'il est équilibré et qu'il reflète au mieux les préoccupations et aspirations de chacun.

Nous avons donc voté ce texte, à la fois didactique et porteur de préconisations essentielles pour poser les limites et les conditions du déploiement de l'IA. Nous considérons que cet avis envoie un signal fort : celui d'une société qui

maîtrise son avenir technologique au service de l'humain.

Nous avons adressé nos remerciements sincères à l'ensemble des membres de la commission temporaire, en particulier les citoyennes et citoyens, les conseillères et conseillers, les administratrices et administrateurs, Monsieur le Président de la commission ainsi que les co-rapporteurs pour leur implication et leur engagement dans ce travail collectif.

Artisanat et professions libérales

Certains voient dans l'IA la solution miracle quand d'autres n'y voient qu'un danger.

Il est vrai que l'IA ouvre la voie à des progrès considérables en matière de recherche, de santé, mais aussi dans les domaines du bâtiment et des infrastructures à travers la maintenance prédictive, et plus largement dans tous les secteurs en réduisant les tâches ingrates et en facilitant les tâches complexes. L'IA n'en présente pas moins de nombreux défauts tels que des biais de natures diverses, de forts impacts environnementaux ou encore des risques de manipulation de l'information et d'atteinte aux droits fondamentaux.

Alors que sa montée en puissance est inéluctable, notre Groupe estime qu'il est essentiel d'aborder cette technologie avec lucidité.

C'est bien, selon nous, l'approche retenue par l'Avis qui relève à la fois les opportunités offertes dans les domaines qui touchent à la vie des citoyens, tout en regardant avec réalisme les risques inhérents à

l'usage de l'IA et en identifiant les leviers pour s'en prémunir.

Notre Groupe soutient cette approche pragmatique et pédagogique, d'autant qu'elle s'inscrit dans un objectif clairement affiché : celui d'appréhender l'IA comme un outil au service de l'humain et de l'intérêt général.

Des artisans et des professionnels libéraux de santé, avocats, architectes, experts-comptables et autres sont sur cette même ligne ; bien conscients que l'IA dont ils sont déjà nombreux à faire usage ne saurait se substituer ni à leur expertise ni à leur responsabilité à l'égard du client ou du patient.

L'IA va impacter fortement la manière dont le travail sera réalisé dans la plupart des emplois. Aussi, il importe de préparer les travailleurs – salariés comme indépendants – à ces évolutions. A ce titre il est crucial d'en faire rapidement un sujet de la formation initiale et continue dans tous les secteurs, mais aussi d'associer les collaborateurs à ces changements à travers le dialogue social. Nous approuvons les propositions de l'Avis en ce sens, tout comme celles appelant à diffuser une culture de l'IA auprès des élèves et des enseignants, et plus largement de l'ensemble des citoyens.

Alors que l'IA est clairement un levier stratégique pour les Etats, il est par ailleurs essentiel que l'Europe - et la France en particulier - se mobilisent de façon déterminée autour de cette priorité, par le soutien à l'innovation et aux chercheurs, mais aussi aux acteurs européens de l'IA auxquels l'Avis propose, à juste titre, d'accorder une préférence dans le cadre de la commande publique.

Enfin, il est fondamental de sensibiliser et d'accompagner les TPE afin qu'elles s'approprient l'IA pour optimiser leurs pratiques et gagner en efficacité, quel que soit leur secteur d'activité. C'est bien l'objectif de nos organisations

représentatives des professions libérales et de l'artisanat, qui encouragent aussi le développement d'applications-métiers basées sur l'IA, telles que celles destinées par exemple aux boulangers-pâtisseries ou aux artisans du bâtiment.

Considérant que l'Avis porte une analyse à la fois complète et équilibrée sur l'IA, le Groupe Artisanat et Professions Libérales a exprimé un vote favorable.

Associations

« Comment mettre le déploiement de l'IA au service de l'intérêt général ? » tel est l'enjeu de cette saisine, qui ne vise pas à épuiser le sujet dans tous ses développements sectoriels. Car il est en effet bien trop tard pour se poser la question de son déploiement !

Si nous sommes seuls à stopper, d'autres bien moins soucieux d'intérêt général et de démocratie ou encore d'égalité, continueront seuls et ce n'est pas envisageable.

C'est pourquoi, le groupe des associations pense qu'il faut en encadrer les usages de l'intelligence artificielle dès maintenant.

Les possibilités offertes par l'IA peuvent apporter des conditions de vie et de travail positives et réduire les inégalités pourvu que ses usages soient aussi pensés pour cela.

Cependant, ce qui nous inquiète par-dessus tout aujourd'hui sont les outils qui permettent la manipulation de l'information.

Or nos démocraties sont basées constitutionnellement sur l'existence de citoyennes et citoyens éclairés.

Comment dès lors se forger une opinion lorsque on ne peut plus avoir confiance en ce que l'on voit, ni s'accorder sur les faits ?

Comment reconnaître une image ou une vidéo fabriquée ?

Si le grand n'importe quoi des réseaux

sociaux aujourd'hui attente déjà à notre capacité commune à vivre ensemble, comment même demain la preuve juridique pourra-t-elle être administrée ?

Car n'en doutons pas, s'il est encore à peu près facile de discriminer une vraie photo ou vidéo d'un contenu fabriqué via l'IA, cela ne durera pas.

L'usage, par exemple, de Watermarks automatiques devrait être dès à présent obligatoire.

De même que la mise en application du règlement européen sur les services numériques et la protection stricte de nos données avec le RGPD ou du droit d'auteur, doit se faire sans concession.

Aussi, nous attendons du gouvernement et de l'Europe qu'ils travaillent à conscientiser les citoyennes et citoyens autour de ces technologies, à les former, les éduquer et à réglementer ces technologies pour nous pour protéger et protéger l'intérêt général.

Il est temps aussi de renforcer les associations d'éducation populaire pour les conforter dans leur mission d'éducation tout au long de la vie. Sur ces questions aussi, elles sont incontournables et déjà très engagées

Le groupe des Associations tient tout particulièrement à remercier les citoyennes et citoyens qui ont travaillé avec les conseillères et conseillers sur cette question. Il a majoritairement voté en faveur de cet avis.

CFDT

Si l'intelligence artificielle est présente depuis longtemps dans notre quotidien, souvent sans

que nous en ayons conscience, la démocratisation des systèmes d'intelligence artificielle générative a fait émerger le sujet dans le débat public. Ainsi l'IA a-t-elle suscité à la fois méfiance, voire rejet, chez les uns, adhésion curieuse, voire enthousiasme chez les autres.

Pour la CFDT, il est important d'objectiver le débat. Comme pour toute nouvelle technologie, le développement de l'IA offre, par exemple, des possibilités d'innover mais il comporte également des risques, que nous pouvons déjà appréhender sans toutefois pouvoir en mesurer l'étendue.

L'avis pointe bien les différents enjeux inhérents au développement de cette technologie qu'ils soient environnementaux, éthiques, de genre, sociétaux, économiques, voire géopolitiques. Il est ainsi équilibré et nuancé, nous l'avons voté.

L'IA a bien évidemment un fort impact sur le travail, comme le pointe l'analyse de controverses de la commission Travail et emploi. Pour la CFDT, si les discussions se concentrent aujourd'hui principalement sur les enjeux d'emploi, il est crucial d'engager un dialogue social et professionnel sur les transformations du travail induites par l'IA. Seule cette perspective plus large permettra d'orienter l'utilisation de l'IA vers l'amélioration des conditions de travail. Cela suppose dès à présent des investissements en formation pour l'adaptation, le renouvellement des compétences et pour des reconversions professionnelles.

Les technologies ne sont jamais neutres : le dialogue social et le débat démocratique doivent en déterminer les usages souhaitables

et ceux à limiter, voire à proscrire. La régulation est donc indispensable pour éviter toute dérive, tout en sécurisant les opportunités.

Enfin, il nous semble évident que le sujet de l'IA n'est pas clos pour le CESE. D'autres commissions seront probablement amenées à approfondir ce sujet infiniment complexe.

CFE-CGC

« Science sans conscience n'est que ruine de l'âme »

La formule de Rabelais s'applique à toutes les nouvelles technologies. L'Intelligence Artificielle est une histoire ancienne mais le phénomène ChatGPT et l'IA générative a mis en lumière cette science informatique.

Nouvelle innovation ou révolution qui bousculera la société et nos façons de vivre. Il est sûrement encore trop tôt pour le savoir. Mais il est bien que le CESE se soit saisi de cette question dans un avis très riche, traitant un vaste panel de questions sociétales et sociales. Que ce soit sur la santé, le handicap, le travail, l'éducation ou l'environnement, l'avis présente les bénéfices potentiels et les risques à maîtriser.

Pour la CFE-CGC, dans l'entreprise, il est essentiel que les salariés soient impliqués dans le déploiement de cette technologie qui peut avoir des impacts sur les emplois, le contenu du travail ou la gestion RH.

La CFE-CGC considère qu'il est très difficile de prévoir les gains de productivité ou les conséquences sur le volume des emplois. Cela dépend de la façon dont chaque entreprise veut utiliser ces nouveaux outils.

Comme lors de l'implémentation de toute nouvelle technologie, les représentants des salariés doivent être informés et des études d'impacts diligentées et discutées pour déterminer les formations

nécessaires ou mesurer les conséquences sur les conditions de travail.

Une technologie imposée se trouvera confrontée à des résistances pouvant limiter les gains attendus.

Dans certains cas, l'utilisation mal maîtrisée de ces nouveaux outils peut conduire à des fuites de données sensibles ou des utilisations erronées des résultats. La formation des salariés est donc une étape essentielle.

Plus généralement, la gestion prévisionnelle des emplois et des compétences doit permettre d'anticiper les conséquences sur le contenu de certains emplois pour mettre en place les formations ou les transitions nécessaires.

Certains logiciels RH traitent de données personnelles et la CFE-CGC estime que la transparence est impérative pour comprendre le fonctionnement de ces outils qui peuvent décider du choix d'un CV ou d'une promotion. Ces aides à la décision doivent être vierges de biais.

L'IA Act posé par l'UE est une régulation nécessaire pour prévenir des utilisations abusives ou dangereuses de ces puissants logiciels, au fonctionnement souvent obscur.

Mais cette technologie se déploie partout et la concurrence fait rage. L'Europe en général, la France en particulier doivent se donner les moyens d'être dans la compétition. Il s'agit donc d'investir dans la formation et dans les technologies matérielles permettant de développer les semi-conducteurs performants, des serveurs puissants, des infrastructures efficaces et des applications conformes aux valeurs de notre continent.

La CFE-CGC considère que les conséquences environnementales doivent aussi être prise en compte dans les choix stratégiques à venir. L'Europe, la France doivent se donner les moyens d'une IA durable est souveraine.

La CFE-CGC analyse cette technologie avec intérêt, parfois enthousiasme mais avec vigilance.

L'Intelligence Artificielle fera partie du monde de demain. Il s'agit de la comprendre, de la contrôler pour qu'elle demeure au service des humains.

La CFE-CGC votera cet avis.

CFTC

En septembre dernier, le groupe CFTC déclarait que, sur la question environnementale, l'IA pouvait être un accélérateur de progrès comme un potentiel accélérateur de catastrophes. Aujourd'hui, à l'issue de cet avis tenant compte de tous les autres aspects de notre société qui seront - ou sont déjà ! - transformés par l'IA, il nous semble que ce constat tient toujours. C'est pourquoi il ne faut pas adopter l'IA dans le seul but de rester dans la course industrielle et technologique. Non, il faut aussi l'adopter dans les buts d'améliorer la condition des travailleurs, d'étoffer notre tissu économique, de favoriser l'inclusion de tous et de redonner de la vitalité à nos services publics.

En clair, il ne faut pas séparer la performance économique de la performance sociale ! Autrement, ce qu'on croyait être des fantasmes deviendront des réalités : perte d'emplois et de sens au travail, subordination de l'Homme à la machine, captation des bénéfices par une minorité d'acteurs économiques, surveillance généralisée et augmentation des discriminations. Pour que l'IA soit une révolution souhaitable, elle doit profiter au plus grand nombre sinon elle ne s'intégrera que dans la défiance et le

rejet.

Parce que l'IA n'est rien d'autre qu'un outil conçu par l'humain, donc faillible, elle doit être strictement encadrée et correspondre aux aspirations démocratiques de notre temps. Ainsi, la CFTC soutient les préconisations visant à accroître la transparence des algorithmes et de l'utilisation des données, à encourager la diversité des profils au sein des métiers de l'IA - notamment la parité entre les femmes et les hommes, à exiger un haut niveau d'accessibilité des outils et à associer les partenaires sociaux dans l'intégration des systèmes d'IA.

En conclusion, c'est bien l'IA qu'on adapte au service de l'intérêt général et non l'humain qu'on adapte à une IA souvent détenue par des groupes privés aux intérêts très particuliers ! La CFTC votera en faveur de l'avis.

CGT

Difficile de cerner l'impact réel de l'IA sur notre société et notre capacité collective à en faire une technologie au service de l'intérêt général, du progrès social et de la lutte contre la crise environnementale.

L'IA permet certes des progrès scientifiques, médicaux ou aide parfois à limiter notre impact environnemental, mais soyons lucides : des milliers de scientifiques et de développeurs alertent sur sa dangerosité. L'IA donne un pouvoir considérable aux quelques géants de la tech. C'est une arme puissante pour déployer des politiques impérialistes, soutenir l'extrême-droite et déstabiliser les démocraties comme le font Donald Trump et Elon Musk. Notre sécurité et nos libertés fondamentales ne sont-elles pas

en jeu ? C'est pourquoi nous partageons pleinement la déclaration du groupe Environnement et Nature.

Notre groupe a fait le choix de l'abstention, qui n'a rien de neutre.

Nos travaux ont confirmé l'impact de l'IA sur les travailleurs du privé ou du public : Amazon les remplace par des robots ; 15 associations ont engagé un recours contre la CAF et son algorithme de notation sociale. À rebours des discours rassurants, l'IA à la DGFIP a été mise en place sans dialogue social et a dégradé les conditions de travail. Le texte intègre bien la philosophie de nos propositions mais sans en afficher une réelle volonté de déploiement dans les entreprises.

Ce contexte lourd de dangers a conduit à l'expression forte d'une demande de moratoire dans la commission. Nous regrettons que l'élaboration du texte final ait été surtout conditionnée par le souhait d'apporter la caution citoyenne du CESE au prochain "Sommet pour l'action sur l'intelligence artificielle" inspiré par les préconisations du rapport Aghion. Celui-ci pourrait déboucher sur une « Organisation mondiale de l'IA pour évaluer et encadrer les systèmes d'IA » au lieu de consolider l'OIT pour mettre en place une norme internationale sur l'IA dans le travail, ce que soutient la CGT.

Nous adressons enfin nos remerciements aux rapporteur.es pour leur patience face à nos nombreuses interventions et à l'ensemble de celles et ceux qui ont contribué à l'organisation du travail - soutenu - de la commission temporaire.

CGT-FO

L'Intelligence Artificielle est un véritable enjeu pour nos démocraties. L'IA peut, sous certaines conditions, constituer une opportunité de développement économique et social. Mais Cette technologie peut aussi nous échapper

et servir à des fins contraires à l'intérêt général et aux libertés nécessaires au bon fonctionnement de nos sociétés. Les questions de sécurité sont par ailleurs essentielles et aucune faille ne doit être tolérée. Pour FO, il est donc important que la société civile s'empare de ce sujet et aide à mieux comprendre l'IA, ses limites et ses risques.

Le groupe FO souscrit aux différentes conclusions de cet avis. Il nous paraît important que le CESE considère que l'IA s'inscrit dans des choix politiques et qu'elle est avant tout une production humaine soumise aux subjectivités de ses choix. L'IA répond aussi à des logiques d'intérêts, de pouvoir et de recherche de domination, c'est un outil géopolitique dont l'intelligence ne se soucie pas de la seule neutralité scientifique. Il faut donc la considérer comme telle et voir comment agir pour une IA plus ouverte, plus utile, plus citoyenne et plus favorable au progrès de nos sociétés.

Les démocraties doivent veiller à ce que l'IA ne constitue pas une atteinte à nos libertés fondamentales, s'assurer de la protection de nos données personnelles et lutter contre la désinformation. Il est de la responsabilité des pouvoirs publics de réglementer l'IA, garantir sa transparence, son objectivité et son ouverture au différents points de vue et à la représentation de l'ensemble de la société. La France et l'Europe ne doivent pas subir l'IA et perdre leur autonomie géostratégique dans ce domaine. Il faut donc des investissements massifs dans tous les domaines de l'IA : R&D, infrastructures, logiciels, etc. et dans tous les secteurs d'activité : santé, industrie, agriculture, information, etc. Il ne peut y avoir d'Europe forte sans maîtrise des technologies liées à l'IA et sans sécurisation et localisation des données des pays de l'UE sur leurs territoires.

Si l'IA peut constituer sous certaines conditions une opportunité de création de

nouveaux emplois un enjeu existe sur la qualité de ces derniers et le niveau de qualifications exigé. Il faut ainsi anticiper les évolutions des métiers et préparer les transitions pour que le monde du travail ne subisse pas cette transformation technologique. Il faut agir pour une IA qui apporte un progrès pour le monde du travail et non une IA comme seulement une opportunité d'augmentation des gains de productivité pour les employeurs.

Le groupe FO remercie le CESE pour ce travail et a voté en faveur de cet avis.

Entreprises

Dans le cadre du sommet sur l'IA prévu en février 2025, le CESE a souhaité s'exprimer sur une technologie qui s'inscrit déjà dans la vie des citoyens français. Les entreprises sont simultanément le réceptacle, les vecteurs et les innovateurs dans ces technologies, sans pour autant se substituer à l'Etat dans son rôle de détermination des nécessaires régulations ; ces évolutions jouent un rôle majeur d'innovation tant en production que dans les process d'organisation, et ce dans un contexte de concurrence mondiale.

Il est essentiel que la société civile exprime ses recommandations et sensibilise aux enjeux de l'IA. D'autant que l'avis présenté aujourd'hui s'articule autour d'autres travaux pour certains déjà réalisés : pour reprendre celui qui traite des impacts de l'IA sur l'environnement, comment ne pas souligner que l'IA est une opportunité pour réduire les risques liés à l'environnement. Pour les travaux en devenir, comme santé et numérique, comment ne

pas remarquer les progrès que cette technologie apporte dans les traitements. Le groupe Entreprises en est persuadé : dans un paysage en constante évolution, l'IA offre des possibilités positives d'évolution.

Le groupe Entreprises souhaiterait souligner quelques aspects sur les nouvelles préconisations du CESE ici présentées :

- Tout d'abord, il invite à une nécessaire législation qui défend et soutient les organisations privées et publiques européennes. Nous encourageons les dispositions permettant un subtil équilibre entre d'une part l'encouragement à l'innovation et la recherche, la prise en compte de la concurrence mondiale et, d'autre part, l'indispensable régulation. Comme l'indique le rapport de Mario Draghi et pour éviter la relégation dans le domaine technologique, il faut investir plus, de façon plus ciblée et il faut moins réguler.

- Le groupe Entreprises se félicite par ailleurs des préconisations visant à développer ou à renforcer l'autonomie stratégique de l'UE : le financement de la recherche et l'harmonisation le droit des affaires sont plus que nécessaires. De même mobiliser les outils du droit à la concurrence pour permettre l'émergence d'opérateurs européens est essentiel.

- Le CESE souligne avec raison l'importance de la formation et de l'attractivité des talents. Si cela est indispensable, le groupe Entreprises estime que le sujet de la formation n'est plus celui d'hier : une formation continue plus agile et centrée sur les capacités des personnes à effectuer une veille se dessine. Cela pose la question

de l'autonomie de ces dernières et du changement managérial induit. Pour les chefs entreprises, la conduite du changement sera primordiale.

- L'emploi est au cœur des préoccupations et nul doute que certains métiers sont et seront impactés, de même que de changements de comportements et de modes de consommation. Sans entrer dans les polémiques des analyses variées, l'emploi s'est toujours transformé. Moyennant des formations renforcées des salariés, les conditions de travail peuvent grandement s'améliorer. La communication et le dialogue social sont des clés importantes pour un bon usage de l'IA. Citons trois outils existants : charte d'utilisation, information du CSE, des négociations obligatoires de type Qualité de Vie et Conditions de Travail voire Gestion Prévisionnelle des Emplois et des Parcours. Sans préjuger de la forme, le dialogue social joue un rôle essentiel.

- Enfin, le groupe Entreprises se félicite de la part importante réservée à l'éducation et la formation initiale dans cet avis. Enjeu démocratique et essentiel pour tous, les pouvoirs publics, l'école, l'enseignement professionnel et supérieur, les entreprises doivent collaborer pour élaborer des programmes. Repenser les apprentissages, reformer les méthodes, outiller les enseignants sont des fondamentaux.

S'il ne partage pas l'ensemble des recommandations, le groupe Entreprises salue les orientations contenues dans cet avis et son ambition. En conséquence, le groupe Entreprises vote favorablement ce texte.

Environnement et nature

Pour nous, une réflexion sur l'IA aurait dû avant tout, s'attacher à faire le tri entre les usages pouvant présenter de vrais avantages collectifs, comme

dans les domaines de la médecine ou de la recherche, et ceux présentant, en revanche, de réels dangers pour la démocratie, les libertés et les droits fondamentaux.

En la matière, les alertes sont nombreuses et l'actualité foisonnante. Déjà, en mai 2023, une tribune signée de scientifiques spécialisés et de géants de la tech appelait à une politique d'atténuation des risques d'extinction humaine que l'IA soulève comme les pandémies ou les armes nucléaires. Auparavant, une autre tribune signée de 33000 scientifiques appelait à un moratoire dans le développement de nouveaux modèles d'IA jusqu'à ce qu'une réglementation puisse être mise en place.

Bien que l'avis reconnaisse que l'IA est un objet politique, son élaboration n'a pas permis de mener un vrai travail de fond, notamment à cause de délais trop contraints. Il en résulte que le texte souffre d'une contradiction irrésolue entre sa finalité de court terme et le traitement raisonné de ces grands enjeux.

Des risques majeurs sont liés directement à la technologie, comme le mensonge, la trahison, ou encore le piratage que certaines IA décident d'elles-mêmes, en violation des règles qui leurs sont imposées par leurs propres concepteurs. L'effet « boîte noire » les amplifie en empêchant d'en comprendre les raisons.

A ceci s'ajoute une orientation marquée. Par exemple, les préconisations concernant l'impératif d'autonomie stratégique relaient les propositions de groupes d'intérêt privés sous couvert d'assurer les droits et la sécurité de tous. Or, seules des régulations rigoureuses garantissant des droits pourront protéger des dérives de l'utilisation et du déploiement de tels outils par quelques-uns.

Concernant les questions environnementales, ce n'est pas non plus en affirmant qu'il faut respecter le

cadre fixé par les limites planétaires que cela va le permettre. Aucune préconisation ne permet à minima d'en envisager la possibilité alors même que les conflits d'usage et l'utilisation des ressources explosent.

Nous regrettons profondément que l'avis n'ait pas pris la mesure de ces défis. La plupart des préconisations sont surtout des vœux pieux, mais non opératoires.

La participation de citoyens à cet avis est aussi pour nous un échec. En effet, nous n'avons pas réussi à permettre leur véritable implication.

Dans une logique de complémentarité, nous tenons à signaler que nous partagerons également la vision et les apports du groupe de la CGT.

Pour ces raisons, notre groupe partagera ses votes entre abstention et opposition.

Familles

Être ou ne pas être pour l'intelligence artificielle, là n'est plus la question. L'IA existe ; elle est à l'œuvre dans de nombreux domaines depuis plusieurs décennies ; elle est désormais présente dans notre quotidien au travers, notamment, de l'IA générative.

Comme le souligne le présent avis, l'IA a permis et continue de permettre des avancées considérables dans des domaines essentiels tels que la santé, l'accessibilité ou l'agriculture. La question qui se pose aujourd'hui est celle de ses limites, limites qui doivent répondre à certaines inquiétudes quant à l'utilisation des ressources naturelles, quant à la protection des données personnelles, quant aux biais

que pourraient comporter les applications ou quant aux impacts qu'elle a ou qu'elle pourrait avoir sur l'emploi, le fonctionnement des entreprises et des services publics. Et c'est à ces inquiétudes que répond une grande part des préconisations.

Le Groupe Familles se soucie particulièrement de l'utilisation qui peut être faite de l'IA dans le domaine de l'éducation. Là aussi doivent être mis en regard aspects positifs et négatifs. L'IA ne soit pas être une solution de facilité. Mais elle peut être un outil au service de l'enseignement. Encore faut-il apprendre à s'en servir et développer l'esprit critique nécessaire pour éviter toute manipulation des esprits.

Au-delà de ces préoccupations sur l'utilisation de l'intelligence artificielle, l'avis insiste sur l'indépendance que doit retrouver notre pays et plus largement l'Europe. Nous ne pouvons pas rester à la remorque des Etats-Unis et de la Chine et nous contenter de réguler un phénomène qui nous échappe. Il est indispensable que soient réunies les conditions d'une autonomie stratégique européenne et que se développe en Europe une industrie de l'IA.

L'avis a su maintenir un équilibre entre les craintes que suscite l'IA et ses aspects positifs et c'est pour cela que le Groupe Familles a adopté l'avis. Il a formulé le vœu en ce début d'année 2025 que les travaux du CESE puissent utilement alimenter les réflexions du Sommet pour l'action sur l'intelligence artificielle (IA) des 10 et 11 février prochains.

Non-Inscrits

L'apparition de l'IA constitue une accélération encore jamais connue des progrès technologiques exerçant une influence majeure sur le fonctionnement des sociétés humaines.

Les débats à son sujet oscillent entre discours alarmistes, façon science-fiction et espérance de voir l'IA résoudre tous les maux dont souffrent notre société.

Avec la puissance numérique et la globalisation mondiale de « l'IA », nous sommes à la croisée d'une nouvelle civilisation : Elle modifie le monde économique, celui du travail, celui des libertés, de la défense de la vie privée, et celui de notre cadre environnemental.

La nécessaire protection des libertés individuelles, - soulignée par l'avis - fait prendre conscience, aux autorités publiques nationales ou européennes, d'un nécessaire rééquilibrage entre la puissance de géants économiques mondiaux, ayant acquis une considérable avance technologique, et les citoyens, les sociétés, ou les États.

Les condamnations rendues par la CNIL Irlandaise (Septembre 2024) par la communauté Européenne (Octobre 2024) ou par la Cour de Justice des communautés (en 2023) viennent rappeler ce besoin de réglementation. Les préconisations 3 et 4 de l'avis, n'en trouvent que plus d'acuité.

Selon les opérateurs, ces interdictions sont un frein au développement des applications « IA » dans l'Union Européenne.

A l'inverse l'IA peut se révéler facteur de progrès dans le domaine médical.

Le prix Nobel de chimie est ainsi venu récompenser les travaux de 2 chercheurs ayant exploité l'IA pour établir un changement de paradigme dans l'ingénierie moléculaire.

C'est donc le débat habituel entre progrès ou liberté, obscurantisme ou lumière.

Chaque fois qu'il s'est posé, comme pour la révolution industrielle du 19ème siècle, ou l'utilisation de l'atome dans les années trente, les penseurs ont toujours fait appel au « Discours de la Méthode » de Descartes pour poser les termes du dilemme : « La réponse tient avant tout dans la perception des conséquences associées à telle ou telle proposition ».

Pour ma part, sur le débat entre préservation des libertés individuelles et liberté économique sans borne, ni mesure, j'en appellerai au Père Lacordaire, penseur chrétien social du 19° : « entre le faible et le fort, c'est la liberté qui opprime et la loi qui protège ».

Le groupe des non-inscrits a voté l'avis.

Et qu'il me soit ici permis, puisque c'est la dernière intervention que je ferai dans cet hémicycle de remercier Yazid CHIR, membre de la commission temporaire, avec lequel nous avons travaillé sur le sujet.

Je remercie également chacun d'entre vous pour les quinze années de dialogue d'amitié et de travaux constructifs que j'aurai passé dans cette institution.

Organisations étudiantes et mouvements de jeunesse

Face à la montée en puissance des systèmes d'intelligence artificielle, certains et certaines font preuve d'une grande politesse à leur égard, s'assurant de bien utiliser les usuels bonjours, merci et s'il vous plaît pour ne pas se retrouver en première ligne de leurs victimes en cas de prise de pouvoir des IA sur les humains. Mais doit-on attendre d'en arriver là pour réfléchir à la place de l'IA dans nos vies ? Est-on vraiment face à un choix binaire de s'en éloigner ou de se faire dépasser ?

Il n'est pas simple de définir une position

claire vis-à-vis d'une nouvelle technologie qui apporte avec elle son lot de fantasmes mais aussi beaucoup de doutes. Pourtant, les nombreux enjeux de société qu'elle soulève nous obligent à ouvrir une réflexion sérieuse sur la manière dont nous voulons nous en emparer.

Tout au long des auditions, s'est dégagé pour notre groupe, un principe de base : l'IA est déjà là ; alors comment faire pour l'encadrer et la régler au mieux afin qu'elle devienne véritablement vectrice de développement humain ? Réfléchir à une éthique de l'IA, c'est s'assurer qu'elle soit configurée pour contrecarrer les biais et stéréotypes humains plutôt que pour les reproduire bêtement (vis-à-vis de l'égalité femmes-hommes par exemple). C'est s'assurer de valoriser ce qui fait la spécificité de l'intelligence humaine, en encourageant par exemple la créativité plutôt que la multiplicité et la rentabilité dans le monde de la culture, ou la relation entre l'enseignant-enseignante et son élève dans l'éducation.

Pour servir à toutes et à tous et pour être véritablement vecteur de bien-être, les révolutions technologiques, comme les révolutions industrielles, doivent aussi venir questionner nos modes d'organisation capitalistes. A qui profitent les gains de productivité permis par l'IA ? A celles et ceux qui pensent leur place dans la société au prisme de la maximisation de leur richesse ou à celles et ceux qui profiteraient bien de quelques heures de travail en moins pour aider leurs proches, participer à la vie collective et associative ou se reposer ?

Le travail a été quelque peu agité en commission pour parvenir à

construire la logique politique de ce texte. Les temporalités imposées n'ayant pas permis de monter correctement en expertise, de structurer clairement toutes les problématiques et d'assurer tous les débats nécessaires, notre groupe a choisi une position de prudence (mais pas de blocage) : l'abstention. L'intérêt de notre institution à se positionner sur le sujet repose justement dans nos différences et notre capacité à créer de la nuance, à valoriser l'innovation tout en posant de sérieux et importants garde-fous, qu'ils soient environnementaux, éthiques, économiques ou sociaux. Aujourd'hui, nous gardons une réserve quant à la portée politique qu'aura ce travail, que nous formulons comme un vœu de vigilance à respecter le consensus trouvé.

Nous remercions toutes les personnes mobilisées sur cet avis, rapporteurs, administrateurs, attachés et toutes les citoyennes et citoyens de cette commission temporaire.

Outre-mer

L'idée que la dimension artificielle de la technique entraînera indubitablement un péril civilisationnel n'est pas nouveau. En 1946, BERNANOS dénonçait déjà le triomphe de la civilisation des machines. Il affirmait même qu'un « monde gagné par la technique est perdu pour la liberté ».

La révolution technologique et numérique qu'est l'intelligence artificielle n'en fait pas exception notamment par la fulgurante et massive émergence de ce qu'on appelle l'IA générative.

Cette révolution technologique où doit-on dire cette évolution tant son existence est ancienne, inquiète tout autant qu'il fascine les optimistes dans sa capacité à changer la face du monde et à affecter tous les domaines d'activité.

Cet avis brosse les grands enjeux de l'IA et de ses impacts dans la société en traitant plus spécifiquement de ses effets sur l'emploi, son intégration en entreprise, au domaine de l'éducation et de la santé. Toutefois, en tant que représentants de la société civile et des collectivités géographiquement éloignées de l'Hexagone, les conseillers du Groupe des Outre-mer auraient souhaité, comme on l'a suggéré, de traiter également l'IA comme vecteur de développement durable et d'attractivité des territoires. On n'a malheureusement pas suffisamment été entendu et on le regrette.

L'émergence de l'Intelligence artificielle dans la société doit être regardée comme autant d'innovation au profit de la réduction des inégalités sociales, de développement durable et collectif. Cependant, ces bénéfices ne se feront ni sans projet politique, ni sans engagement collectifs et encore moins sans investissement public et privé.

Si la France et l'Europe ne veulent pas être tributaire, sous domination anglo-saxonne à appliquer les technologies des autres, elle doit impérativement actionner plusieurs leviers. Celui de l'accessibilité des données, du déploiement des capacités d'investissement, de formation mais surtout de la mobilisation de notre communauté de chercheurs tout en se dotant d'outils juridique et de protection.

L'IA appelle à faire preuve de réalisme, de pragmatisme et d'engagement collectif car comme le disent les experts de la Commission de l'Intelligence artificielle : le premier des enjeux consiste peut-être à dédramatiser l'IA sans pour autant l'idéaliser.

Le Groupe des Outre-mer a voté cet avis.

Santé et citoyenneté

Le développement accéléré de l'IA, la désinformation et la propagation des extrêmes idéologiques, la reproduction, voire l'intensification, des inégalités engendrées par les algorithmes et l'IA générative, nous font craindre pour l'avenir de notre démocratie.

L'enjeu démocratique apparaît déterminant pour nos sociétés et leur avenir tant la vitesse des (r)évolutions technologiques nous dépasse et transforme profondément l'espace public, et ce, en plus des enjeux environnementaux, de souveraineté et de sécurité, économiques et sociaux, éthiques et philosophiques ou encore éducatifs et culturels.

Renforcer la transparence et la traçabilité des algorithmes tout en favorisant la diversité des données pour prendre en compte la diversité des populations, mais aussi la diversité des profils de celles et ceux qui font l'IA, apparaît donc comme un préalable. C'est dans ce sens que notre avis s'attache à rechercher comment supprimer les biais qui perpétuent une société inégalitaire.

L'IA est sexiste car elle se base sur des clichés sexistes.

L'IA accroît les discriminations car les données qu'elle utilise sont elles-mêmes porteuses des biais de notre société.

Bien sûr les technologies du numérique, dont l'IA, ouvrent également des voies d'espérance, des voies de progrès, comme dans le domaine de la santé, où elle peut améliorer la qualité des soins et où elle doit faciliter l'accès aux soins.

Mais dans ce domaine, comme dans d'autres, le groupe Santé et Citoyenneté rappelle l'importance des relations humaines et de la protection des données sensibles.

Par ailleurs, il nous faut sensibiliser et échanger sur la culture des données et de l'IA pour renforcer l'accessibilité, l'égalité et lutter contre la désinformation ... et in fine préserver et défendre notre démocratie. Notre avis participe à cette démarche d'acculturation.

En conclusion, le groupe Santé et Citoyenneté fait sienne les propos du sociologue Dominique Cardon : «Face aux enjeux de l'IA, comme pour toutes les innovations technologiques, la question n'est pas d'être pour ou contre, mais de dessiner un cadre et des règles pour que son utilisation soit vertueuse».

Le groupe Santé & Citoyenneté a voté l'avis.

UNSA

Dans le prolongement du déploiement du numérique et de la capacité de traitement de données de grand volume, l'Intelligence Artificielle apparaît comme une étape de virtualisation et d'intermédiation modifiant profondément nos référentiels et nos équilibres.

Ses impacts sont perçus comme autant d'opportunité de progrès (comme la santé, l'inclusion) qu'un potentiel danger violent contre nos démocraties et l'ensemble des relations interpersonnelles y compris dans le cadre professionnel. A ce titre le domaine d'application, mais plus encore l'intention, les objectifs qui y sont associés doivent d'être posées, explicitées et faire l'objet d'une réflexion commune entre les différentes parties prenantes.

L'usage, et ses impacts sur l'environnement social et les ressources naturelles sont les

axes essentiels à interroger, à analyser collectivement pour définir un chemin commun respectant les valeurs fondamentales de notre société : protection de la vie privée, diversité et régulation de la « propagande », droit à une information fiable et de qualité, droit à la dignité, etc En outre la consommation grandissante en énergie et eau, interpelle sur la nécessité de planifier équitablement sur les territoires et d'instruire les conditions d'une juste répartition des besoins pour éviter des conflits d'usages. A ce titre, l'UNSA rejoint les préconisations, à la fois de l'avis de la commission environnement et les préconisations du présent avis en insistant sur le caractère démocratique des décisions.

Si la régulation, notamment des données, se construit à l'échelle européenne, l'UE se trouve assez démunie, par son absence en tant qu'acteur dans la structuration primaire du secteur, aussi par sa faiblesse industrielle et technologique pour déployer massivement cette technologie selon ses propres codes et normes explicites comme implicites. Outre les aspects commerciaux, ce sont aussi des modèles de société qui sont véhiculés par l'instruction des IA, surtout quand cette dernière accentue les stéréotypes et discriminations, et l'actualité nous montre que les intentions peuvent être aussi des plus belliqueuses !

A une échelle plus micro, l'IA peut tout autant faire augmenter les effets délétères d'une hyper segmentation et prédétermination des situations, des femmes et hommes, des tâches, rendant les relations humaines apaisées et libres comme une situation particulièrement privilégiés,

et l'UNSA s'inquiète d'une vision productiviste et étriquée de l'IA dans le monde du travail, alors que l'IA peut être l'occasion idoine pour réinterroger le modèle productif, les relations comme les hiérarchisations en mettant plus en avant l'apport humain dans sa créativité, sa subtilité et aussi les savoirs faire manuels et relationnels versus un reporting et management vertical sur des référentiels sociaux déjà obsolètes.

Mais cette réflexion n'est possible que si l'ensemble du secteur se prête à cette remise en question et ne cherche pas un ultime refuge dans l'hyper subordination, et un remplacement tâche par tâche niant les risques de conflits de rationalité.

Le travail mené au sein de la commission travail emploi avec l'étude de controverses doit être souligné, notamment car il permet d'apporter de la rationalité dans un débat qui en manque cruellement. L'ensemble du travail aurait pu constituer une préconisation, il se retrouve diluer dans plusieurs, ce qui est dommage.

L'IA produit d'autres besoins, et organisation, entretien, alimentation et adaptation progressive de l'outil, chaîne de valeur, et seul un dialogue social technologique avec implication des travailleurs au processus d'innovation dès le début permet de formaliser le « compromis de rationalité » en partant du travail réel, et en faisant des travailleurs des « instructeurs » de l'IA.

Pour cela, la formation initiale et continue doit être à la hauteur d'une véritable assimilation des enjeux et risques, et donc nécessite un déploiement autrement plus conséquent, notamment dans les établissements scolaires, de moyen technique et ressources en instructeur qualifiés.

Les potentialités, bonnes ou mauvaises, induites par le déploiement massif des systèmes d'intelligence artificielle doit nous interroger, nous questionner sur

l'impact de ces systèmes sur nos vies, personnelles et professionnelles. Les choix fait aujourd'hui détermineront en grande partie notre capacité à nous épanouir demain, individuellement et collectivement. Cet impact, peut être le plus massif depuis des décennies, doit pouvoir être analysé, débattu démocratiquement, avec l'ensemble des personnes concernées, dépassant les cadres prédéfinis du dialogue, s'élargissant alors pour que chacun décide, en conscience, qui reste le propre de l'être humain.

Scrutin

**Scrutin sur l'ensemble de l'avis.
Le CESE a adopté.**

**Nombre de votantes
et de votants : 134**

Pour : 102

Contre : 10

Abstentions : 22

Ont voté pour : 102

GROUPE	COMPOSITION
Agir autrement pour l'innovation sociale et environnementale	Mme Djouadi, MM. El Jarroudi, Hammouche, Mmes Roux de Bezieux, Tordeux Bitker, Tutenuit.
Agriculture	MM. Amécourt (d'), Biès-Péré, Mme Blin, MM. Coué, Dagès, Durand, Férey, Mmes Lion, Vial.
Alternatives sociales et écologiques	Mmes Calmels, Gondard-Lalanne, Groison, M. Meyer.
Artisanat et professions libérales	MM. Anract, Chassang, Fourny, Mmes Munoz, Niakaté, M. Repon, Mme Vial.
Associations	Mme Belhaddad, MM. Bobel, Deniau, Deschamps, Mme Martel, M. Miribel, Mmes Monnier, Sivignon, M. Thomasset, Mme Thoury.
CFDT	Mme Blancard, M. Cadart, Mmes Caillet, Duboc, Gresset Bourgeois, MM. Guihéneuf, Lautridou, Mariani, Mmes Meyling, Pajarès y Sanchez, M. Ritzenthaler, Mme Thiéry.
CFE-CGC	Mme Biarnaix-Roche, M. Nicoud.
CFTC	Mmes Chatain, Coton, M. Heitz.

CGT-FO	MM. André, Cambou, Mmes Delaveau, Marot, M. Sabot.
Coopération	M. Landriot, Mme Lienemann.
Entreprises	MM. Asselin, Blachier, Chevée, Mme Couderc, M. Creyssel, Mme Dubrac, MM. Gardinal, Goguet, Mme Guerniou, MM. Julier, Kling, Moisselin, Mme Pauzat, M. Ruchenstain, Mmes Salvadoretti, Tomé-Gertheinrichs, Ullern, MM. Vermot Desroches, Vidor.
Familles	Mme Balducci, MM. Desbrosses, Erbs, Mme Gariel, M. Marmier, Mmes Picardat, Quaranta.
Non-inscrits	Mme Beaufls, MM. Breton, Joseph, Noël.
Outre-mer	Mme Bouchaut-Choisy, MM. Leung, Marie-Joseph, Yan.
Santé et Citoyenneté	MM. Boroy, Da Costa, Mme Joseph, M. Raymond.
UNSA	Mme Arav, MM. Darwane, Truffat, Mme Vignau.

Contre : 10

GRUPE	COMPOSITION
Environnement et nature	M. Abel, Mme Claveirole, M. Gatet, Mmes Journée, Marsaud, Martinie-Cousty, MM. Mayol, Métails, Mmes Popelin, M. Richard..

Se sont abstenus : 22

GRUPE	COMPOSITION
Associations	Mme Doresse Dewas.
CGT	Mmes Barth, Belghazi, M. Fontanié, Mme Hoeffling, MM. Le Corre, Michaud, Rabhi, Mmes Rouchy, Tatot, M. Vire.
Environnement et nature	MM. Beauvais, Chabason, Compain, Mme Ostria.
Non-inscrits	MM. Bazot, Chir, Pouget.
Organisations étudiantes et mouvements de jeunesse	M. Eyriey, Mme Hamel, MM. Jeanne-Rose, Occansey.

Annexes

1

Composition de la commission temporaire

Président

Thierry CADART

Agir autrement pour l'innovation sociale et environnementale

Marianne TORDEUX BITKER

Samira DJOUADI (suppléante)

Saïd HAMMOUCHE (suppléant)

Agriculture

Henri BIÈS-PÉRÉ

Catherine LION

Anne-Claire VIAL

Alternatives sociales et écologiques

Érik MEYER

Cécile GONDARD-LALANNE (suppléante)

Artisanat et professions libérales

Michel CHASSANG

Associations

Marie-Claire MARTEL

Jean-Pascal THOMASSET

CFDT

Thierry CADART

Pascal GUIHÉNEUF

Marie-Hélène MEYLING

Patricia BLANCARD (suppléante)

Jean-Yves LAUTRIDOU (suppléant)

Franck TIVIERGE (suppléant)

CFE-CGC

Fabrice NICOU

Djamel SOUAMI (suppléant)

CFTC

Pascale COTON

CGT

Maurad RABHI

Fabienne ROUCHY

Fabienne TATOT

CGT-FO

Dominique DELAVEAU

Hervé QUILLET

Jean-Yves SABOT (suppléant)

Sylvia VEITL

Coopération

Jacques LANDRIOT

Christophe GRISON (suppléant)

Entreprises

Dominique CARLAC'H

Gilles VERMOT DESROCHES

Pierre-Olivier RUCHENSTAIN

Environnement et nature

Florent COMPAIN

Venance JOURNÉ

Cécile OSTRIA (suppléante)

Familles

Pierre ERBS

Marie-Pierre GARIEL (suppléante)

Non inscrits

Yazid CHIR

Corine BEAUFILS (suppléante)

**Organisations Étudiantes et
Mouvements de jeunesse**

Évanne JEANNE-ROSE

Helno EYRIEY (suppléant)

Outre-mer

Éric LEUNG

Pierre MARIE-JOSEPH (suppléant)

Santé et citoyenneté

Dominique JOSEPH

Gérard RAYMOND (suppléant)

UNSA

Fanny ARAV

Jean-Marie TRUFFAT (suppléant)

Citoyennes et citoyens

Habiba A.

Maud A.

Martine D.

Myriam F.

Hicham I.

Saïd K.

Jonathan L.

Frédéric M.

Yannick S.

Lucas V.

2

Liste des auditionnés

En vue de parfaire son information, les rapporteurs et les membres de la Commission temporaire « Intelligence artificielle » ont reçu en audition les personnalités suivantes :

Théo ALVES DA COSTA

Directeur de l'unité « IA et transition écologique » dans le cabinet français spécialisé en data Ekimetrics, co-président de l'ONG « Data For Good »»

Philippe ARCHIAS

Directeur conseil au Cabinet Julhiet Sterwen

Frédéric BARDEAU

Président et co-fondateur de l'Institut des futurs souhaitables (IFS)

Adrien BASDEVANT

Avocat spécialisé en droit des nouvelles technologies, membre du Conseil national du numérique (CNUM), fondateur d'Entropy, cabinet d'avocats spécialisé en technologies de rupture

Bilel BENBOUZID

Maitre de conférences spécialiste de la sociologie de l'action publique et de la sociologie digitale à l'université Gustave Eiffel

Nathalie BESLAY

Avocate, fondatrice de Naaia, plateforme d'aide à la mise en conformité de l'intelligence artificielle (IA) avec l'AI Act européen

Benjamin BESSE

Oncologue médical spécialiste de la prise en charge des cancers thoraciques, directeur de la recherche clinique de la clinique Gustave Roussy

Nicolas BLANC

Secrétaire national à la transition économique Contributeur au Policy Brief du GPAI sur IA et emploi CFE-CGC

Delphine BLUMEREAU

Responsable du Pôle opportunités de l'IFS

Jérémie BOROY

Membre du CESE, Président du Conseil national consultatif des personnes handicapées (CNCPPH)

Benjamin BÜRBAUMER

Économiste, maître de conférences au Centre Émile Durkheim de Sciences Po Bordeaux

Céline COLUCCI

Déléguée générale de l'association Les Interconnectés, réseau des territoires connectés

Axel DAUCHEZ

Entrepreneur, fondateur de Make.org, plateforme de mobilisation citoyenne

Ysens DE FRANCE

Docteure en droit chargée de mission IA à la Direction générale de la Gendarmerie nationale du ministère de l'Intérieur et des Outre-mer

Sandra DEMARCO

Secrétaire nationale du syndicat Solidaires Finances Publiques

Laure DE ROUCY-ROCHEGONDE

Responsable du centre Géopolitique des technologies de l'Institut français des relations internationales (IFRI)

Marlène DOLVECK

Directrice générale de SNCF Gares & Connexions et Directrice générale adjointe du groupe SNCF en charge de la Transformation et du Numérique

Laëtitia DOMANGE

Directrice du pôle agricole de la Coopérative Maïsadour

Maxime EFOUI-HESS

Coordinateur de projet technologies numériques, industries, matières chez The Shift Project

Thomas FAUVEL

Ingénieur Télécoms pour Orange et président de « 100 % Handinamique » au CNCPH

Benjamin GANDOUIN

Secrétaire national du syndicat Solidaires Finances Publiques

Albane GRANDJEAN

Consultante chez Gleamer, plateforme d'intelligence artificielle qui assiste les radiologues

Jean GUO

Directrice et co-fondatrice de Konexio

Pierre JANNIN

Conseiller municipal à la ville de Rennes, délégué au numérique et à l'innovation

Christine LAVARDE

Sénatrice des Hauts-de-Seine, Présidente de la délégation à la prospective du Sénat

Yann LEHELLE

Informaticien, entrepreneur co-fondateur de Probabl, entreprise à Mission de Souveraineté Industrielle et Numérique

Laure LUCCHESI

Entrepreneure, ex-Directrice générale d'Etalab

Dominique MÉDA

Sociologue et philosophe du travail

Tristan MENDÈS-FRANCE

Enseignant dans le domaine du numérique, spécialiste des cultures numériques et de l'extrémisme en ligne à l'Observatoire du conspirationnisme

Grégory MOULIS

Directeur productions végétales à la Coopérative Maisadour

Edlira NANO

Informaticienne, membre de La Quadrature du Net

Julien NICOLAS

Directeur numérique du groupe SNCF

Aurélien PALIX

Sous-directeur des réseaux et des usages du numérique à la Direction générale des entreprises (DGE) du ministère de l'Économie et des finances

Sébastien PICARDAT

Directeur général d'AgParisDataHub

Hugo RUGGIERI

Directeur juridique et affaires publiques de la plateforme d'intelligence juridique Doctrine.fr

Franca SALIS-MADINIER

Secrétaire nationale de la CFDT Cadres, chargée de l'Europe et de l'international, du numérique et de l'intelligence artificielle, vice-présidente du Comité économique et social (CES) européen

Arthur TALPAERT

Chef de produit IA, assistant de consultation chez Doctolib

Isabelle THOUMIN

Membre du CESER Pays-de-la-Loire

Félix TRÉGUER

Chercheur associé au Centre Internet et société du CNRS, membre de La Quadrature du Net

Martin ULBRICH

Expert confirmé de l'unité A2 « Règlement et Conformité en matière d'intelligence artificielle » de la Direction générale (DG) CONNECT, bureau européen de l'IA, de la Commission européenne

Camille VAZIAGA

Directrice des affaires publiques chez Doctolib

Renaud VEDEL

Préfet, conseiller d'État, ex-coordonateur de l'IA, ex-directeur de cabinet des ministres du numérique
Jean-Noël BARROT et Marina FERRARI

Anne-Sophie VIERA

Directrice de la communication
Coopérative Maisadour,

Flora VINCENT

Docteure en sciences, chercheuse
à l'European Molecular Biology
Laboratory (EMBL)

Le Président, les rapporteurs ainsi que les membres de la Commission temporaire remercient ces personnes pour leur contribution à l'élaboration de cet avis.

3

Compte-rendu des modalités d'association des citoyennes et citoyens aux travaux de la commission temporaire « intelligence artificielle » annexé à l'avis conformément à l'article 26 du règlement intérieur

• Objectifs :

L'IA est omniprésente dans nos vies depuis trente ans sous différentes formes et ce sans que nous en soyons avertis. L'IA soulève de nombreux enjeux et concerne ou va concerner chacune et chacun dans sa vie quotidienne. C'est pourquoi le CESE a choisi d'intégrer 12 citoyennes et citoyens à ses travaux pour enrichir la réflexion des organisations de la société civile au sein d'une commission temporaire « mixte », composée de 12 citoyennes et citoyens et 29 conseillères et conseillers.

L'objectif de la participation de citoyennes et citoyens à ces travaux était d'enrichir la réflexion de la commission par la diversité de leurs situations, expériences vécues, territoires : leur quotidien et les impacts et utilisations possibles de l'IA pour elles et eux, etc.

Afin de garantir une bonne intégration et de favoriser la participation effective, le Président de la commission et les co-rapporteurs se sont engagés envers les citoyennes et citoyens sur les points suivants :

- les associer à l'orientation stratégique et à l'élaboration du projet d'avis ;
- mettre en œuvre des méthodologies leur permettant d'être à l'aise pour la prise de parole (par exemple : des temps de travail en sous-groupes, des temps d'accompagnement en visio, etc.) ;
- afficher clairement leur contribution à l'avis en leur ouvrant un espace d'expression libre dans le texte (courte déclaration sur leur expérience, leur

opinion sur le texte, etc. qui sera annexée à l'avis) ;

- les informer et les associer à la phase de portage de l'avis.

• Modalités de recrutement des citoyennes et citoyens :

Pour recruter les 12 citoyennes et citoyens, un prestataire a été missionné conformément aux règles des marchés publics : État d'Esprit - Stratis. Le recrutement des citoyennes et citoyens a été réalisé selon une méthode permettant d'assurer l'unité des méthodes de recrutement, et avec un objectif de diversité, la représentativité étant impossible à atteindre sur un si petit panel.

Le recrutement a été réalisé selon la méthode suivante :

En amont, 5 critères ont été retenus et définis :

Le genre :

Femmes : 6

Hommes : 6

L'âge :

18 – 27 ans : 3

28 – 40 ans : 3

41 – 54 ans : 3

55 – 65 ans : 3

L'origine géographique, selon deux axes :

La représentation de *toutes les régions métropolitaines*, à l'exception de la Corse ;

La diversité des *territoires urbains et ruraux* :

commune urbaine : 4

commune péri-urbaine : 4

commune rurale : 4

La catégorie socio-professionnelle :

Agriculteurs exploitants : 1

Artisans, commerçants et chefs d'entreprises : 1

Cadres et professions intellectuelles : 1 ou 2

Professions intermédiaires : 1 ou 2

Employés : 1 ou 2

Ouvriers : 1 ou 2

Retraités : 1 ou 2

Étudiant : 1 ou 2

Sans activité : 1 ou 2

Le niveau d'étude :

Aucun diplôme ou certificat d'études primaires : 2 ou 3

BEPC, brevet des collèges, DNB : 1

CAP /BEP ou équivalent : 2 ou 3

Baccalauréat, brevet professionnel ou équivalent : 1 ou 2

Diplôme de l'enseignement supérieur de niveau bac +2 : 1 ou 2

Diplôme de l'enseignement supérieur de niveau bac +3 ou bac+4 : 1

Diplôme de l'enseignement supérieur de niveau bac+5 ou plus : 1 ou 2

Le tirage au sort a ensuite été réalisé en 4 étapes majeures, entre mi-juillet et mi-août 2024 :

→ 1. Le prestataire dispose d'une large base de données de personnes ayant déclaré un intérêt pour participer à des consultations citoyennes (fichier constitué par des campagnes de communication régulières : réseaux sociaux, recrutement de terrain, etc.) ;

→ 2. De cette base de données sont exclues les personnes ayant déjà fait partie d'un dispositif de participation citoyenne ou ne répondant pas aux critères du panel ;

→ 3. Un tirage au sort est réalisé au sein du fichier, composé de plusieurs dizaines de milliers de personnes. Les personnes tirées au sort reçoivent un mail afin de récolter leur intérêt et de compléter les informations de profil nécessaires à la constitution d'un panel ;

→ 4. Les personnes ayant répondu au mail sont contactées de manière aléatoire, par téléphone, afin d'être sélectionnées pour rejoindre le panel. Les personnes qui s'engagent sont retenues et les personnes correspondantes au même profil ne sont plus contactées, jusqu'à remplissage des 12 sièges citoyens au sein de la Commission.

Étapes de l'association des citoyennes et citoyens aux travaux, missions confiées, accompagnement et ressources mises à leur disposition :

Afin de s'assurer de la bonne intégration des citoyennes et citoyens, une "personne-ressource" a été mise à leur disposition pour les accompagner dans leur intégration, leur compréhension du contexte institutionnel et pour mener à bien leurs missions. Des garants ont également été nommés afin de vérifier que le processus répondait à des exigences de sincérité, d'égalité, de transparence et d'impartialité, notamment en se tenant à l'écoute des citoyennes et citoyens en cas de difficultés éventuelles.

Les citoyennes et citoyens ont été intégrés dès le début des travaux de la Commission Temporaire, afin de pouvoir participer pleinement à toutes les étapes du travail en commission. Elles et ils ont donc participé à toutes les réunions de travail de la commission, en présentiel ou de manière exceptionnelle en visioconférence lorsque cela était nécessaire. Le taux de présence des citoyennes et citoyens aux réunions de travail de la commission a été de 94%.

Tout au long des travaux, Les citoyennes et citoyens ont été rendus destinataires des mêmes documents de travail que ceux adressés aux membres du CESE désignés dans la commission.

Plusieurs temps dédiés ont été mis en place afin d'assurer l'intégration des citoyennes et des citoyens au sein de la commission, certains uniquement entre citoyennes et citoyens, et d'autres mixtes avec les conseillères et conseillers.

- **Accueil**

→ *Enjeux en matière d'intégration*

L'accueil des citoyennes et citoyens recouvre plusieurs enjeux. Il s'agit tout d'abord de présenter le CESE, ses missions, le travail en commission et ses différentes étapes, les acteurs et lieux avec lesquels ils vont évoluer tout au long de la démarche. L'objectif de l'accueil est également de permettre aux citoyennes et citoyens de bien comprendre le mandat et les missions qui leur sont confiés au sein de la commission, et de se sentir à l'aise et légitimes pour participer aux travaux. Enfin, il s'agit de créer du lien entre toutes et tous et de rencontrer les conseillères et conseillers avec qui elles et ils vont travailler durant plusieurs mois.

→ *Leviers mis en place*

Trois jours consécutifs au tout début du dispositif ont été mis en place selon le programme suivant :

- **Présentation du CESE, présentation du travail en commission et des missions des citoyennes et citoyens**
- **Présentation des acteurs du dispositif**
- **Visite du Palais d'Iéna**
- **Temps conviviaux pour entrer dans le sujet et faire connaissance : exposition sur le travail et l'IA, pièce de théâtre sur l'IA, buffet avec les conseillères et conseillers**
- **Plusieurs temps d'entrée dans le sujet : présentation par les co-rapporteurs, appropriation du socle documentaire qui leur avait été mis à disposition, atelier de prospective sur l'IA avec une composante participative.**

Détail du déroulé des journées d'accueil :

Mardi 10 septembre 2024		
12h - 13h30	Accueil des citoyennes et citoyens autour d'un buffet	Salon Eiffel
13h30 - 14h (citoyens)	<p>Mots de bienvenue</p> <p>→ Mot d'introduction par Thierry BEAUDET, président du CESE</p> <p>→ Mot d'accueil par Thierry CADART, président de la Commission temporaire sur l'intelligence artificielle</p> <p>→ Mot d'accueil par Marianne TORDEUX-BITKER et Erik MEYER, corapporteurs de l'avis</p>	Salon Eiffel
14h - 14h30 (citoyens)	<p>Présentation du CESE</p> <p>→ Kenza OCCANSEY, vice-président du CESE en charge de la participation citoyenne et rapporteur de l'avis « Engagement et participation démocratique des jeunes »</p>	Salon Eiffel
14h30 - 15h (citoyens)	Interconnaissance entre les citoyennes et citoyens	Salon Eiffel

15h - 16h30 (citoyens)	<p>Présentation de la mission des citoyennes et citoyens et du travail en commission</p> <p>→ <i>Retour sur le mandat</i> ; les objectifs ; le fonctionnement du travail en commission (étapes, calendrier, etc.) ;</p> <p>→ <i>Présentation des documents d'accueil</i> (charte, livret, socle documentaire).</p> <p>→ <i>Rédaction collective de questions</i> afin d'identifier les craintes et de répondre aux premières questions</p>	Salon Eiffel
16h30 - 17h	Pause	Salon Eiffel
17h - 18h15 (citoyens + membres)	Visite de l'exposition immersive (APEC) « Travailler demain quels futurs se dessinent »	Hypostyle
18h15 - 19h30 (citoyens + membres)	Pièce de théâtre (INRIA) « <u>Qui a hacké Garoutzia ?</u> »	Hémicycle
19h30 - 19h40 (citoyens + membres)	Clôture de la journée et présentation du programme du lendemain	Hémicycle
19h45 - 21h30 (citoyens + membres)	Buffet dînatoire réunissant les membres de la commission temporaire	Salon Eiffel ou terrasse
Mercredi 11 septembre 2024		
9h - 9h30	Accueil café	Place du village

9h30 - 10h (citoyens)	Point d'étape : comment nous sentons-nous à ce stade ? Bilan des premières présentations, liste des sujets à clarifier et échanges libres.	Place du village
10h - 11h (citoyens + membres)	Présentation des acteurs présents autour de la CT : garants, évaluateurs, communication, personne-ressource → Jacques Commaille et Stéphanie Lacour, garants du dispositif → Emmanuel Rivat, Agence Phare (évaluation) → Manuella Rubio et Lucile Duclaux, Direction de la communication	Place du village
11h - 11h30	Pause	Place du village
11h30 - 12h30 (citoyens + membres)	Temps d'échanges sur le mandat de participation et le travail en commission	Place du village
12h30 - 14h	Déjeuner + photo pour les badges	Cafétéria
14h - 15h (citoyens + membres)	Première approche pédagogique du sujet de l'IA → Thierry CADART , président de la Commission temporaire sur l'intelligence artificielle → Marianne TORDEUX BITKER et Erik MEYER , corapporteurs du projet d'avis	Salle 225
15h - 17h15 (citoyens + membres)	Lecture collective du socle documentaire	Salle 225 et 245

17h15 - 17h45	Pause	Palier du 2ème étage
17h45 - 18h30 (citoyens)	Visite du Palais d'Iéna Visite guidée par un guide conférencier	Hypostyle
18h30 - 18h45	Clôture et présentation du programme du lendemain	Cafétéria
19h - 21h	Repas	Cafétéria
Jeudi 12 septembre 2024		
9h - 9h30 (citoyens + membres)	Accueil café et interconnaissance entre conseillers et citoyens	Salon Eiffel
9h30 - 12h30 (citoyens + membres)	Atelier prospective	Salon Eiffel
12h30 - 14h	Déjeuner	Cafétéria
14h - 17h (citoyens + membres)	Réunion de la Commission → Retour sur les ateliers prospectives (celui de la matinée et celui du 28 mai) en petits groupes → Échanges sur les objectifs collectifs de la commission.	Salle 67

En plus, et afin de se projeter dans l'aboutissement de leur travail au sein de la commission temporaire, les citoyennes et citoyens ont été invités à assister à la séance plénière sur l'avis "Impacts de l'intelligence artificielle : risques et opportunités pour l'environnement".

- **Auditions**

- *Missions confiées aux citoyennes et citoyens*

Les citoyennes et citoyens ont participé aux auditions de la Commission temporaire au même titre que les conseillères et conseillers. Elles et ils ont pu poser des questions aux auditionnés, proposer des noms de personnes ou structures à auditionner, réagir lors des débats et discussions à la suite des auditions, et proposer leurs propres éléments de retours, constats, et propositions lors de ces temps d'échanges.

- *Enjeux en matière d'intégration*

Durant la phase d'auditions, les enjeux sont de permettre aux citoyennes et citoyens de se sentir suffisamment à l'aise et outillés pour poser des questions pendant les auditions, et participer aux débats suivant ces auditions.

- *Leviers mis en place*

Durant la phase d'amendements, des visio-conférences facultatives d'accompagnement sur les auditions ont été mises en place avec l'administration de la commission, la personne ressource et le président, afin de permettre aux citoyennes et citoyens de revenir sur les auditions passées, et de préparer les auditions de la réunion à venir.

Les méthodes de travail classiques de commissions ont été adaptées afin de permettre une plus grande implication des citoyennes et citoyens : mise en place de travail en sous-groupes de 3 ou 4 personnes (conseillers et citoyens mélangés) après les auditions afin de formuler de retours sur les auditions et lancer le débat, mise en place d'ateliers de réflexion en sous-groupes pour

l'élaboration du plan de l'avis, tours de table à plusieurs reprises, notamment au tout début de la première lecture du texte, afin de donner *de facto* la parole à toutes et tous.

Enfin, des temps conviviaux ont été mis en place afin de créer du lien entre tous les membres conseillers et citoyens de la commission : notamment un "buffet des régions".

- **Examen du texte**

- *Missions confiées aux citoyennes et citoyens*

Durant la phase d'examen du texte, les citoyennes et citoyens avaient pour mission de lire le texte afin de présenter leurs retours et formuler des propositions d'évolutions à l'oral lors de la première lecture du texte. Lors de l'examen en seconde lecture du texte, les citoyennes et citoyens, comme les conseillères et conseillers, ont pu déposer des amendements sur le texte, puis réagir lors des réunions de commission afin d'apporter leurs retours sur les amendements déposés par toutes et tous.

Les citoyennes et citoyens ont choisi lors de cette phase d'intégrer au texte des encadrés citoyens intitulés « Recommandations de citoyennes et citoyens », et qui visaient à présenter de manière courte et simplifiée les principaux éléments de chaque partie.

- *Enjeux en matière d'intégration*

Cette phase est cruciale en matière d'accompagnement pour les citoyennes et citoyens intégrés car la dynamique du travail en commission évolue beaucoup à ce moment du dispositif. Il s'agit donc d'outiller les citoyennes et citoyens pour

comprendre comment va se dérouler cette phase, quelles sont les différences entre la première et la deuxième lecture, comment remplir le cahier d'amendements, comment défendre leurs positions, etc.

→ *Leviers mis en place*

Une demie matinée de réunion de commission a été dédiée à une formation à l'examen d'un texte (première lecture, deuxième lecture, processus d'amendements, etc). Ce temps en commission a été l'occasion de préciser le cadre de travail de cette nouvelle étape, et les conseillères et conseillers ont pu apporter des exemples concrets et leur expérience aux citoyennes et citoyens pour préparer au mieux cette nouvelle étape du travail en commission.

Afin de préparer au mieux les citoyennes et citoyens à la phase d'examen du texte, l'après-midi à la suite de la formation à l'examen d'un texte a été dédié à des exercices pratiques uniquement entre citoyennes et citoyens pour s'approprier les méthodes de travail de cette nouvelle phase. Cela a notamment permis de travailler à formuler des amendements sur un texte fictif et de travailler à construire un argumentaire pour défendre un avis ou un amendement précis.

Les méthodes de travail en commission ont également été adaptées durant cette phase, notamment afin de permettre un tour de table au tout début de l'examen en première lecture du texte, ce qui a permis à toutes et tous d'exprimer un avis sur la première version du texte.

→ *Adoption du texte*

Le Règlement intérieur du CESE prévoit que seuls les conseillères et conseillers ont le droit de voter. Un vote consultatif des citoyennes et citoyens est cependant organisé au moment du vote d'adoption du texte en Commission, pour la transmission de l'avis au Bureau. Les résultats de

ce vote sont indicatifs pour le vote des conseillères et conseillers, et sont annexés à l'avis.

→ *Portage de l'avis*

Les citoyennes et citoyens seront également associés à la phase de portage de l'avis.

• **Garantir les critères de réussite de la démarche :**

En plus du rôle des garants qui sont à disposition des citoyennes et des citoyens, une évaluation du dispositif a été mise en place afin de vérifier la plus-value apportée par ces modalités de travail à la construction d'un avis du CESE, et d'apporter au CESE un regard réflexif sur le dispositif.

Cette évaluation a été confiée au cabinet Agence Phare avec pour objectifs de vérifier la pertinence, la cohérence et l'efficacité du dispositif et de proposer des pistes d'amélioration pour d'éventuelles futures commissions mixtes.

L'évaluation du dispositif s'est déroulée selon plusieurs étapes et modalités :

- observation par le cabinet de plusieurs temps (journées d'accueil, auditions, séances de première et deuxième lectures)
- 2 entretiens collectifs avec chacun des citoyens : un au début du dispositif, juste après les journées d'accueil et un à la fin du dispositif
- 1 entretien collectif avec des conseillères et conseillers
- 1 entretien avec les garants du dispositif
- 1 questionnaire d'évaluation des journées d'accueil
- 1 questionnaire d'évolution de l'opinion envoyé à trois moments clés du dispositif : au début, à la fin de la phase d'auditions et à la fin de la phase d'examen du texte.

4

Expressions libres des citoyennes et citoyens intégrés à la Commission Intelligence artificielle

1. Frédéric M.

Dans un monde soumis à l'accélération technologique, l'intelligence artificielle est inévitablement vouée à prendre le contrôle de l'ensemble des domaines de la vie. Vouloir la réguler est une vue de l'esprit. Bien au contraire, ce qui est en jeu, c'est la souveraineté du pays et de l'Europe qui doivent rapidement mettre les conditions favorables à l'émergence de développeurs capables d'assurer pour notre compte, la défense de nos valeurs fondamentales. L'avis ne reprend finalement que peu cette vue plus libérale de la situation et comme pour tous les aspects de la vie, ne fragmente que trop l'aspect simplement éthique de l'IA demain au même titre qu'internet hier. Il ne s'agit pas d'être contre l'IA comme il est possible de le lire mais seulement contre les faussaires qui peuvent se servir de l'IA. Là est le seul sujet.

2. Hicham L.

A mon sens ce travail du CESE en collaboration entre conseillers et citoyens a été une réussite. En effet, je retrouve mon avis général dans le rendu final : la problématique des risques et menaces que peut soulever l'IA est bien présente tout en expliquant en détail ces menaces, leurs origines, ainsi qu'une notion clé, la donnée personnelle du consommateur qui contribue à créer ces IA et qui peut poser problème lorsque ces données sont dites sensibles.

J'aurai cependant à mon sens peut être mis l'accent légèrement plus sur les enjeux environnementaux mais ce qui est présent dans l'avis est déjà relativement fourni, cependant la diversification des échelles, notamment à celle locale aurait été intéressante dans ces discussions.

3. Lucas 34ans. Citoyen

L'IA étant un sujet complexe et forcé de par son émergence soudaine et sa prise rapide dans les rouages de nos sociétés, il est difficile d'émettre un avis l'encadrant totalement. C'est ce défi qu'a essayé de réaliser notre commission temporaire. Je suis globalement d'accord avec ce qui est ressorti des débats et mis en écrit. Le plus important pour moi est le regard que l'on peut avoir sur cette technologie, le paradigme dans lequel on se trouve pour en faire la promotion, il est trop tard pour reculer car tout le monde s'engouffre dans cette brèche juteuse. La question la plus importante aujourd'hui est : comment doit-on faire de l'IA ? C'est une technologie qui a certes des avantages mais aussi bons nombres de défauts non négligeables que ce soit sur le plan humain, éthique et surtout écologique. Si je devais directement parler à nos décideurs sur le sujet, ma phrase serait courte : ne nous tirons pas une balle dans le pied. Les crises climatiques et humaines sont déjà bien avancées. Pour cela les cas d'usages sont très importants. Une IA écoresponsable utilisée comme un outil avec parcimonie et intelligence, promu dans les secteurs où le besoin est réel, oui.

Une IA utilisée à tout va à des fins de compétitivité économique, sans réflexion sur son usage et son impact, non.

4. Maud, 29 ans, agricultrice, Picardie

Dans l'ensemble je suis plutôt contente des préconisations émises par le CESE, elles retracent bien nos points énumérés lors de nos réunions.

Néanmoins, ça ne change pas mon avis sur le sujet en lui-même, qui était négatif avant la commission et l'est encore plus à la suite de celle-ci.

En effet, il y a des aspects positifs à cette nouvelle technologie comme l'avancée des recherches médicales qui concernent un bien communautaire international et aussi une aide florissante dans mon métier d'agricultrice avec de nouvelles idées et avancées par rapport à la gestion des intrants et du parcellaire mais qui a ses limites d'appropriation par rapport au prix et risque de créer un fossé entre grands et petits exploitants.

Aussi, pour les autres secteurs je ne vois certes de mon point de vue que des aspects dégradant l'être humain et le laissant en arrière-plan, tout en polluant deux fois plus notre planète déjà bien atteinte par notre propre cancer.

Je crains aussi que le développement de l'IA ne se fasse que pour des raisons commerciales.

Ma question est :

Une fois que l'être humain aura parcouru de fond en comble les ultras limites de l'IA, quel sera le prochain engouement ?

5. Saïd K.

Tout d'abord je tiens à dire que j'étais très heureux de participer à la commission sur l'Intelligence Artificielle, c'est un vaste sujet que nous devons sans cesse approfondir.

Les pouvoirs publics et la société en général doivent se l'approprier, car il y a trop d'enjeux. Ce texte bien entendu n'est pas le fruit de de l'IA. mais d'un citoyen.

La chose primordiale concernant cette thématique est d'informer et de former les citoyens et élus sur ce qu'est réellement l'IA.. La formation et l'éducation sont la base de tout, comment éviter les « biais » si nous n'avons pas conscience de la matière.

Cela permettra de lever tous les préjugés sur ce qui reste une avancée technologique de premier ordre. Il faut prendre en considération que cette technologie nous est étrangère.

Son implication touche tous les domaines : à elle seule, la Commission temporaire ne peut répondre à toutes les questions. Elle nous donne des pistes de réflexions, et bien entendu tout un chacun suivra une voie selon son ou ses obédiences, sa trajectoire de vie, ce qui fera notre force c'est de respecter toutes les idées. Je salue d'ailleurs le travail des rapporteurs et du Président qui n'a pas été un long fleuve tranquille, et n'oublions pas non plus tous les collaborateurs.

Me concernant, je dois dire que je fus novice en la matière, mais que je suis arrivé grâce à la commission à cerner l'IA.. Elle ne cessera d'évoluer et de nous interroger.

Un mot que je pense essentiel fut répété maintes fois lors des débats, REGULER. Nous n'avons pas eu le temps nécessaire de développer sur ce point, mais la régulation nous permettra de maîtriser l'IA. dans l'intérêt de la société et des plus démunis face à la technologie. L'intelligence artificielle est là, épousant tous les domaines, elle peut être une révolution saine mais sans contrôle ce sont notre mode de vie et notre liberté qui sont en jeu. Nous ne pouvons également nous mettre en marge de l'évolution du monde. En effet l'industrie, l'agriculture et

les services font appel depuis des années à l'IA, qu'il s'agisse de ce qui vont devenir les emplois et notre système social : des enquêtes démontrent que des emplois vont être sacrifiés.

A nous, de relever le défi et de rendre ce monde plus juste, sans se priver de ce qui pourrait être une aubaine. Soyons acteurs de notre avenir, l'IA peut également aider à mieux soigner, éduquer, rationaliser... La seule chose importante c'est de garder la maîtrise de cet outil.

6. Yannick S.

Tout d'abord, je tiens à exprimer ma fierté et mon honneur d'avoir pu participer, en tant que citoyen, à cette commission. Cette expérience m'a ouvert les yeux sur la richesse et la complexité des travaux menés par le CESE, mais aussi sur l'importance de cette institution pour éclairer les décideurs publics sur des sujets aussi fondamentaux que l'IA et les nouvelles technologies. J'ai découvert à quel point ces enjeux exigent une réflexion collective et un dialogue entre experts, citoyens et institutions.

L'expression « intelligence artificielle » est usurpée. Issue de l'anglicisme "artificial intelligence", elle ne traduit pas fidèlement ce qu'elle représente : une technologie qui repose avant tout sur l'organisation et le traitement de l'information, et non sur une véritable intelligence au sens humain du terme. Dans ce domaine, la seule intelligence réelle est celle des humains : celle des créateurs, des utilisateurs et des décideurs publics. Ce sont eux qui déterminent comment l'IA est conçue, réglementée et employée.

Cette distinction est essentielle car elle met en lumière notre responsabilité collective. L'IA n'est pas une entité autonome, mais un outil qui reflète nos priorités et nos choix. Pour qu'elle serve réellement l'intérêt général, nous devons utiliser notre intelligence humaine pour encadrer son usage. Cela passe par plusieurs axes prioritaires :

• 1. Une régulation éthique et inclusive

Les biais algorithmiques présents dans les systèmes d'IA sont le reflet des discriminations sociétales existantes. Ces biais peuvent être amplifiés si nous n'y prêtons pas attention. Une régulation proactive est donc nécessaire pour garantir que l'IA ne reproduise pas ces injustices mais contribue à une société plus équitable.

• 2. Une responsabilité environnementale

L'impact écologique de l'IA est considérable : consommation énergétique des centres de données, fabrication de matériels et multiplication des usages superflus. Par exemple, une recherche effectuée par un système d'IA peut consommer plusieurs fois l'énergie nécessaire pour une recherche internet classique. Il est crucial d'optimiser l'implantation géographique des centres de données et de sensibiliser les utilisateurs aux coûts cachés de ces technologies.

• 3. Des usages ciblés et bénéfiques

L'IA, lorsqu'elle est bien utilisée, peut transformer des vies. Pour les personnes en situation de handicap, elle ouvre des perspectives

extraordinaires : outils de reconnaissance vocale pour les troubles de la motricité, applications d'assistance visuelle pour les malvoyants, traduction automatique en langue des signes, ou encore prothèses intelligentes. Dans le domaine médical, elle permet des avancées majeures, notamment en imagerie. Les systèmes d'IA peuvent détecter précocement des cancers ou des anomalies cardiaques, offrant ainsi aux professionnels de santé des outils puissants pour sauver des vies.

• **4. Une technologie au service de la société**

Pour que l'IA soit un levier de progrès, elle ne doit ni diviser les populations ni exacerber les inégalités. Elle doit être un outil transparent, compréhensible et contrôlable par tous. L'éducation et la formation des citoyens sur les enjeux et limites de l'IA sont déterminantes pour favoriser une adhésion collective et un usage responsable.

Ce que l'IA nous enseigne avant tout, c'est l'importance de notre propre intelligence. Elle nous force à nous interroger sur nos valeurs et notre vision du futur. En faisant preuve d'ambition et de prudence, nous avons l'opportunité de façonner un avenir où la technologie reste au service de l'humanité. Cet avis collectif est un pas dans cette direction, et j'invite chacun à s'en saisir.

7. Myriam F.

Rencontrer l'IA dans différents domaines de ma vie quotidienne était chose admise et acquise.

Venir au CESE approfondir mes connaissances a été une expérience intense et enrichissante.

L'IA affole ou enchante les êtres humains. Certains y voient une avancée technologique incroyable au service de l'homme, d'autres voient en elle une manne financière source de richesse et

d'avantages personnels.

L'IA doit rester un outil dont les usages nécessitent une grande prudence. Il est du devoir des instances gouvernantes d'en démocratiser la connaissance et son utilisation. Mais avant tout d'en définir son utilité. Elle ne doit pas remplacer l'humain, elle doit le servir.

L'avis du CESE démontre les multiples avancées et dérives possibles. Le monde a besoin de l'IA pour progresser dans bien des domaines. Restent le savoir et la sagesse qui seuls permettront à tous de conserver une attitude prudente et une bonne adaptation à cette révolution technologique, en limitant ainsi les dangers.

8. Martine, Indre-et-Loire

Je trouve l'avis bien fait et bien détaillé.

Il couvre notamment bien tous les secteurs, et je trouve positif que les deux parties entre enjeux et impacts soient bien distinctes.

Le rapport présente pour moi quelques points faibles. Concernant les suites tout d'abord : on ne sait pas ce qu'il va donner (accord par les pouvoirs publics, présentation Sommet mondial de l'IA). La partie sur la formation ne me semble pas suffisamment explicite. D'autres aspects manquent également : plus de détails sur qui peut s'emparer de cet outil, et quels sont les moyens financiers nécessaires, plus parler des pouvoirs qu'a Elon Musk avec l'IA.

Les préconisations prioritaires sont pour moi :

• **Renforcer la réglementation européenne (préconisation 1) :**

Aujourd'hui, la réglementation spécifique à l'IA en Europe n'est pas suffisante, alors que c'est un domaine en constante évolution. L'intelligence artificielle arrive dans tous les secteurs. Il faut réglementer l'IA et ses usages pour garantir le respect

de nos droits fondamentaux. L'IA offre des opportunités significatives, mais ses impacts négatifs sur les libertés, l'équité et la démocratie nécessitent des mesures de contrôle strictes. Il faut une régulation proactive pour garantir un usage responsable et respectueux des droits de tous les citoyens. L'IA consomme trop de ressources (data center) pour son fonctionnement. Il faut donc plus de transparence envers les utilisateurs. Il faut planifier une utilisation responsable de l'énergie et favoriser les infrastructures durables.

- **Renforcer le financement de la recherche, mobiliser les investisseurs privés en faisant attention aux risques (préconisation 9) :**

La France et l'Europe doivent avoir un rôle crucial dans la course à l'IA face aux USA et la Chine. Mais nous sommes freinés par nos infrastructures et le sous-investissement dans la recherche. Il faut investir davantage dans les technologies propres et durables.

- **Harmoniser le droit des affaires dans l'UE (préconisation 10)**

- **Retenir et attirer en France tous les talents (préconisation 12) :**

Il faut faire évoluer les mentalités européennes pour les femmes reviennent dans les métiers du numérique pour qu'elles puissent « encoder l'égalité », et promouvoir la féminisation dans le secteur de l'IA, en mettant en place des initiatives spécifiques pour attirer et soutenir les femmes dans ce domaine.

- **Faire une base pour le dialogue social sur l'IA (préconisation 13)**

- **Soutenir la formation au numérique et à l'IA de la population à tout âge (préconisation 24)**

Il faut aussi mettre en place des formations pour préparer les travailleurs aux nouveaux métiers et intégrer l'IA. Comme l'IA reste compliquée pour ceux qui ne maîtrisent pas le numérique et aggrave les inégalités, il faut renforcer les formations pour tous les âges et donner les moyens à chacun.

- **Développer la formation des élèves à ce qu'est l'IA (préconisation 27)**

Je pense qu'il faut déployer les systèmes d'IA à condition qu'ils améliorent la qualité du service public et les conditions de travail des agents. Ils doivent être plus efficaces et accessibles mais sans remplacer l'humain. Il faut prendre en compte les besoins de tous.

Dans la santé, l'IA permet des avancées rapides et améliore la qualité des soins. Mais il faut maintenir une relation humaine dans les soins et on insiste surtout sur la protection des données. Plusieurs cas de piratage d'hôpitaux ont permis d'avoir les données des patients.

D'autre part, l'IA peut améliorer la vie des personnes handicapées et il faut donc renforcer leur accessibilité pour que tout le monde en profite.

La France pourrait promouvoir la création d'un Fonds international pour l'IA d'intérêt public. Ce fonds contribuerait à faire émerger un ensemble de projets bénéfiques pour l'humanité : des services d'IA libres et gratuits (open source), des projets de recherche indépendante,

des innovations (dans l'environnement, la science, la santé...)

- **Mettre l'IA au service du bien commun**
- **Favoriser le dynamisme de l'innovation et l'ouverture de l'écosystème**
- **Préparer l'avenir du travail avec l'IA**
- **Promouvoir des systèmes d'IA sûrs et sécurisés**

Pour ma part, j'étais un peu contre l'IA au début car je n'avais pas assez de connaissances. Par la suite, avec les entretiens, je suis devenue pour, mais à certaines conditions. Le plus important pour moi ce sont nos données personnelles : que vont-elles devenir avec tous ces piratages ? De plus, l'Europe est peut-être déjà bien retard vis-à-vis des 2 grands : on peut se poser la question de si les préconisations du CESE ne seront pas obsolètes dans 1 an ou plus.

9. Jonathan, 41 ans, Ile-de-France, Gestionnaire d'actifs financiers pour compte propre

Je trouve bien que cet avis soit centré sur l'intérêt général, avec la nécessité d'une autonomie stratégique européenne. Deux préconisations sont pour moi prioritaires :

- Celle concernant la gouvernance démocratique et le contrôle éthique de l'IA : cela garantit que le déploiement de l'IA soit aligné avec les valeurs sociétales, tout en évitant une concentration de pouvoir technologique entre les mains de quelques acteurs.
- Celle concernant la régulation de l'impact environnemental : l'urgence climatique impose de prioriser les solutions technologiques qui respectent les limites planétaires.

Cependant, il manque d'après moi certains aspects dans l'avis. En effet, l'avis se concentre principalement sur l'Europe

et l'Occident, je trouve qu'il manque une analyse de l'impact de l'IA sur les pays en développement, que ce soit en termes d'opportunités ou de risques. Je trouve également que l'avis ne parle pas assez d'argent (bénéfices pour la population, bénéfices pour les entreprises, bénéfices pour les investisseurs, etc.).

Aussi, je ne suis pas d'accord avec l'approche un peu trop alarmiste dans les passages concernant les biais et les impacts négatifs de l'IA. Je pense que les préoccupations sont valables mais cela peut donner l'impression d'un manque de neutralité. Je regrette que le texte manque parfois de détails (notamment lorsqu'il mentionne des acteurs comme Google, Amazon, Microsoft, sans évoquer l'impact économique de l'IA sur les PME ou les pays en développement), qu'il soit parfois trop répétitif sur certains thèmes (comme les biais de l'IA ou les impacts environnementaux) et parfois difficile d'accès à un public non expert (du fait d'un langage technique ou académique).

Pour moi, de futurs travaux sur le sujet de l'IA devraient interroger jusqu'où autoriser les systèmes autonomes (véhicules, drones, robots chirurgicaux) à prendre des décisions sans intervention humaine directe.

10. Habiba, Marseille

Pour ma part l'avis est conforme aux auditions présentées tout au long des séances : j'ai voté pour.

Les préconisations prioritaires sont pour moi :

- **Celle sur la CNIL (préconisation 3), parce qu'il faut renforcer la protection de nos données**
- **Favoriser la diversité des données pour réduire les biais (préconisation 2), tout simplement parce que c'est de là que partent les fausses données, et qu'il faut de la diversité.**

Je crois que dans les travaux futurs, il faudrait approfondir la transparence, et l'information auprès des citoyens. En effet, pour ma part cela concerne les data centers, je savais qu'il y en avait un à Marseille (d'où je viens) et cela me semblait positif, mais au fil des auditions, j'apprends qu'il faut beaucoup de ressources pour les faire fonctionner (eau, électricité), et que cela rejette de la chaleur. Bien sûr pour l'environnement, c'est nocif. Je pense qu'il serait bien de tenir informé chaque habitant des départements, et régions concernés avec un courrier nominatif expliquant l'implantation d'un data center et peut-être même leur permettre de voter sur cette implantation.

J'avoue qu'au début de la commission, j'étais sceptique, pour moi l'IA c'était un sujet trop complexe. Au fil des auditions mon avis a évolué, surtout concernant la santé. Pour moi il faut en faire une priorité et donner les moyens pour développer l'IA pour la santé et pour les domaines où l'IA a un impact positif : l'IA est une bonne chose dans ces domaines.

Cependant il faudrait que les politiques s'investissent : il faut qu'ils régulent lorsque l'IA atteint l'humain et nos protections ou quand elle est dangereuse (le cas des biais par exemple).

Il faut aussi approfondir l'enseignement au plus tôt pour les enfants et mettre à disposition des « professeurs d'IA » dès le collège (1h par semaine) pour que les élèves sachent ce qu'est l'IA, puissent identifier les fake news et qu'ils ne soient pas perdus. Cela permettrait aussi que ceux qui n'ont pas d'ordinateur à la maison découvrent cet outil : l'école doit donner les mêmes chances à tout le monde.

5

Analyse de controverses : Intelligence artificielle, travail et emploi – Synthèses des tableaux d'analyse¹⁹¹

A. Synthèse du tableau des arguments de la question 1 : l'intelligence artificielle va-t-elle augmenter les inégalités au travail ?

Les inégalités constituent un ensemble plus vaste que la question plus restreinte des discriminations telles que définies dans le Code du travail. L'analyse ne s'attachera pas au caractère légal ou illégal des inégalités, mais aux enjeux d'équité dans le déploiement des systèmes d'IA.

1. Éléments introductifs

Le développement et l'utilisation de l'IA doivent impérativement être encadrés, car ils soulèvent des préoccupations éthiques. La lutte contre les risques associés à l'IA, en particulier ceux relatifs au renforcement des inégalités et des discriminations, nécessite en effet des choix politiques, juridiques et de société associant l'ensemble des acteurs (pouvoirs publics, développeurs, chercheurs, entreprises, partenaires sociaux...) afin de créer des systèmes d'IA plus équitables et impartiaux au service de la société et d'en assurer un contrôle régulier. L'absence de transparence du fonctionnement des systèmes d'IA pose une question fondamentale : celle d'un usage malveillant des algorithmes, à l'insu des travailleurs et travailleuses pour évaluer leur comportement, leurs performances professionnelles ou leur état émotionnel ou de santé (pour désavantager, par exemple, les candidats anxieux, avec un handicap ou en perte d'autonomie). À l'inverse, de tels outils

pourraient être utilisés par les managers pour détecter des salariés en difficulté.

Les principales controverses sont présentées par thématique.

Les biais algorithmiques reflètent, perpétuent, voire amplifient les inégalités sociétales et les stéréotypes de genre qui se reproduisent dans tout le cycle de vie des systèmes algorithmiques, soit de leur conception à leur utilisation. Les algorithmes et autres formes d'intelligence artificielle peuvent avoir des effets discriminatoires. L'IA peut en outre procéder à des différenciations injustes (mobilisation de critères qui semblent être neutres en apparence) de groupes sociaux faisant déjà l'objet de discriminations (femmes, personnes en situation de handicap, seniors...) renforçant ainsi les inégalités. Malgré l'effet « boîtes noires » des systèmes algorithmiques, des solutions techniques existent pour prévenir et corriger ces biais algorithmiques (agir sur les critères de sélection, éliminer les biais de jugement humains, redresser les données d'apprentissage, études d'impact...) en supprimant par exemple les formulations stéréotypées de genre dans les offres d'emploi ou en diversifiant les équipes de développeurs dans le milieu du numérique.

¹⁹¹ Synthèse de l'étude rapportée au nom de la Commission Travail et emploi par Jean-Marie Truffat.

L'IA conduit à reproduire ou à renforcer les représentations genrées dans le travail, notamment en matière de recrutement. Par exemple, si les candidates sont sous-représentées dans les données d'apprentissage de l'outil algorithmique, la proportion de femmes recrutées sera faible. Mais l'IA est aussi un atout pour l'égalité de genre, car valoriser la diversité peut constituer une opportunité pour les entreprises.

Les outils fondés sur l'IA ne contribuent pas toujours à réduire les discriminations à l'embauche, et peuvent même engendrer de nouveaux biais de jugement. Ces outils algorithmiques d'aide au recrutement tendent par exemple, à reproduire les caractéristiques des personnes déjà employées limitant ainsi la diversité et aggravant les discriminations liées à l'âge ou aux origines sociales, ethniques et géographique (etc.). Pourtant, l'IA permet aussi un recrutement plus inclusif et peut être utilisé comme un moyen de renforcement de l'égalité de traitement.

L'IA conduit à de plus grandes inégalités de revenus et de compétences. Pour certains, l'IA favorise les salariés hautement qualifiés, alors que, pour d'autres, l'IA accompagne également les travailleurs et travailleuses débutants ou peu qualifiés, susceptible d'entraîner un rattrapage de revenus et de compétences.

L'utilisation et l'appropriation de l'IA varient selon la taille, l'appétence, les moyens de l'entreprise et les secteurs d'activité. Elle offre un avantage significatif pouvant accroître les écarts entre les entreprises. Il est à noter toutefois que des solutions permettent de

rendre l'IA plus accessible aux petites entreprises.

L'IA est aussi une opportunité pour les services publics de l'emploi, car elle vise à libérer du temps administratif afin de se consacrer à l'accompagnement personnalisé des demandeurs d'emploi. Pour d'autres, l'IA entraîne des inégalités et discriminations dans l'accompagnement des demandeurs d'emploi. Le profilage algorithmique de ces personnes à des fins de contrôle est susceptible d'entraîner des classifications stigmatisantes des personnes particulièrement en situation de précarité. Ces approches risquent de biaiser l'évaluation des demandeurs d'emploi en fonction de critères discriminants.

2. Conclusion

Plusieurs points de vigilance ou angles morts doivent être signalés, pouvant être porteurs de risques ou d'opportunités sur des sujets que l'analyse n'a pas été en mesure d'aborder, notamment en l'absence de sources précises ou suffisamment fiables :

- modifier les perspectives de travail des plus vulnérables, notamment des personnes moins qualifiées ou qui ont plus de difficulté à intégrer de manière stable le marché du travail ;
- faciliter l'inclusion des groupes vulnérables sur le marché du travail ;
- identifier des salariés considérés comme les plus aptes via les critères retenus dans le paramétrage des algorithmes de recrutement ;
- améliorer le couplage entre l'offre et la demande de travail basé sur

de multiples critères (tels que le métier, les compétences, la géolocalisation, le type de contrat, les horaires de travail...);

- favoriser le management bienveillant et apporter une solution aux difficultés des collaborateurs ;
- cibler prioritairement certaines populations de demandeurs d'emploi

(les populations les plus vulnérables comme les personnes isolées ou à faible revenu) dans les politiques de lutte contre la fraude aux versements des allocations chômage versées par France Travail ;

- avoir un impact sur les allocataires des minimas sociaux.

B. Synthèse du tableau des arguments de la question 2 : l'intelligence artificielle est-elle un progrès pour l'organisation des conditions de travail et pour la santé au travail ?

1. Le sens de la question

Cette question de controverse renvoie aux effets des usages des intelligences artificielles sur les conditions de travail et la santé au travail. Plus fondamentalement, elle met en jeu l'articulation entre le développement d'une technologie nouvelle - dans laquelle certains voient une véritable rupture - et le travail lui-même dans la réalité de son organisation et de son incidence sur les conditions et la santé au travail. Dans le travail collectif de la controverse, les arguments ont été recherchés en « pour » et en « contre » c'est-à-dire en soutien et en opposition au présumé de la question : l'IA conçue comme un « progrès ».

La question est formulée au présent, car, loin d'être seulement un sujet de prospective, l'intelligence artificielle est déjà très présente dans le monde du travail. Si l'IA générative des grands modèles de langage apparaît effectivement d'une assez grande nouveauté, des systèmes experts ont été développés depuis plus longtemps dans des univers industriels, à l'appui de l'automatisation.

2. Quatre grands enjeux

L'enchaînement des arguments « pour » et « contre » retenu dans le tableau conduit à identifier quatre thématiques principales :

- 1) intérêt, qualité du travail et autonomie,
- 2) santé au travail et charge mentale
- 3) formation et adaptation/reconversion,
- 4) régulation de l'IA en rapport avec le travail : norme légale et dialogue social.

2.1 LA CAPACITÉ DE L'IA À AMÉLIORER LA QUALITÉ DU TRAVAIL, SON INTÉRÊT POUR CELLES ET CEUX QUI L'UTILISENT, LEUR DEGRÉ D'AUTONOMIE

Techniquement, l'IA peut libérer de tâches fastidieuses pour concentrer les interventions humaines sur des activités de plus grand intérêt et à plus fort contenu relationnel. À l'inverse, dans des organisations néotayloriennes, l'IA n'élimine pas la répétition et ne combat pas l'ennui, mais réduit drastiquement, dans une recherche de plus grande efficacité, la capacité d'initiative de l'opérateur humain et intensifie le travail en supprimant les temps morts. Les grandes entreprises du numérique ont elles-mêmes développé des formes d'emplois

dégradés et parcellisés, souvent délocalisés, pour nourrir et entraîner les algorithmes.

• IA et santé au travail

Dans les entreprises de type manufacturier, l'hyper rationalisation du travail par un système d'IA peut conduire à une réduction des risques physiques en limitant par exemple les déplacements et le port de charges lourdes au strict nécessaire. En outre, les outils numériques les plus évolués peuvent être mobilisés, par des services de santé au travail ou des organismes de prévoyance, très efficacement pour prévenir les risques professionnels.

Par ailleurs, pour éviter tout détournement de finalité, un risque sur la protection et la conservation des données de santé au travail doit être traité.

En revanche, les systèmes d'IA peuvent contribuer à accroître les risques pour la santé mentale de celles et ceux qui travaillent dans des situations très différentes, voire opposées : lorsque les marges d'autonomie des travailleurs sont très fortement réduites (l'opérateur humain soumis aux prescriptions de la machine, perd le contrôle de son activité) ; lorsque, au contraire, le travail des humains est exclusivement constitué de tâches complexes et/ou très engageantes assorties de fortes attentes en matière de résultats ; enfin, lorsque celui ou celle qui travaille ne comprend pas ou mal le fonctionnement du SIA qu'il doit utiliser ou encore lorsqu'il doit pallier son dysfonctionnement (effet boîte noire de l'intelligence artificielle). L'ensemble des atteintes à la santé mentale, dont l'IA pourrait être à l'origine, en milieu professionnel, mériterait à ce stade d'être mieux

documenté : des travaux de recherches seraient à susciter dans ce domaine.

• Formation, adaptation et reconversion

La formation interne des personnels est déterminante pour une mise en place réussie des dispositifs d'IA en milieu de travail. L'acquisition de connaissances sur ces systèmes en situation d'usage peut seul permettre l'appropriation de l'outil et le développement de la confiance quant à son utilisation. Or, les dispositifs conçus pour les reclassements et l'accompagnement à la transformation profonde des professions (avec la possibilité de recourir à des formations longues pouvant aller jusqu'à la reconversion) sont à ce jour insuffisamment développés dans notre pays. Par ailleurs, l'implantation rapide des outils induit une acquisition de compétences qui ne sont encore pas ou mal reconnues.

Il existe aussi un véritable paradoxe de la formation lorsqu'elle est envisagée pour répondre au développement des IA génératives. Car si cette transformation technique induit des besoins importants en matière de formation des personnels, il est aussi attendu des salariés qu'ils participent à l'adaptation-formation de ces intelligences artificielles qui doivent être entraînées et encadrées par les opérateurs humains lors de leur mise en place. Or, les compétences mobilisées pour former l'IA seraient aujourd'hui peu visibles et très faiblement reconnues.

- **Régulation de l'IA en relation avec le travail**

Sur le terrain de la norme législative, l'Union européenne (UE) a accompli des efforts importants de régulation des usages du numérique et sans équivalent (RGPD en 2016, directive sur le travail de plateformes de 2024, règlement sur l'intelligence artificielle en octobre 2024). Il s'agit de protéger les citoyens de l'UE et leurs droits fondamentaux, dans tous les domaines d'activité, contre les dérives de systèmes potentiellement très intrusifs. Toutefois, l'échelle de temps nécessaire à l'adoption et la mise en œuvre de ces normes n'est plus celle du développement ou de la diffusion des technologies numériques, beaucoup plus rapide, et le palier franchi par les IA génératives annonce sans doute une nouvelle accélération, y compris dans la transformation de la relation contractuelle du travail. Dans le monde du travail, une régulation efficace et fondée sur la confiance dans le développement de ces techniques ne peut être envisagée sans la participation active des principaux acteurs concernés.

Les collectifs de travail ont un rôle majeur à jouer dans la régulation de l'IA. Un dialogue social actif sur ce sujet est d'autant plus nécessaire que l'IA constitue un enjeu managérial de première

importance : si l'usage de l'IA dans ce domaine est parfois présenté comme une garantie d'objectivité et d'impartialité, l'effacement de la compétence humaine et décisionnaire est aussi un risque dénoncé par des managers de proximité qui craignent d'être sommés de suivre des propositions algorithmiques, pouvant être conçues avec des paramètres biaisés.

Le dialogue et la négociation, notamment d'entreprises en lien avec l'introduction des intelligences artificielles connaissent des débuts modestes, même si le nombre d'accords d'entreprises sur le sujet est en sensible augmentation depuis environ cinq ans. Beaucoup reste à faire, notamment pour outiller les acteurs et leur permettre d'accéder à une information fiable sur le contenu des algorithmes et leur permettre de s'en approprier les enjeux.

3. Conclusion

Les arguments et leurs sources mettent en lumière le caractère ambivalent de l'IA. Son déploiement présente des risques et des opportunités pour l'ensemble des activités humaines dans les champs économiques et sociaux. Pour conjurer les risques, tel que la miscibilité entre les espaces professionnel et privé, il existe un espace, le dialogue social, pour ne pas subir, mais débattre, négocier et décider.

C. Synthèse du tableau des arguments de la question 3 : l'intelligence artificielle aura-t-elle des impacts positifs sur l'emploi ?

1. Réflexion sur la question et la méthode

La question posée était initialement la suivante : « L'IA sera-t-elle créatrice d'emplois de qualité ? ». Toute latitude avait été laissée dans les modalités d'appropriation de cette question, y

compris, si nécessaire, en procédant à une reformulation.

Partant de ce principe, il a été considéré que la notion de « qualité » devait être abordée comme un sous-thème d'une question générique sur la « création d'emplois ». Les dimensions temporelles de cette création - court, moyen, long

terme - ont ensuite été débattues, mais il a été considéré qu'elles étaient nécessairement sous-entendues. La formulation « l'IA aura-t-elle des impacts positifs sur l'emploi ? » a finalement été retenue.

Le groupe s'est efforcé de travailler conformément à la méthode annoncée : recueil des arguments « pour » et « contre » et recherche des sources, mise en correspondance des arguments et des contre-arguments, regroupement des arguments/contre-arguments par thématique. Les six thématiques ci-dessous ont été retenues.

2. L'IA va-t-elle créer ou détruire de l'emploi ?

Le déploiement de l'intelligence artificielle va générer des activités nouvelles qui seront créatrices d'emplois, elle pourra dans certains cas compenser des évolutions destructrices d'emplois ou accélérer d'autres évolutions porteuses d'emplois nouveaux.

Cependant, il est incontestable que beaucoup seront perdus par effet de substitution. Leur nombre ne peut être anticipé avec exactitude, pas plus que l'impact sur l'emploi de l'amélioration de nombreuses tâches. Il est probable que le contenu de nombre d'emplois devra être repensé. La rapidité avec laquelle l'IA sera intégrée aux processus de production est aussi une source d'incertitude dans les estimations.

3. L'IA va-t-elle améliorer ou détériorer la qualité de l'emploi ?

Le déploiement de l'IA permet la prise en charge automatisée d'un grand nombre de tâches répétitives. Les personnes dont l'emploi est ainsi transformé peuvent se consacrer à des tâches plus valorisantes. Le temps libéré offre la possibilité d'effectuer un travail plus épanouissant. Les travailleurs sont en mesure de

développer toutes leurs capacités pour s'intégrer dans une économie fondée sur l'IA.

Cependant, le niveau de compétences, le caractère intellectuel du métier exercé, ne suffisent déjà plus à protéger l'emploi. L'esprit de synthèse, la capacité d'analyse en vue de la décision, qualités exigées pour certains métiers et qui en faisaient l'intérêt, n'est pas hors de portée de l'IA. Le temps libéré par l'IA devra être consacré à des efforts accrus de productivité : il existe un risque d'accroissement de l'intensité du travail.

4. L'IA produira-t-elle un partage de la valeur favorable à l'emploi et aux salaires ?

L'IA va générer des gains de productivité, elle sera source de valeur ajoutée. Les travailleurs qui se verront concernés par de nouvelles tâches devraient pouvoir bénéficier de hausses de salaire. Les gains de productivité seront pour partie affectés à des emplois de qualité.

Il n'est cependant pas acquis que ces gains profiteront à l'emploi et aux salaires. La situation ne sera pas favorable, en cas de négociation, aux salariés. La réduction du nombre de tâches disponibles peut avoir un effet négatif sur les salaires.

5. L'IA est-elle bien acceptée par les personnes en emploi ?

Les personnes qui utilisent en lien avec leur emploi des systèmes d'intelligence artificielle sont majoritairement satisfaites. Le degré d'acceptation varie en fonction de l'emploi, du degré d'utilisation de l'IA, de ses effets sur l'intensification du travail et de la nature des interactions avec le système.

Dans les mêmes proportions, l'IA est perçue par les travailleurs interrogés comme une menace pour l'emploi à une échéance de dix ans environ - par substitution à l'emploi existant. Ce rejet

de l'IA, s'il persistait, pourrait freiner son intégration dans les processus de production, contrariant ainsi les perspectives de développement de l'emploi.

6. Les besoins en formation nécessaires au développement des emplois induits par l'IA sont-ils satisfaits ?

L'IA nécessite un effort de formation important, en premier lieu au profit des travailleurs qui l'utilisent déjà, et cet effort est bien engagé par un grand nombre d'entreprises. L'IA a des effets positifs tout d'abord parce qu'elle suscite une montée en compétences qui concerne tous les niveaux de qualification, ensuite parce qu'elle permet le développement de nouveaux outils et qu'elle facilite les actions de formation.

Cependant, le manque de compétences demeure un obstacle majeur à l'adoption de l'IA par les entreprises. Ce manque de compétences est une contrainte importante à tous les niveaux de responsabilité. L'adéquation entre les emplois, les compétences requises et les formations nécessaires pour répondre

aux besoins n'est pas assurée.

7. Est-ce que toutes les entreprises ont la capacité d'investir dans l'IA, avec les conséquences que cela implique concernant l'emploi ?

Toutes les entreprises sont prêtes à investir dans l'IA, car elle est perçue comme une « technologie générique » au même titre qu'internet et qu'il n'apparaît pas possible de ne pas s'en saisir. Les investissements privés dans l'IA continuent de se multiplier.

Une majorité d'employeurs interrogés par l'OCDE pointe néanmoins le coût potentiellement élevé de la technologie selon les solutions envisagées. Pour cette raison, ou à cause d'autres priorités, toutes les entreprises n'ont pas la capacité d'investir dans l'IA. Pour certaines d'entre elles, principalement des PME ou TPE sous-traitantes, le risque de dépendance à l'égard de leurs donneuses d'ordres va s'accroître. L'économie numérique étant beaucoup plus faible en Europe qu'aux États-Unis, il y a un risque de décrochage de la France et de l'UE en ce domaine.

6

Contribution de la délégation aux droits des femmes et à l'égalité*

* Contribution rapportée au nom de la Délégation aux droits des femmes et à l'égalité par Fabienne Tatot)

L'intelligence artificielle : dangers et avancées pour les droits des femmes et l'égalité

Points clés de la contribution

Constats :

- Les femmes sont aujourd'hui largement minoritaires dans le secteur du numérique, ce qui implique que les hommes sont les principaux développeurs d'algorithmes.
- De façon erronée, les algorithmes sont perçus comme neutres et objectifs.
- Néanmoins, par la quasi-absence de femmes dans la création des algorithmes et l'utilisation de bases de données non représentatives, l'IA reproduit, amplifie et rend possibles les biais notamment de genre.
- Ces biais peuvent porter préjudice aux femmes dans leur carrière et constituent aussi un risque pour leur santé. Pourtant, une intelligence artificielle égalitaire pourrait être un atout pour la carrière des femmes et une opportunité pour mieux les soigner.
- Le développement des intelligences artificielles dans le domaine de l'intime tend à ancrer les femmes dans des rôles traditionnels et à les sexualiser. Cette sexualisation passe notamment par l'utilisation de robots sexuels qui constituent le « continuum de la domination masculine en contexte sexuel »¹.
- En l'état, l'intelligence artificielle apprend à partir de situations existantes, ce qui fait que dans une société inégalitaire les biais de perception sont reproduits. La création de nouveaux champs relationnels et de désirs ne seront jamais proposés par une IA.
- L'IA n'est pas seulement le reflet de nos biais. Le miroir qu'elle nous tend est à double sens, il nous transforme en retour².

¹ Pauline Noiseau, « Les enjeux éthiques de la robotique sexuelle : une perspective critique féministe » (2020), <https://doi.org/10.13140/RG.2.2.15844.71042>, p. 2.

² Propos tenus par Vanessa Nurock lors de son entretien avec la Délégation aux droits des femmes et à l'égalité du CESE, le 10 octobre 2024.

Propositions :

Les conditions à la conception d'une intelligence artificielle plus égalitaire

- L'orientation des filles vers les filières scientifiques permettra une meilleure intégration des femmes dans les processus de développement des IA.
- La mise en place d'équipes de conception pluridisciplinaires et la sensibilisation des concepteurs et développeurs aux risques de biais.
- La diversification des bases de données pour qu'elles soient plus représentatives et ne reproduisent pas les inégalités existantes.

L'émergence d'une réelle volonté politique visant à transformer les enjeux économiques et politiques actuels derrière le développement de l'IA, dans la perspective d'une société plus égalitaire. **Faire de l'intelligence artificielle un levier pour l'égalité**

- Mesurer les inégalités grâce aux IA pour justifier des politiques publiques ciblées pour les combattre.
- Manifester une volonté politique de faire de l'IA un levier pour l'égalité.
- Utiliser les algorithmes pour diversifier les bases de données, qu'elles soient textuelles ou iconographiques – en créant par exemple automatiquement des images de synthèse³.
- « Encoder l'égalité »⁴ : écrire l'égalité dans le code lui-même, pour que l'IA promeuve l'égalité au-delà des mentalités qui demeurent conservatrices.
- Si les pratiques liées aux robots sexuels demeurent, les penser de sorte qu'elles ne promeuvent pas la culture du viol, voire, au contraire, permettent une éducation au consentement.

³ Adrien Basdevant nous parle de la possibilité de créer synthétiquement des données grâce à l'intelligence artificielle pour diversifier les bases de données et les rendre plus représentatives ; lors de son audition devant la Commission temporaire « Intelligence artificielle » du 19 septembre 2024.

⁴ Expression employée par Flora Vincent, lors de son audition devant la commission temporaire Intelligence Artificielle le 10 octobre 2024.

Introduction

Du XIX^e siècle jusqu'à la fin des années 1970, les femmes occupent une place majeure au sein du champ de la programmation, représentant la moitié des effectifs du secteur dans les années 1950 avec des figures emblématiques comme Ada Lovelace⁵. Cependant, dans les années 1990, la montée en prestige des métiers informatiques et l'utilisation croissante des ordinateurs personnels entraînent une chute à 12 % de leur représentation, ainsi qu'une invisibilisation de leur contribution.

Parce que le secteur de la programmation devient puissant, il se construit progressivement autour de codes virils qui attirent davantage les hommes que les femmes⁶. Cette sous-représentation des femmes dans la conception de l'intelligence artificielle engendre des biais et peut contribuer à la montée des violences, comme le cyberharcèlement. Il semble donc essentiel d'explorer les menaces et les opportunités que l'IA représente pour les femmes, ainsi que les conditions nécessaires à la réalisation de ces opportunités.

Compte-tenu du travail très large et documenté de la commission temporaire, la Délégation aux droits des femmes et à l'égalité a fait le choix d'interroger les impacts de l'IA sur les femmes, et en particulier sur leur place dans la société, au regard de trois thématiques majeures : le travail, la santé et les relations

intimes.

Dans un premier temps, la présente contribution établira le constat que l'IA s'inscrit dans la société inégalitaire que nous connaissons, la reflète et produit des impacts en retour sur la carrière et la santé des femmes. Au-delà, la contribution montre que l'IA tend à conformer socialement les femmes et à les enfermer dans une vision fortement sexualisée. Pourtant, comme le propose la dernière partie, l'IA peut être vectrice d'émancipation, en suggérant des applications favorables à l'égalité professionnelle et pour leur santé.

⁵ La Grande École du Numérique, « Les chiffres clés sur les femmes et la tech », 16 septembre 2024, <https://www.grandeecolenumerique.fr/le-numerique-et-les-femmes/les-chiffres-cles-sur-les-femmes-et-la-tech>.

⁶ D'après Vanessa Nurock, entendue en entretien par la Délégation aux droits des femmes et à l'égalité du CESE, le 10 octobre 2024.

I - La création et l'utilisation de l'IA dans une société inégalitaire risquent de renforcer les inégalités

A. Des algorithmes biaisés à la source...

1. De l'orientation scolaire à la sous-représentation des femmes dans les métiers du numérique

D'après l'INSEE, en 2021-2022, les femmes représentent seulement 24 % des personnes employées dans le numérique en France⁷. Plus spécifiquement, les femmes représentent 12 % des chercheurs en intelligence artificielle et 6 % des développeurs⁸.

Les femmes présentes dans le secteur du numérique sont en moyenne plus jeunes et plus diplômées que dans les autres secteurs – en effet, deux fois plus de femmes de ce secteur ont un diplôme de l'enseignement supérieur long. Pourtant, 58 % des femmes du secteur du numérique ont moins de 5 ans d'ancienneté, et sont moins souvent cadres que les hommes – elles le sont dans 62 % des cas contre 70 % des cas pour les hommes⁹.

Les stéréotypes contribuent à éloigner les filles du numérique. Les parents comme les professeurs orientent davantage les garçons que les filles vers des filières scientifiques. D'après Elyes Jouini¹⁰, ces différences émergent dès le CP, avec un décrochage notable des filles des matières scientifiques à l'entrée en CE1¹¹. Par la suite, quand seulement 33 % des filles sont encouragées par leurs parents à se diriger vers les métiers du numérique, c'est le cas de 61 % des garçons¹². De la même manière, quand 66 % des lycéens envisagent de s'orienter vers une école informatique ou d'ingénieur, c'est le cas de seulement 37 % de lycéennes¹³. Les stéréotypes intériorisés par les individus ont d'ailleurs de forts effets sur la réussite à l'école en fonction du genre et sur les choix d'orientation. Plusieurs études montrent que lorsque les exercices sont qualifiés de problèmes mathématiques ou géométriques, les filles réussissent moins bien que lorsqu'on les décrit comme des

7 La Grande École du Numérique, « Les chiffres clés sur les femmes et la tech », 16 septembre 2024. <https://www.grandeecolenumerique.fr/le-numerique-et-les-femmes/les-chiffres-cles-sur-les-femmes-et-la-tech>.

8 « La place des femmes dans l'Intelligence Artificielle », Compte-rendu d'évènement (Assemblée Nationale, 2024), https://www.assemblee-nationale.fr/dyn/16/rapports/ots/116b2658_rapport-information, p. 15.

9 La Grande École du Numérique, « Les chiffres clés sur les femmes et la tech », 16 septembre 2024, <https://www.grandeecolenumerique.fr/le-numerique-et-les-femmes/les-chiffres-cles-sur-les-femmes-et-la-tech>.

10 Titulaire de la chaire Unesco « Femmes et Science » à l'Université Paris-Dauphine, entendu dans le cadre du rapport n°2658 rendant compte de l'évènement consacré à la place des femmes dans l'Intelligence Artificielle, organisé conjointement par le Sénat et l'Assemblée nationale.

11 « La place des femmes dans l'Intelligence Artificielle », Compte-rendu d'évènement (Assemblée Nationale, 2024), https://www.assemblee-nationale.fr/dyn/16/rapports/ots/116b2658_rapport-information, p. 17.

12 La Grande École du Numérique, « Les chiffres clés sur les femmes et la tech », 16 septembre 2024, <https://www.grandeecolenumerique.fr/le-numerique-et-les-femmes/les-chiffres-cles-sur-les-femmes-et-la-tech>.

13 J. Champetier, « Le numérique ou l'absence des femmes », 2023, <https://euradio.fr/emission/KYQZ-la-cause-des-femmes-en-europe-jade-champetier/ZMB7-le-numerique-ou-labsence-des-femmes>.

jeux ou des exercices de dessin¹⁴.

En outre, la filière du numérique étant largement dominée par les hommes, elle se caractériserait par une « forte culture sexiste »¹⁵. Selon le rapport Gender Scan 2022, 46 % des femmes de la tech déclarent avoir été victimes de comportements sexistes, 8 % de plus que dans tous les autres secteurs¹⁶. Ces comportements problématiques se retrouvent dès les études. En effet, les écoles d'ingénieurs font partie des établissements où le plus de violences sexuelles sont recensées, avec 7 % des étudiantes et étudiants déclarant avoir été victimes d'agressions sexuelles et 10 % témoins¹⁷.

2. Un reflet de la société

D'après Mathilde Saliou¹⁸, les biais algorithmiques peuvent avoir trois sources distinctes : les données d'entraînement, les équipes en charge de leur conception ou encore les objectifs de leur création¹⁹.

Les données utilisées par les algorithmes au cours de leur phase d'apprentissage tendent à manquer de représentativité, on parle alors de biais statistiques²⁰. Les algorithmes apprennent à partir de données réelles, puis leurs modèles de prédiction reproduisent, voire renforcent ces biais²¹. Or, si l'on prend l'exemple de la profession de cadres, qui plus est supérieurs, les hommes y sont sur-représentés, quand les femmes sont bien moins présentes ; ce faisant, l'algorithme à qui l'on demande de représenter un cadre, et en particulier supérieur²², montrera un homme plutôt qu'une femme. L'étude menée par Bloomberg en 2023 illustre bien la reproduction des biais, en demandant à l'IA de Stable Diffusion de générer 5000 images selon des

14 C. Perronet, « Mathématiques et lutte contre les stéréotypes sexués. » (Ministère de l'Éducation nationale, de la jeunesse et des sports, s. d.),

15 Haut Conseil à l'Égalité entre les femmes et les hommes, « La Femme Invisible dans le numérique : le cercle vicieux du sexisme », 2023, <https://www.haut-conseil-egalite.gouv.fr/parite/travaux-du-hce/article/rapport-la-femme-invisible-dans-le-numerique-le-cercle-vicieux-du-sexisme>.

16 « Gender Scan » (Global Contact, 2022), https://www.genderscan.org/Docs/Gender_Scan_Synthese_2022_France.pdf, p. 13.

17 Observatoire étudiant des violences sexuelles et sexistes dans l'enseignement supérieur, « Baromètre des violences sexuelles et sexistes dans l'Enseignement supérieur », 2023, <https://observatoire-vss.com/wp-content/uploads/2023/04/Rapport-detaille-Barometre-2023-des-violences-sexistes-et-sexuelles-dans-lEnseignement-superieur.pdf>, p. 100.

18 Journaliste spécialisée dans le numérique.

19 C. Pasteau, « IA et biais algorithmiques : décryptage avec Mathilde Saliou - Hello Future Orange », Hello Future, 26 janvier 2024, <https://hellofuture.orange.com/fr/biais-algorithmiques/>.

20 *Ibid.*

21 James Holdsworth, « Qu'est-ce que le biais de l'IA ? », IBM, 2023, <https://www.ibm.com/fr-fr/topics/ai-bias>.

22 Même lorsque « supérieure » est au féminin.

consignes précises²³. Le résultat est que « le monde selon Stable Diffusion est dirigé par des PDG masculins blancs. Les femmes sont rarement médecins, avocates ou juges. Les hommes à la peau foncée commettent des crimes, tandis que les femmes à la peau foncée préparent des hamburgers. »²⁴. **Si la contribution s'articule ici principalement autour des biais de genre, cet exemple illustre également les biais produits envers les minorités, qui sont tout aussi problématiques.**

De plus, le manque de femmes dans les métiers du numérique et, par extension, le manque de femmes responsables de la conception des algorithmes, est une source importante des biais discriminatoires – que l'on peut qualifier de biais cognitifs²⁵. Selon Hélène Deckx van Ruys²⁶, « 88 % des algorithmes sont créés par des hommes qui, inconsciemment ou non, reproduisent leurs biais »²⁷. D'après Jean Guo²⁸, si les employés du numérique sont à 24 % des femmes, le domaine de l'IA ne compte lui que 12 % de femmes. De ce fait, **80 %**

des algorithmes seraient conçus par des équipes majoritairement masculines et la parité n'est pas à espérer avant 2100. C'est notamment par la sélection et l'annotation des bases de données que les équipes en charge de leur conception peuvent introduire leurs propres stéréotypes et être à la source de l'amplification des biais par l'algorithme²⁹. L'opacité dans la conception des algorithmes et dans les données sur lesquelles leur apprentissage se fonde est d'ailleurs problématique puisqu'elle rend plus difficile l'identification des sources de discrimination et donc la lutte contre celles-ci.

Enfin, les finalités de la création des algorithmes auraient un impact sur les biais qu'ils produisent ou au contraire ne produisent pas³⁰. Ainsi, si l'objectif est de gagner des clients et de l'argent, avec un modèle économique reposant sur la publicité, la volonté de se prémunir des biais sera moindre que si l'objectif est de fournir un service à forte utilité sociale et en libre accès.

23 Leonardo. Nicoletti et Dina Bass, « Humans Are Biased. Generative AI Is Even Worse », *Bloomberg.Com*, 23 septembre 2024, <https://www.bloomberg.com/graphics/2023-generative-ai-bias/>.

24 Leonardo Nicoletti et Dina Bass, « Humans Are Biased. Generative AI Is Even Worse », *Bloomberg.Com*, 23 septembre 2024, <https://www.bloomberg.com/graphics/2023-generative-ai-bias/> ; In James Holdsworth, « Qu'est-ce que le biais de l'IA ? », *IBM*, 2023, <https://www.ibm.com/fr-fr/topics/ai-bias>.

25 James Holdsworth, « Qu'est-ce que le biais de l'IA ? », *IBM*, 2023, <https://www.ibm.com/fr-fr/topics/ai-bias>.

26 Directrice RSE et copilote du groupe Femmes et IA au Laboratoire de l'égalité.

27 Alexis Graillot, « Femmes et IA : « L'intelligence artificielle reflète les biais de la société » », *Public Sénat*, 8 mars 2024, <https://www.publicsenat.fr/actualites/societe/femmes-et-ia-intelligence-artificielle-reflete-les-biais-de-la-societe>.

28 Fondatrice de Konexio, entendue par le CESE le 3 octobre 2024 dans le cadre de la Commission temporaire « Intelligence artificielle ».

29 « Home | lapassanselles », consulté le 9 octobre 2024, <https://iapassanselles.wixsite.com/presentation>.

30 C. Pasteau, « IA et biais algorithmiques : décryptage avec Mathilde Saliou - Hello Future Orange », *Hello Future*, 26 janvier 2024, <https://hellofuture.orange.com/fr/biais-algorithmiques/>.

B. ... qui ont des impacts invisibles mais réels sur les femmes

1. Un frein insidieux pour la carrière des femmes

Ces biais entraînent des conséquences directes dans le monde réel, notamment en termes de discrimination dans la recherche d'emploi ou l'embauche. Le 12 juin 2023, la Fondation des Femmes, l'association Femmes Ingénieures et l'ONG Global Witness ont saisi la Défenseure des droits et la CNIL de deux plaintes relatives à la discrimination sexiste opérée par les algorithmes de Facebook. Ces plaintes découlent des tests réalisés par l'ONG, montrant que lors de la publication d'offres d'emploi, l'algorithme opère une sélection discriminatoire et sexiste dans les personnes qui sont ciblées par les annonces : **92 % des femmes sont visées par l'annonce du poste de secrétaire, quand 85 % des hommes sont ciblés pour des postes de pilote de ligne.**

Les discriminations peuvent également être présentes dans la sélection des candidates et candidats par les entreprises qui automatisent leur système de recrutement³¹. En 2015, lorsqu'Amazon a automatisé son processus de recrutement, les CV de femmes – comportant le mot « women », ou faisant mention

d'universités réservées aux femmes – étaient systématiquement moins bien notés que les CV d'hommes.

Ces biais découlaient des données utilisées pour entraîner l'algorithme, majoritairement composées de CV d'hommes quand il s'agissait des postes de cadres.

Le recrutement sur les aptitudes comportementales des candidats et candidates peut aussi constituer un risque supplémentaire de discrimination. En effet, les traits qui sont plus souvent associés aux hommes – tels que le **leadership fort ou à la compétitivité** –, seront davantage favorisés par l'IA du fait de son apprentissage sur des données déséquilibrées, favorisant ainsi les candidats masculins³².

D'après l'enquête Indeed « Diversité et inclusion »³³, l'intelligence artificielle peut constituer un frein nouveau pour la carrière des femmes du fait des biais présents dans l'apprentissage, ou encore des dimensions plus fines qui ne peuvent être saisies par l'algorithme. 53 % des personnes interrogées déclarent être conscientes de la présence de ces biais³⁴.

L'exclusion des femmes du secteur numérique a des conséquences économiques importantes pour ces dernières en les privant

31 Patrice Bertail et al., « Algorithmes : biais, discrimination et équité », 2019, Algorithmes : biais, discrimination et équité (telecom-paris.fr).

32 Clementine Collett, Gina Neff, et Livia Gouvea, « The Effects of AI on the Working Lives of Women » (Inter-American Development Bank, 7 mars 2022), The Effects of AI on the Working Lives of Women (iadb.org).

33 Indeed, « Diversité et inclusion : l'IA peut-elle vous aider ? », consulté le 9 octobre 2024, <https://fr.indeed.com/lead/diversite-et-inclusion-ia-peut-elle-vous-aider>.

34 *Ibid.*

d'opportunités dans un secteur où l'emploi croît 2,5 fois plus vite que dans les autres secteurs. C'est aussi les éloigner de statuts stables, carrières prestigieuses et fortement rémunérées³⁵. Si les femmes sont peu représentées dans le secteur du numérique, **elles sont aussi celles dont les emplois sont les plus menacés du fait même du numérique**. En effet, elles auraient 1,5 fois plus besoin de trouver un nouveau poste que les hommes - puisque leurs emplois seraient plus susceptibles d'être automatisés et donc effectués par des algorithmes³⁶.

2. Reconnaître les risques pour la santé des femmes

En santé, l'intelligence artificielle, en l'état, tend aussi à défavoriser les femmes. Depuis 2014 en Espagne, certains hôpitaux ont introduit l'intelligence artificielle dans leurs méthodes pour prédire la compatibilité des donneurs/ donneuses et des receveurs/ receveuses afin d'aider les médecins dans le cadre de greffes de foie. Si cette méthode semble très novatrice et pertinente, elle peut néanmoins comporter des biais en défaveur des femmes. En effet, deux ans plus tard, les médecins se sont rendu compte qu'**aucun organe n'avait été attribué à une femme**. L'apprentissage de

l'algorithme avait été réalisé sur la base d'études qui n'avaient pris en compte que peu de femmes et n'avaient pas permis de constater la différence des biomarqueurs impliqués dans les greffes impliquant des hommes et des femmes³⁷.

Plus largement, bien que les femmes vivent plus longtemps, elles sont aussi plus souvent victimes de maladies chroniques³⁸. Les femmes sont moins bien soignées³⁹ parce qu'il existe des biais dans les diagnostics en fonction du genre, avec la perception de certaines maladies comme « masculines », à l'image des maladies cardiovasculaires⁴⁰.

Ainsi, l'introduction des algorithmes en santé, si ceux-ci apprennent sur des données déjà produites - qui tendent à invisibiliser les particularités de la santé des femmes -, on ne peut que supposer une amplification des discriminations à l'encontre des femmes dans ce domaine.

35 La Grande École du Numérique, « Les chiffres clés sur les femmes et la tech », 16 septembre 2024, <https://www.grandeecolenumerique.fr/le-numerique-et-les-femmes/les-chiffres-cles-sur-les-femmes-et-la-tech>.

36 McKinsey, « Generative AI and the future of work in America », consulté le 9 octobre 2024, <https://www.mckinsey.com/mgi/our-research/generative-ai-and-the-future-of-work-in-america>.

37 Josephine Lethbridge, « Oui, l'IA est sexiste – qu'est-ce qu'on peut faire ? », *Les Glorieuses*, 29 avril 2024, <https://lesglorieuses.fr/oui-lia-est-sexiste/>.

38 Muriel Salle et Catherine Vidal, *Femmes et santé, encore une affaire d'hommes ?*, Belin, coll. Égale à Égal, , 2019, p. 11-12.

39 *Ibid*, p. 21.

40 *Ibid*, p. 36.

II - Comment l'IA sculpte socialement les femmes

A. L'IA à l'image et au service de la société patriarcale

1. Une technologie qui pousse les femmes à s'adapter

En s'entraînant sur des échantillons majoritairement composés de données masculines, l'intelligence artificielle peine à servir les femmes aussi bien qu'elle sert les hommes.

À titre d'exemple, **certains logiciels de reconnaissance vocale ne reconnaissent pas bien les voix féminines et invitent les femmes à masculiniser leur voix pour se faire comprendre**⁴¹. Si la voix des femmes est moins bien reconnue que celle des hommes, c'est aussi le cas de leur visage – comme le montre l'enquête *Gender Shades* publiée en 2018 par le *MIT Lab*⁴². L'enquête montre que les trois compagnies étudiées - Microsoft, Face++ et IBM – avaient de meilleurs résultats sur les hommes que sur les femmes, avec des différences dans les taux d'erreur variant de 8,1 à 20,6 %⁴³. Les écarts sont encore plus importants quand on introduit la **variable d'origine ethnique**. Effectivement, les femmes noires sont celles qui souffrent le plus des lacunes de la reconnaissance faciale, avec un écart de 34,4 %

pour l'algorithme d'IBM entre le taux d'erreur pour les hommes blancs et les femmes noires. Depuis, l'IA apprend et améliore ses reconnaissances, mais la vigilance reste de mise.

2. Cantonner les femmes à des sphères et caractéristiques stéréotypées

En parallèle de la moindre reconnaissance des femmes par les logiciels de reconnaissance vocale, ces logiciels sont incarnés par des prénoms et voix féminines⁴⁴. Si ce double constat peut sembler sans importance, il reflète pourtant un paradigme de la société dans laquelle les **« comportements de commandement »** sont associés aux hommes, quand les **« comportements d'assistant »** sont réservés aux femmes⁴⁵.

Par le même mécanisme, les assistants vocaux - associés aux robots ménagers qu'on dote de caractéristiques corporelles féminines - tels qu'ils sont conçus, reflètent la vision d'une société dans laquelle le travail domestique relève des femmes et est invisibilisé -

41 Graeme McMillan, « It's Not You, It's It: Voice Recognition Doesn't Recognize Women », Time, 1 juin 2011, <https://techland.time.com/2011/06/01/its-not-you-its-it-voice-recognition-doesnt-recognize-women/> ; In Vanessa Nurock, « L'intelligence artificielle a-t-elle un genre ? », PUF, 2019, p. 66.

42 « Gender Shades », consulté le 9 octobre 2024, <https://gendershades.org/>.

43 *Ibid.*

44 Vanessa Nurock, « L'intelligence artificielle a-t-elle un genre ? », PUF, 2019, pp.66-67.

45 *Ibid.*

virtualisé dans ce cas⁴⁶.

Les grands modèles de langage d'Open AI et de Meta, tendent aussi à reproduire des représentations stéréotypées en inscrivant les individus dans des sphères distinctes en fonction de leur genre⁴⁷. Les prénoms féminins sont associés à

la sphère privée et familiale – par l'utilisation de mots tels que « foyer », « famille », « enfants » et « mariage » – quand les prénoms masculins sont rattachés à la sphère publique et professionnelle – avec des mots tels qu'« entreprise », « cadre », « salaire » et « carrière »⁴⁸.

B. Une technologie qui entérine un système de domination jusqu'aux relations intimes

1. Conformisme et ancrage dans des rôles traditionnels : l'exemple des sites de rencontre

Dans le cas des sites de rencontre, ce sont à la fois **les intentions des concepteurs et les logiques informatiques qui semblent problématiques**.

D'une part, les algorithmes des sites de rencontre découlent d'une vision qui tend à reproduire le modèle patriarcal des relations hétérosexuelles⁴⁹. En 2019, Jessica Pidoux⁵⁰ montre que le brevet déposé par Match Group LLC et Match Group Inc en 2018⁵¹, intègre

une vision traditionnelle des rôles de genre : les jeunes femmes peu diplômées et avec un revenu faible ont tendance à être proposées à des hommes plus vieux, diplômés et avec un revenu plus important. L'inverse n'est pas vrai⁵².

En outre, le système d'application de rencontre contraint les femmes à réduire leur apparence à des critères prédéfinis - comme : la taille, la couleur des yeux, la longueur, la couleur ou la coupe de cheveux, la morphologie, le poids et la taille de soutien-gorge - qui fondent leur valeur⁵³. Cette réduction du monde social – et par extension du corps

46 *Ibid.*

47 UNESCO, « Generative AI: UNESCO Study Reveals Alarming Evidence of Regressive Gender Stereotypes », 2023, <https://www.unesco.org/en/articles/generative-ai-unesco-study-reveals-alarming-evidence-regressive-gender-stereotypes>.

48 *Ibid.*, In Josephine Lethbridge, « Oui, l'IA est sexiste – qu'est-ce qu'on peut faire ? », *Les Glorieuses*, 29 avril 2024, <https://lesglorieuses.fr/oui-lia-est-sexiste/>.

49 Jessica Pidoux, « X. Toi et moi, une distance calculée : les pratiques de quantification algorithmiques sur Tinder », in *Carte d'identités*, Hermann, 2019, pp. 249-267, <https://doi.org/10.3917/herm.povoa.2019.01.0249>.

50 Auditionnée devant la Délégation aux droits des femmes et à l'égalité du CESE, le mardi 8 octobre 2024.

51 Sean Rad et al., *Matching Process System And Method*, United States US20140074824A1, filed 21 octobre 2013, et issued 13 mars 2014, <https://patents.google.com/patent/US20140074824/en?q=US2014%2f0074824+A1>.

52 D'après les propos tenus par Jessica Pidoux, auditionnée par la Délégation au droit des femmes et à l'égalité, le 8 octobre 2024.

53 Jessica Pidoux et Pascale Kuntz, « Digitamorphosis of feminine bodies on affective dating applications », 2024, <https://www.openstarts.units.it/entities/publication/8ee00e1e-4dbb-459d-9311-9e5ae5f8fa7c>.

des femmes⁵⁴ - découlerait de contraintes techniques imposant un nombre de variables restreint, mais aussi de choix humains. Ces valeurs moyennes seraient le produit « d'un récit hétéronormatif sur les normes de beauté de la société concernant l'attrait des femmes »⁵⁵. Jessica Pidoux, ayant réalisé une étude comparative de trente applications, explique que ces applications tendent à normaliser un certain mode de rencontre d'après les valeurs de concepteurs et développeurs - qui sont principalement masculins. Par la suite, on assiste à une standardisation de la mise en relation des personnes par la diffusion massive des applications de rencontre et la circulation des algorithmes⁵⁶ – posant dès lors un cadre normatif uniforme sur la manière de s'appareiller en ligne⁵⁷.

2. La reproduction d'une vision sexualisée des femmes

C'est donc une vision de la société dans lesquelles les femmes sont

sexualisées, leurs corps érotisés et assignés à des normes de beauté propres au regard masculin⁵⁸, qui peut être véhiculée par les intelligences artificielles. En 2019, l'UNESCO publie un rapport⁵⁹ dans lequel est montré que les assistantes personnelles reproduisent et donc promeuvent à leur tour le système de domination patriarcale, en étant programmées pour encourager le flirt même quand celui-ci devient agressif et insultant à leur égard⁶⁰.

Les robots sexuels participent aussi à la diffusion de cette vision de la société. En effet, d'après Ezekiel Kwetchi Takam⁶¹, ces robots permettent de réactualiser la position dominante des consommateurs à l'aune du mouvement #meetoo, et constituent « le dernier bastion de leur pouvoir »⁶². Conçu « par les hommes, pour les hommes », ces « sexbots » entretiennent l'idée selon laquelle le rôle des femmes est de « satisfaire les désirs masculins ». En somme, les « robotesses sexuelles » constituent le « continuum de la domination

54 D'après Jessica Pidoux, les femmes seraient décrites bien plus en détail que les hommes – *Adopteunmec* allant même jusqu'à leur faire renseigner le type de culottes qu'elles portent.

55 Jessica Pidoux et Pascale Kuntz, « Digitamorphosis of feminine bodies on affective dating applications », 2024, <https://www.openstarts.units.it/entities/publication/8ee00e1e-4ddb-459d-9311-9e5ae5f8fa7c>, p. 236.

56 S'expliquant notamment par l'environnement technologique et le fonctionnement économique : Google et Apple fournissent en effet des logiciels types qui contribuent à la standardisation des applications ; In Jessica Pidoux et Pascale Kuntz, « Concevoir des applications informatiques respectueuses des humains », *The Conversation*, 30 décembre 2021, <http://theconversation.com/concevoir-des-applications-informatiques-respectueuses-des-humains-161950>.

57 Propos issus de l'audition de Jessica Pidoux du 8 octobre 2024.

58 Ou « male gaze », concept théorisé par Laura Mulvey ; In Laura Mulvey, « Plaisir visuel et cinéma narratif », 1975.

59 UNESCO, « I'd blush if I could: closing gender divides in digital skills through education », *EQUALS*, 2019, <https://unesdoc.unesco.org/ark:/48223/pf0000367416>.

60 Vanessa Nurock, « L'intelligence artificielle a-t-elle un genre ? », PUF, 2019, pp. 66-67.

61 Doctorant en éthique théologique à l'Université de Genève.

62 Florise Vaubien, « Sexbots: une pratique sexuelle qui se démocratise », infoSport, rts.ch, 18 mars 2024, <https://www.rts.ch/info/sciences-tech/2024/article/sexbots-une-pratique-sexuelle-qui-se-democratise-28438897.html>.

masculine en contexte sexuel »⁶³.

Le développement des IA dans le domaine de l'intime, pose en outre un enjeu fort en termes de consentement. D'après Anna Choury⁶⁴, le problème est que l'IA - n'étant pas dotée de conscience - ne peut ni consentir ni ne pas consentir. Or, ce « consentement par défaut » s'inscrirait à l'encontre des valeurs que l'on tente d'inculquer aux jeunes générations⁶⁵.

Il est aussi intéressant d'observer la manière dont l'offre dans ce secteur découle d'une conception stéréotypée des attentes des consommateurs femmes et hommes. Si l'on regarde par exemple le site Dollsfrance, l'offre de poupées féminines est largement supérieure et vante leur morphologie et orifices permettant toutes pratiques sexuelles quand l'offre de poupée masculine est bien plus restreinte et s'articule autour d'une publicité en lien avec la tendresse et la possibilité de se blottir dans les bras des mannequins.

De plus, si les intelligences artificielles s'incarnent dans des corps, elles sont avant tout présentes dans les *smartphones* et

principalement pour les hommes, au travers des *AI girlfriends*. **Au-delà des questionnements que pose le développement de ces IA en termes d'évolutions sociétales⁶⁶, ce sont aussi les normes qu'elles véhiculent qui sont problématiques⁶⁷.** En effet, c'est l'image d'une petite amie idéale toujours disponible, d'humeur égale et n'existant que pour plaire, ne pas contredire, qui est ainsi véhiculée.

Dans le cas des IA de l'intime, le **modèle économique** au sein duquel elles se développent et les conditions de la conception de ces IA est problématique. En effet, l'économie de l'attention⁶⁸ s'enrichit des rapports de domination hommes-femmes et les données d'apprentissage reflètent des valeurs dominantes problématiques de notre société⁶⁹.

63 Pauline Noiseau, « Les enjeux éthiques de la robotique sexuelle : une perspective critique féministe » (2020), <https://doi.org/10.13140/RG.2.2.15844.71042>, p. 2.

64 Experte en intelligence artificielle à l'Institut national des sciences appliquées de Toulouse. Auditionnée par la Délégation aux droits des femmes et à l'égalité du CESE, le mardi 8 octobre 2024.

65 D'après les propos d'Anna Choury, auditionnée par la Délégation aux droits des femmes et à l'égalité, le 8 octobre 2024.

66 La rue devient un espace de danger plutôt que de rencontre, quand la vie numérique est foisonnante. Le développement des IA dans le domaine de l'intime est à la fois la cause et la conséquence de ces évolutions.

67 Audition d'Anna Choury devant la Délégation aux droits des femmes et à l'égalité du CESE du 8 octobre 2024.

68 Cette expression désigne un modèle dans lequel l'attention humaine est considérée comme une ressource rare et précieuse que les entreprises et médias se disputent ; In Yves Citton, *L'économie de l'attention*, La Découverte, 2014, <https://shs.cairn.info/l-economie-de-l-attention--9782707178701>.

69 Dans le cas du domaine de l'intime, les IA apprennent notamment sur du contenu pornographique qui amplifie domination, culture du viol, racisme, etc.

III - Comment développer une IA vectrice d'émancipation ?

A. Gommer la (fausse) neutralité de l'IA

1. En faveur de l'égalité professionnelle

La croissance 2,5 % plus rapide des emplois du secteur du numérique peut constituer une opportunité pour les femmes **si ces dernières sont davantage incitées à se diriger vers ce type de professions**. Cela suppose aussi une représentation du secteur d'activité dont les finalités sont plus diversifiées que ce que laisse supposer la culture « geek ». Une plus grande pluridisciplinarité des équipes de conception pourrait participer à un changement de culture.

Bien que les femmes utilisent pour l'instant moins l'IA dans leur travail que leurs homologues masculins⁷⁰, la maîtrise de cette technologie pourrait aussi devenir un critère d'employabilité dont les femmes seraient en mesure de se saisir.

Mais pour ce faire, il est nécessaire de développer des outils qui ne produisent pas de discriminations à leur encontre, et en lesquels il leur sera donc possible d'avoir confiance. Plusieurs projets voient le jour pour former les femmes à l'utilisation de l'IA et les intégrer dans leur contrôle⁷¹.

D'après Frédéric Bardeau⁷², l'IA peut constituer un « outil de libération » des femmes en leur donnant une forme d'« autonomie » et d'« [in] dépendance »⁷³. En outre, selon Hélène Chinal⁷⁴, par les nouvelles compétences que l'IA amène à acquérir au fil du temps, elle pourrait permettre à toutes et tous d'accéder à des postes de plus grande qualité - notamment en automatisant les tâches inintéressantes⁷⁵.

D'après Laura McGee⁷⁶, l'IA aurait aussi le potentiel de **minimiser l'influence des relations**

70 Nitzan Engelberg, « Les femmes utilisent moins l'IA et mettent en danger leur carrière », 2024, <https://www.welcometothejungle.com/fr/articles/femmes-utilisent-moins-ia-mettent-en-danger-leur-carriere>.

71 Comme se propose de le faire *Women Defining AI* ; In « Home », Women Defining Ai, consulté le 9 octobre 2024, <https://www.womendefiningai.com>.

72 Président et co-fondateur du centre de formation Simplon.

73 Sophy Caulier, « Les femmes et les hommes sont-ils égaux face à l'avènement de l'intelligence artificielle dans les entreprises ? », 24 janvier 2024, https://www.lemonde.fr/emploi/article/2024/01/24/les-femmes-et-les-hommes-sont-ils-egaux-face-a-l-avenement-de-l-intelligence-artificielle-dans-les-entreprises_6212659_1698637.html.

74 Coordinatrice du groupe de travail sur la formation IA et biais de genre au sein d'Impact AI, un collectif engagé pour une intelligence artificielle éthique et responsable ; In « Accueil », impact-ai, consulté le 11 octobre 2024, <https://www.impact-ai.fr/>.

75 Géraldine Russell, « L'intelligence artificielle, meilleur coach professionnel des femmes ? », *Intelekto*, 14 mars 2024, <https://intelekto.fr/36438-intelligence-artificielle-carriere-femmes/>.

76 PDG et fondatrice de Diversio, une plateforme de pour l'inclusion et la diversité sur le lieu de travail, et qui propose des formations diverses sur ces thématiques ; In « Diversity & Inclusion Experts, DEI Platform & Training », Diversio, consulté le 11 octobre 2024, <https://diversio.com/>.

interpersonnelles qui peuvent beaucoup peser dans l'évolution de carrières des individus, en défaveur des femmes. Effectivement, des études ont prouvé que les temps informels et le présentéisme - par exemple tard le soir - favorisent la cooptation entre hommes, alors que les femmes, en ayant encore largement la charge de leur famille, tendent à être exclues de ces temps-là⁷⁷. Ainsi, **l'intelligence artificielle aurait le pouvoir d'endiguer le poids de la cooptation**. Dès lors, si la carrière des individus tenait compte des évaluations basées sur l'IA, les salariées et salariés seraient jugés sur leurs résultats, et non en fonction de leurs relations avec leurs collègues⁷⁸. De la même manière, Jean Guo⁷⁹ indique que 68 % des recruteurs pensent que l'IA peut réduire les biais dans les processus de sélection lors des recrutements.

Plusieurs initiatives qui ont pour vocation de rendre l'IA plus égalitaire ont vu le jour. Le laboratoire de l'Égalité propose

notamment un *Pacte pour une intelligence artificielle égalitaire*, avec un volet portant sur la conception de bases de données et d'algorithmes plus représentatifs, et un autre sur la formation et l'insertion des femmes dans les métiers du numérique⁸⁰. Le laboratoire de l'Égalité a également développé un *guide de bonnes pratiques pour une IA égalitaire* qui porte sur l'enseignement supérieur, la recherche et les entreprises⁸¹. Le *Réseau des femmes pour une IA éthique* de l'UNESCO prend aussi plusieurs initiatives pour favoriser la participation des femmes au développement des intelligences artificielles⁸². Enfin, *WomenInIA* organise des formations, du mentorat et du réseautage pour les femmes intéressées par des carrières dans ce domaine⁸³.

Dans son rapport⁸⁴, Indeed préconise d'utiliser l'IA de manière responsable dans le domaine de l'emploi en prêtant attention à la diversité au sein des entreprises

77 Lætitia Vitaud, « Quelle politique de cooptation pour favoriser la diversité ? », *Welcome to the Jungle*, consulté le 14 octobre 2024, <https://www.welcometothejungle.com/fr/articles/politique-cooptation-diversite>; Laurence Chabrun, « Présentéisme : un phénomène discriminant pour les femmes », *Novethic* (blog), 27 juin 2014, <https://www.novethic.fr/actualite/social/discrimination/isrse/les-femmes-lesees-par-le-presenteisme-142625.html>.

78 « Diversité et inclusion : l'IA peut-elle vous aider ? », Indeed, consulté le 9 octobre 2024, <https://fr.indeed.com/lead/diversite-et-inclusion-ia-peut-elle-vous-aider>.

79 Fondatrice de Konexio, une entreprise qui prône l'insertion sociale et professionnelle par le numérique ; In « Insertion Sociale & Professionnelle par le Numérique », Konexio, consulté le 11 octobre 2024, <https://www.konexio.eu>.

80 Laboratoire de l'Égalité, « Le Pacte pour une intelligence artificielle égalitaire », 2021, <https://www.laboratoiredeegalite.org/le-pacte-pour-une-intelligence-artificielle-egalitaire-entre-les-femmes-et-les-hommes/>.

81 Laboratoire de l'Égalité, « Le guide de bonnes pratiques pour une Intelligence artificielle égalitaire entre les femmes et les hommes – laboratoiredeegalite », 2023, <https://www.laboratoiredeegalite.org/le-guide-de-bonnes-pratiques-pour-une-ia-egalitaire-entre-les-femmes-et-les-hommes/>.

82 UNESCO, « Intelligence artificielle : l'UNESCO lance le Réseau des femmes pour une IA éthique, plateforme d'expertes pour faire progresser l'égalité des genres », consulté le 9 octobre 2024, <https://www.unesco.org/fr/articles/intelligence-artificielle-lunesco-lance-le-reseau-des-femmes-pour-une-ia-ethique-plateforme>.

83 « Women in AI (#WAI) », consulté le 9 octobre 2024, <https://www.womeninai.co>.

84 Indeed, « Diversité et inclusion : l'IA peut-elle vous aider ? », consulté le 9 octobre 2024, <https://fr.indeed.com/lead/diversite-et-inclusion-ia-peut-elle-vous-aider>.

qui produisent les IA utilisées, en ne se fiant pas seulement aux compétences affichées sur les CV et en faisant place au jugement humain – **l'IA ne devant constituer qu'une aide à la décision.**

2. Pour mieux prendre en charge la santé des femmes

En matière de santé, comme cela a été souligné en première partie, les femmes sont moins bien prises en charge et soignées⁸⁵ du fait des biais dans les diagnostics en fonction du genre. En ce sens, **la prise en compte des discriminations de genre en santé est un prérequis essentiel pour mieux soigner les femmes**⁸⁶. C'est seulement en ayant conscience des différences biologiques – mais aussi les vécus sociaux et économiques distincts - entre femmes et hommes, et en se développant dans un terreau neutre de discriminations et de biais que l'intelligence artificielle peut constituer une opportunité pour une meilleure santé des femmes. Pour cela, les recherches – qui seront ensuite mobilisées pour l'apprentissage des algorithmes – doivent être plus représentatives.

La formation des professionnelles et professionnels de santé est

également primordiale⁸⁷, parce que c'est à partir de leur diagnostic d'aujourd'hui que pourraient être déduits les diagnostics de demain.

De nombreuses institutions et organisations ont initié un changement dans les pratiques⁸⁸. C'est le cas du programme « Gender innovation », hébergé par l'université de Stanford et dirigé par Londa Schiebinger, qui développe des programmes d'intelligence artificielle féminins et prend en compte les effets conjoints du sexe et du genre dans la recherche⁸⁹. Les recherches innovantes de l'*Office of Research on Women's Health* des National Institutes of Health, permettront également d'améliorer la santé des femmes.

Un autre exemple inspirant est celui de la multiplication d'initiatives pour **l'introduction de l'intelligence artificielle en radiologie pour détecter les cancers du sein**. Développée notamment en France par l'institut Marie Curie et Ibex depuis 2022, sous le nom de *Galen Breast*⁹⁰, aux Etats-Unis avec *AsymMirai*⁹¹, ou encore au Danemark au travers du programme *Transpara*, cette technologie pourrait avoir des effets très positifs sur la santé

85 Muriel Salle et Catherine Vidal, *Femmes et santé, encore une affaire d'hommes ?*, Belin, coll. Égale à Égal, 2017, p. 21.

86 *Ibid.*, p. 62.

87 Alyson McGregor, *Le sexe de la santé. Notre médecine centrée sur les hommes met en danger la santé des femmes*, Erès, 2021.

88 *Ibid.*, p. 187.

89 « Gendered Innovations », Stanford University, consulté le 9 octobre 2024, <https://genderinnovations.stanford.edu/> ; In : Alyson McGregor, *Le sexe de la santé. Notre médecine centrée sur les hommes met en danger la santé des femmes*, Erès, 2021, pp. 181-182.

90 Institut Curie, « Diagnostic du cancer du sein : l'intelligence artificielle d'Ibex bientôt réalité clinique à l'Institut Curie », 2022, <https://curie.fr/actualite/publication/diagnostic-du-cancer-du-sein-intelligence-artificielle-dibex-bientot-realite>.

91 Jon Donnelly et al., « AsymMirai: Interpretable Mammography-based Deep Learning Model for 1-5-year Breast Cancer Risk Prediction », *Radiology* 310, n° 3 (mars 2024), <https://doi.org/10.1148/radiol.232780>.

des femmes. En effet, *Transpara* a permis : la réduction de la charge de travail des radiologues de 33,5 %, l'amélioration de la qualité du dépistage avec une baisse du taux de rappel de 20,5 %, l'augmentation du taux de détection de cancer et la valeur prédictive positive et la réduction du taux de faux positifs⁹². Ces initiatives témoignent d'un réel engouement sur la question de l'amélioration du traitement des femmes grâce à l'intelligence artificielle.

D'après Mélanie Heard⁹³, « l'IA peut ainsi apporter un gain majeur pour la santé des femmes aussi bien d'un point de vue de la recherche médicale que des parcours de santé. En effet, de nombreuses solutions sont développées dans le champ de la santé reproductive, la santé sexuelle, la prise en charge et la

meilleure connaissance de maladies chroniques comme l'endométriose, mais également le suivi de la maternité et du post partum ou encore l'oncologie. Ces domaines de déploiement de l'IA sont essentiels en matière de santé publique et peuvent permettre de rattraper le retard existant en matière de santé des femmes dans le cadre d'un écosystème FemTech France en structuration rapide. »⁹⁴.

Si l'intelligence artificielle peut être bénéfique pour la santé des femmes comme des hommes, à terme, il semble d'abord primordial de développer la formation médicale sur les différences de sexe et de genre et de conduire des recherches à ce sujet⁹⁵, afin que les données utilisées pour développer des algorithmes soient exemptes de biais sexistes.

B. Les leviers pour y parvenir

1. Un outil au service de l'égalité...

Le 19 septembre 2024, auditionné par le CESE⁹⁶, Adrien Basdevant⁹⁷ présente la technique des données de synthèse⁹⁸ comme solution

potentielle aux échantillons non représentatifs - qui par conséquent, comportent et reproduisent des biais. Permettre l'apprentissage des algorithmes sur des échantillons plus représentatifs et encourager la production de données de synthèse

92 Mélanie Heard, « IA et Santé : pourquoi l'action publique ne peut plus attendre » (Terra Nova, 3 juillet 2024), <https://tnova.fr/societe/sante/ia-et-sante-pourquoi-laction-publique-ne-peut-plus-attendre/>, pp. 38-39.

93 Responsable du pôle santé du think tank Terra Nova.

94 *Ibid.* En outre, le think tank Terra Nova - producteur de ce rapport - apporte par ailleurs des préconisations concrètes pour poser les bases d'une politique publique de l'intelligence artificielle au service de la prévention et de la santé.

95 Alyson McGregor, *Le sexe de la santé. Notre médecine centrée sur les hommes met en danger la santé des femmes*, Erès, 2021, p. 187.

96 Audition menée dans le cadre de la Commission temporaire intelligence artificielle, le 19 septembre 2024.

97 Avocat en droit des nouvelles technologies, membre du Conseil National du Numérique, co-auteur de *L'Empire des données*, 2018.

98 Les données de synthèse, sont des données fabriquées de toutes pièces, similaires aux données humaines et qui permettent l'entraînement des algorithmes.

serait un moyen pertinent d'agir contre les biais de genre.

De plus, comme l'indiquent Aude Bernheim et Flora Vincent⁹⁹, l'IA peut être utilisée pour produire des statistiques sur les inégalités de représentation dans le monde réel – parce qu'il est important de chiffrer ces dernières pour les rendre tangibles –, afin de mieux les combattre par la suite¹⁰⁰. La fondatrice de l'entreprise de formation Konexio, Jean Guo, souligne par ailleurs la possibilité d'utiliser les algorithmes pour détecter et corriger les biais au sein des bases de données¹⁰¹.

En outre, l'intelligence artificielle peut permettre de donner plus de visibilité aux femmes. Elle peut par exemple être utilisée pour générer automatiquement des bibliographies sur Wikipédia pour des femmes ayant été trop peu mises en avant¹⁰².

Enfin, l'IA introduit de nouvelles formes de criminalité dont certaines visent principalement les femmes, **on peut par exemple parler des « deepfakes »¹⁰³, qui sont à 96 %**

des « deepnudes », et ciblent

dans 99 % des cas les femmes. À ce sujet, Ysens de France¹⁰⁴ nous indique en revanche que les autorités développent des applications basées sur l'IA, pour détecter ce type de contenus. Au-delà des « deepfakes », l'IA pourrait être utile pour lutter contre la cybercriminalité ou les contenus illicites, et aider à réduire l'accès des mineurs aux sites pornographiques.

2. ... à condition de répondre à certaines exigences structurelles et politiques

Sur le plan des régulations, si l'*AI Act*¹⁰⁵ ne comprend pas de volet spécifique pour limiter les violences de genre, il pourrait en revanche, par son classement des intelligences artificielles en fonction du risque qu'elles présentent, contribuer à l'interdiction ou à la mise en conformité d'un certain nombre d'entre elles. Cependant, il semble largement insuffisant au regard des débats politiques et idéologiques que l'IA sous-tend¹⁰⁶.

99 Autrices de *L'Intelligence artificielle, pas sans elles* ; In Aude Bernheim et Flora Vincent, *L'Intelligence artificielle, pas sans elles*, Belin, Égale à Égal, 2019.

100 « Home | lapassanselles », consulté le 9 octobre 2024, <https://iapassanselles.wixsite.com/presentation>.

101 Propos recueillis dans le cadre de la Commission temporaire intelligence artificielle du 3 octobre 2024.

102 Comme l'indique Aude Bernheim et Flora Vincent sur le site web de leur ouvrage *Intelligence artificielle, pas sans elles* ; « Home | lapassanselles », consulté le 9 octobre 2024, <https://iapassanselles.wixsite.com/presentation>.

103 Ces « hypertrucages » en français, sont définies par Vie publique comme des « vidéos, images et autres contenus, notamment à caractère sexuel, visant à nuire générés par intelligence artificielle ».

104 Docteure en droit, IA et robotique militaire, Chargée de mission IA auprès du Coordonnateur IA pour la Gendarmerie nationale.

105 *L'AI Act* est une réglementation européenne adoptée le 13 mars 2024 qui vise à réglementer les produits d'intelligence artificielle commercialisés sur le marché européen. La réglementation adopte une approche « par les risques » dont elle définit le niveau ; In « Qu'est-ce que l'« AI Act » ? », info.gouv.fr, consulté le 23 octobre 2024, <https://www.info.gouv.fr/actualite/quest-ce-que-lai-act>.

106 Au cours de son audition par la CT IA, Bilel Benbouzid nous invite par ailleurs à nous méfier de l'*AI Act*, qui serait un « droit de la compliance » (ou de la conformité) dans la mesure où son application par les agences le rend trop peu contraignant.

Il est d'ailleurs complété par le règlement européen sur les services numériques¹⁰⁷, dont le but est de rendre la loi applicable en ligne, et de responsabiliser les plateformes sur lesquelles s'exercent les cyberviolences. Au niveau national, si l'on reprend l'exemple des « deepfakes », ces derniers sont traités par la loi dite SREN¹⁰⁸, qui étend la régulation sur le numérique et introduit d'importantes sanctions financières pour punir leurs auteurs¹⁰⁹. Dès lors, si diverses réglementations existent, elles demeurent méconnues et leur application est insatisfaisante comme le souligne Laure Salmona¹¹⁰. Elles méritent donc d'être largement diffusées et renforcées pour porter un projet plus global portant sur la lutte contre les violences et la promotion de l'égalité.

En outre, lutter contre les biais de genre dans l'IA pourrait passer par une intégration plus grande

des femmes à la conception des intelligences artificielles – notamment en luttant pour une orientation scolaire moins genrée. A titre d'exemple, on peut mentionner certaines femmes qui promeuvent des IA plus représentatives¹¹¹, comme Fei-Fei Li créatrice d'ImageNet¹¹², ou encore de Timnit Gebru, fondatrice de Black in AI¹¹³. Mais pour que l'accroissement du nombre de femmes dans la conception produise ses effets, encore faudrait-il que les femmes n'aient pas elles aussi intériorisé les biais, les comportements et les systèmes de pensée qui entérinent le sexisme. En effet, être une femme ne suffisant pas à s'émanciper des stéréotypes¹¹⁴, la formation et la sensibilisation des développeurs des algorithmes à la question de la reproduction des stéréotypes de genre par ces derniers, semble indispensables pour produire des algorithmes moins biaisés¹¹⁵, tout comme la mise en

107 « Règlement du Parlement européen et du Conseil relatif à un marché unique des services numériques », Pub. L. No. 2022/2065 (2022), <https://eur-lex.europa.eu/legal-content/FR/TXT/PDF/?uri=CELEX:32022R2065>.

108 LOI n°2024-449 du 21 mai 2024.

109 « Art. 22681. – Est puni de deux ans d'emprisonnement et de 60 000 euros d'amende le fait de porter à la connaissance du public ou d'un tiers, par quelque voie que ce soit, un montage à caractère sexuel réalisé avec les paroles ou l'image d'une personne, sans son consentement. Est assimilé à l'infraction mentionnée au présent alinéa et puni des mêmes peines le fait de porter à la connaissance du public ou d'un tiers, par quelque voie que ce soit, un contenu visuel ou sonore à caractère sexuel généré par un traitement algorithmique et reproduisant l'image ou les paroles d'une personne, sans son consentement. » ; In « LOI n° 2024-449 du 21 mai 2024 visant à sécuriser et à réguler l'espace numérique - Dossiers législatifs - Légifrance », consulté le 11 octobre 2024, <https://www.legifrance.gouv.fr/dossierlegislatif/JORFDOLE000047533100/>.

110 Militante féministe et spécialiste des cyberviolences de genre, auditionnée par la commission Education, Culture et Communication du CESE, le 15 octobre 2024.

111 B. Raphael, « Ces 7 femmes redéfinissent l'avenir de l'IA », *Génération IA*, 2024, <https://generationia.flint.media/p/les-7-femmes-qui-reinventent-lintelligence-artificielle>.

112 Une base de données pour la reconnaissance d'images - et militante pour une approche plus inclusive.

113 Une organisation prônant l'intégration des personnes noires aux débats et à la production de l'intelligence artificielle.

114 Comme l'a bien expliqué Jessica Pidoux au sujet de la conception des algorithmes des sites de rencontre, dans son audition du 8 octobre 2024.

115 Jessica Pidoux préconise par exemple l'intégration de sociologue et psychologue formé, à la réflexion pour une méthode formelle d'appariement des individus sur les sites de rencontre.

place d'équipes pluridisciplinaires. De plus, comme le souligne Mathilde Saliou¹¹⁶, en fonction des objectifs qui sous-tendent le développement des intelligences artificielles, les biais varient très largement¹¹⁷. Dès lors, dans le cadre de l'entreprise de formation Konexio, qui place les enjeux de diversité au cœur de ses objectifs, l'usage de l'IA s'inscrit dans la réalisation de cet objectif.

L'intelligence artificielle peut ainsi constituer un levier pour l'égalité si la volonté politique derrière sa conception s'inscrit dans cette perspective¹¹⁸. Comme l'indique Flora Vincent¹¹⁹, produire une IA moins biaisée que l'humain n'est qu'une question de volonté, puisqu'il est techniquement possible d'« **encoder l'égalité** »¹²⁰.

Conclusion

Prendre conscience des biais et construire une gouvernance de l'IA permettant de s'en prémunir est possible, mais il nous faut aussi tenir compte du fait que l'artificialisation

des biais tend à les justifier¹²¹. L'IA, aujourd'hui, est le produit de biais préexistants et n'est ni objective ni impartiale et n'a en aucun cas la compétence de « juger »¹²².

L'infrastructure tout entière et les croyances que nous avons à son sujet sont problématiques. Plus qu'un simple miroir qui reflèterait nos biais, ce que l'IA nous montre tend à nous transformer et donc à transformer la société tout entière¹²³. Parce qu'elle est un reflet du patriarcat et s'inscrit dans un système capitaliste ultralibéral, l'intelligence artificielle pose de réelles questions politiques et démocratiques.

Alors que l'intelligence artificielle se déploie déjà, sa gouvernance doit se traduire par des politiques vectrices d'éthique et d'égalité. Construire un monde qui n'aggrave pas les fractures de genre et les inégalités, voire les résorbe, implique de questionner le rapport de la société à l'IA, d'adopter une gouvernance plus démocratique de l'ensemble de sa chaîne de valeur et d'éduquer sur ses usages.

116 Journaliste française, spécialisée dans les enjeux d'inégalités en lien avec le digital.

117 C. Pasteau, « IA et biais algorithmiques : décryptage avec Mathilde Saliou - Hello Future Orange », Hello Future, 26 janvier 2024, <https://hellofuture.orange.com/fr/biais-algorithmiques/>.

118 D'après les propos tenus par Flora Vincent lors de son audition du 10 octobre 2024 devant la commission temporaire intelligente artificielle.

119 Co-auteurice de *L'Intelligence artificielle, pas sans elles* ; In Aude Bernheim et Flora. Vincent, *L'Intelligence artificielle, pas sans elles.*, Belin, Égale à Égal, 2019.

120 Cet encodage consisterait par exemple à indiquer à l'IA de proposer les mêmes rémunérations pour les femmes et les hommes quand elle est utilisée pour prédire des salaires, à traduire automatiquement des textes en langage épicènes, ou encore à faire remonter autant de CV de femmes que d'hommes dans les logiciels de tri des ressources humaines.

121 Parce que les résultats sont produits par un outil (donc artificiels) supposé neutre, les biais qu'il génère de manière artificielle ne sont pas perçus comme découlant du travail de l'outil puisqu'il n'est pas censé faire « d'erreur » contrairement à un humain. Considérer ces biais comme n'étant pas des erreurs, concourt à les justifier. Par exemple, dans le cas d'une sélection de CV par l'IA, si davantage d'hommes sont retenus, le déséquilibre ne sera pas perçu comme problématique puisque le choix étant fait par une machine soi-disant neutre, ce dernier ne devrait pas être questionné, alors que s'il avait été réalisé par un recruteur la remise en cause aurait été plus évidente.

122 Entretien de Vanessa Nurock, 10 octobre 2024.

123 *Ibid.*

5

Bibliographie

- Actutem, Le marché des puces électroniques pour l'IA promis à une fulgurante croissance jusqu'en 2033, 2024
- ADEME/ARCEP, Analyse prospective publiée en mars 2023
- ADEME, Évaluation de l'impact environnemental du numérique en France et analyse prospective, état des lieux et pistes d'évaluation, janvier 2022
- AFNOR SPEC 2314. (2023). Référentiel général pour l'IA frugale
- Agence du Numérique en Santé, Le Ségur du numérique en santé
- Agence du Numérique en Santé, Les chiffres clés du déploiement du Ségur du numérique en santé
- Agence du Numérique en Santé, Observatoire des incidents de sécurité des systèmes d'information pour les secteurs santé et médico-social 2023, mai 2024
- Professeure Stéphanie Allassonnière et docteur Jean-Louis Fraysse (sous la coordination de), Données de santé artificielles : analyse et pistes de réflexion, Livre blanc, Université Paris Cité, BOT Design, avril 2024
- ANCT, Plan national pour la formation des aidants et médiateurs numériques, mars 2021
- Artificial Intelligence and the Climate Emergency: Opportunities, Challenges, and Recommendations, CellPress
- Assemblée nationale, Proposition de loi, n° 406 visant à réguler l'usage des écrans en présence des enfants de moins de trois ans, 17e législature, octobre 2024
- Assemblée Nationale, « La place des femmes dans l'Intelligence Artificielle », Compte-rendu d'évènement, 2024, https://www.assemblee-nationale.fr/dyn/16/rapports/ots/l16b2658_rapport-information
- Autorité de la concurrence, AVIS 24-A-05 DU 28 juin 2024 relatif au fonctionnement concurrentiel du secteur de l'intelligence artificielle générative
- Pauly Awad et Soraia Oueida, The Potential Impact of Artificial Intelligence on Education: Opportunities and Challenges, SpringerLink, mars 2024
- Baromètre France Digitale - EY sur la performance économique et sociale des startups et fonds de capital-risque, 2024
- Baromètre France Num 2024 : perception et usages du numérique par les TPE et PME, mise à jour décembre 2024
- Baromètre IA en santé : alliée ou menace ?, PulseLife, 2024
- Adrien Basdevant et Jean-Pierre Mignard, *L'Empire des données – Essai sur la société, les algorithmes et la loi*, éd. Don Quichotte, 2018

Antonin Bergeaud, Exposition à l'intelligence artificielle générative et emploi : une application à la classification socio-professionnelle française, INSEE, 2024

Aude Bernheim et Flora Vincent, <https://iapassanselles.wixsite.com/presentation>

Patrice Bertail et al., « Algorithmes : biais, discrimination et équité », 2019, Algorithmes : biais, discrimination et équité (telecom-paris.fr)

Marianne Bléhaut, Jérôme Clerget, Matthieu Serreau, Pascal Plantard, La société numérique française : définir et mesurer l'éloignement numérique, Crédoc avril 2023

Sophy Caulier, « Les femmes et les hommes sont-ils égaux face à l'avènement de l'intelligence artificielle dans les entreprises ? », 24 janvier 2024, https://www.lemonde.fr/emploi/article/2024/01/24/les-femmes-et-les-hommes-sont-ils-egaux-face-a-l-avenement-de-l-intelligence-artificielle-dans-les-entreprises_6212659_1698637.html

Mauro Cazzaniga et al. Gen-AI, Artificial Intelligence and the Future of Work, Fonds monétaire international, 2024

Cahiers pédagogiques, Comment on faisait avant ? IA et pédagogie, mai 2024

Cahiers pédagogiques, Intelligence artificielle et pédagogie, Mickael

Bertrand, Dépasser la crainte de la triche ; Alexandra Maurice, La question n'est pas la bonne, peut-être ?, mai 2024

CFHE, Handicap et intelligence artificielle, mai 2024

Laurence Chabrun, « Présentéisme : un phénomène discriminant pour les femmes », Novethic (blog), 27 juin 2014, <https://www.novethic.fr/actualite/social/discrimination/isr-rse/les-femmes-lesees-par-le-presenteisme-142625.html>

Luc Chagnon, De Facebook à Meta, cinq chiffres qui montrent comment le réseau de Mark Zuckerberg est devenu un géant économique, FranceTV Info, février 2024

Jade Champetier, « Le numérique ou l'absence des femmes », 2023, <https://euradio.fr/emission/KYQz-la-cause-des-femmes-en-europe-jade-champetier/ZMB7-le-numerique-ou-labsence-des-femmes>

Yves Citton, L'économie de l'attention, La Découverte, 2014, <https://shs.cairn.info/l-economie-de-l-attention--9782707178701>

Code des relations entre le public et l'administration, Article L312-1-3 - Légifrance

Clementine Collett, Gina Neff, et Livia Gouvea, « The Effects of AI on the Working Lives of Women », Inter-American Development Bank, 7 mars 2022, The Effects of AI on the Working Lives of Women (iadb.org)

Commission de l'Intelligence artificielle, IA : notre ambition pour la France, mars 2024

Communiqué de presse] Discrimination sexiste des algorithmes de Facebook - Fondation des Femmes | Fondation des Femmes, 12 juin 2023

CNIL, Quelles formalités pour les traitements de données de santé, 2024

CNIL, Accompagnement renforcé : la CNIL sélectionne 3 entreprises du numérique à fort potentiel, 2023

Conseil constitutionnel, Décision du 3 avril 2020, n° 2020-834 QPC

CESE, *Droits sociaux : accès et effectivité*, avis 2024-021 rapporté par Isabelle Doresse et Catherine Pajares y Sanchez au nom de la Commission Affaires sociales et santé, novembre 2024

CESE, *Priorités du CESE pour la nouvelle mandature européenne*, résolution du Bureau 2024-019 rapportée par Catherine Lion sur proposition de la Commission Affaires européennes et internationales, novembre 2024

CESE, *Mieux connecter les Outre-mer*, avis 2024-016, rapporté par Danielle Dubrac et Pierre Marie-Joseph au nom de la délégation aux Outre-mer, octobre 2024

CESE, *Impacts de l'intelligence artificielle : risques et opportunités pour l'environnement*, avis 2024-014, rapporté par Fabienne Tatot et Gilles Vermot Desroches au nom de la commission Environnement, septembre 2024

CESE, *Consommation durable : favoriser une économie de la sobriété pour passer de la prise de conscience aux actes*, avis 2023-021 rapporté par Alain Bazot et Marie-Hélène Meyling au nom de la Commission Économie et finances et de la commission Environnement, juillet 2023

CESE, *Ambition et leviers pour une autonomie stratégique de l'Union européenne dans le domaine économique*, avis 2022-013, rapporté par Catherine Lion et Olivier Mugnier au nom de la Commission Affaires européennes et internationales, septembre 2022

Conseil d'État, Intelligence artificielle et action publique : construire la confiance, servir la performance, août 2022

Johann Corric, Les femmes, les jeunes et les financiers semblent les plus exposés à l'IA, AGEFI, 2024

Cour des comptes, L'intelligence artificielle dans les politiques publiques : l'exemple du ministère de l'économie et des finances, 22 octobre 2024

Défenseure des droits, Algorithmes, systèmes d'IA et services publics : quels droits pour les usagers ? Points de vigilance et recommandations, rapport, novembre 2024

Défenseure des droits, Dématérialisation des services publics : trois ans après, où en est-on ?, rapport, 2022

Thomas Deneux, L'intelligence artificielle dans l'éducation, Académie de Paris

Direction interministérielle de la transformation publique, Label de qualité pour les services publics

Jon Donnelly *et al.*, « AsymMirai: Interpretable Mammography-based Deep Learning Model for 1-5-year Breast Cancer Risk Prediction », *Radiology* 310, n° 3, mars 2024, <https://doi.org/10.1148/radiol.232780>

Mario Draghi, EU competitiveness: Looking ahead - European Commission, 2024

Les Échos, Impôts : des taxes sur les logements vacants envoyées par erreur, 18 novembre 2024

Les Échos, Santé : comment l'IA va révolutionner la médecine, 11 novembre 2024

Econpol, Bocconi Università, Toulouse School of Economics, EU Innovation Policy - How to Escape the Middle Technology Trap, 2024

- EdHu2050 - Le mouvement d'une Éducation centrée sur l'Humain
- Nitzan Engelberg, « Les femmes utilisent moins l'IA et mettent en danger leur carrière », 2024, <https://www.welcometothejungle.com/fr/articles/femmes-utilisent-moins-ia-mettent-en-danger-leur-carriere>
- Euractiv, Malgré ses efforts, le retard de l'Europe en matière d'IA se creuse, novembre 2024
- Ferguson *et al.*, L'IA au travail : propositions pour outiller la confiance, Conférence Nationale sur les Applications Pratiques de l'IA, 2022
- Marina Ferrari, « Qu'est-ce que l'« AI Act » ? », Qu'est-ce que l'« AI Act » ? | info.gouv.fr
- Forum économique mondial. Future of Jobs Report. 2023
- France Digitale et Wavestone, Guide de mise en conformité à l'AI Act : les clés pour comprendre et appliquer la loi sur l'IA, février 2024
- France Digitale, Des puces aux applications : l'Europe peut-elle être une puissance de l'IA générative ? avril 2024
- France Digitale, Mapping des startups de l'IA, mars 2023
- Morgan R. Frank, David Autor, James E. Bessen et Iyad Rahwan, Toward understanding the impact of artificial intelligence on labor, 25 mars 2019.
- Carl Benedict Frey et Michael Osborne. The future of employment : how susceptible are jobs to computerisation ?, 2013
- Joris Galland, Internet des objets et m-santé, octobre 2020
- « Gendered Innovations », Stanford University, <https://genderedinnovations.stanford.edu/>
- Global Contact*, « Gender Scan », 2022 https://www.genderscan.org/Docs/Gender_Scan_Synthese_2022_France.pdf
- Pawel Gmyrek, Janine Berg et David Bescond. Generative AI and Jobs : a global analysis of potential effects on job, quantity and quality, Organisation mondiale du travail, 2023
- Professeur Pierre-Antoine Gourraud, L'intelligence artificielle en santé vue depuis l'an 2044, Le Monde, septembre 2024
- Alexis Graillot, « Femmes et IA : « L'intelligence artificielle reflète les biais de la société » », Public Sénat, 8 mars 2024, <https://www.publicsenat.fr/actualites/societe/femmes-et-ia-lintelligence-artificielle-reflete-les-biais-de-la-societe>
- Le Grand Continent, Les États-Unis devançant l'Union et la Chine en matière d'IA, avril 2024
- Grand Duché du Luxembourg, QTRobot, le compagnon des enfants autistes
- La Grande École du Numérique, « Les chiffres clés sur les femmes et la tech », 16 septembre 2024, <https://www.grandeecolenumerique.fr/le-numerique-et-les-femmes/les-chiffres-cles-sur-les-femmes-et-la-tech>
- HCE, « La Femme Invisible dans le numérique : le cercle vicieux du sexisme », 2023, <https://www.haut-conseil-egalite.gouv.fr/parite/travaux-du-hce/article/rapport-la-femme-invisible-dans-le-numerique-le-cercle-vicieux-du-sexisme>

Mélanie Heard, « IA et Santé : pourquoi l'action publique ne peut plus attendre », Terra Nova, 3 juillet 2024, <https://tnova.fr/societe/sante/ia-et-sante-pourquoi-laction-publique-ne-peut-plus-attendre/>

J. M. Hoc, *From Human-Machine Interaction to Human-Machine Cooperation*, 2000

James Holdsworth, « Qu'est-ce que le biais de l'IA ? », IBM, 2023, <https://www.ibm.com/fr-fr/topics/ai-bias>

IEA, *Global Conference on Energy & AI – Event*

Ifop pour Talan, « Baromètre 2024 Les Français et les IA génératives », avril 2024

Indeed, « Diversité et inclusion : l'IA peut-elle vous aider ? », <https://fr.indeed.com/lead/diversite-et-inclusion-ia-peut-elle-vous-aider>

Institut Curie, « Diagnostic du cancer du sein : l'intelligence artificielle d'Ibex bientôt réalité clinique à l'Institut Curie », 2022, <https://curie.fr/actualite/publication/diagnostic-du-cancer-du-sein-lintelligence-artificielle-dibex-bientot-realite>

INSERM, *Intelligence artificielle et santé, Des algorithmes au service de la médecine*, publié le 6 juillet 2018, modifié le 25 janvier 2024

Institut Sphinx pour le logiciel *Compilatio* en novembre 2023

IRES, *Rapport du projet SeCoIA Deal*, mars 2021-février 2023

Dominique Javel, *Un défi collectif*, dans les Cahiers pédagogiques, IA et pédagogie, mai 2024

Enkelejda Kasneci *et al.*, 2023 dans *Intelligence artificielle et éducation : apports de la recherche et enjeux pour les politiques publiques : édition trilingue 2024*

Gaspard Koenig, *La fin de l'individu - Voyage d'un philosophe au pays de l'intelligence artificielle*, Éditions de l'Observatoire Le Point, Humensis, 2019

Pascale Kuntz et Jessica Pidoux, « Digitamorphosis of feminine bodies on affective dating applications », 2024, <https://www.openstarts.units.it/entities/publication/8ee00e1e-4dbb-459d-9311-9e5ae5f8fa7c>

Pascale Kuntz et Jessica Pidoux, « Concevoir des applications informatiques respectueuses des humains », *The Conversation*, 30 décembre 2021, <http://theconversation.com/concevoir-des-applications-informatiques-respectueuses-des-humains-161950>

Laboratoire de l'Égalité, « Le guide de bonnes pratiques pour une Intelligence artificielle égalitaire entre les femmes et les hommes – *laboratoiredelegalite* », 2023, <https://www.laboratoiredelegalite.org/le-guide-de-bonnes-pratiques-pour-une-ia-egalitaire-entre-les-femmes-et-les-hommes/>

Laboratoire de l'Égalité, « Le Pacte pour une intelligence artificielle égalitaire », 2021, <https://www.laboratoiredelegalite.org/le-pacte-pour-une-intelligence-artificielle-egalitaire-entre-les-femmes-et-les-hommes/>

Joséphine Lethbridge, « Oui, l'IA est sexiste – qu'est-ce qu'on peut faire ? », *Les Glorieuses*, 29 avril 2024 <https://lesglorieuses.fr/oui-lia-est-sexiste/>

Enrico Letta, *Much More Than a Market*, Report, 2024

Lifen Documents, *Gestion sécurisée de vos comptes rendus médicaux*

LOI n° 2024-449 du 21 mai 2024 visant à sécuriser et à réguler l'espace numérique - *Dossiers législatifs – Légifrance*

LOI n° 2024-449 du 21 mai 2024 visant à sécuriser et à réguler l'espace numérique <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000049563368>

Alyson McGregor, *Le sexe de la santé. Notre médecine centrée sur les hommes met en danger la santé des femmes*, Erès, 2021

McKinsey, « Generative AI and the future of work in America », consulté le 9 octobre 2024, <https://www.mckinsey.com/mgi/our-research/generative-ai-and-the-future-of-work-in-america>

Graeme McMillan, « It's Not You, It's It: Voice Recognition Doesn't Recognize Women », Time, 1 juin 2011, <https://techland.time.com/2011/06/01/its-not-you-its-it-voice-recognition-doesnt-recognize-women/>

Stéphanie Marquès, *L'intelligence des élèves plus forte que l'intelligence artificielle*, 2024

MIT Media Lab, *Project Gendershades*, 2020

Philippe Meirieu, *Enseignement : qui a peur des robots conversationnels ?*, 2024

Philippe Meirieu, *Le danger de ChatGPT n'est pas dans la fraude qu'il permet mais dans le rapport aux connaissances qu'il promet*, Le Monde, mars 2023

Ministère de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche, *Le Programme national de formation 2024-2025*, 2024

Ministère de l'Éducation nationale et de la recherche, *IA et éducation, enjeux de la recherche et de la politique publique*, janvier 2024

Le Monde, *Proflage et discriminations : enquête sur les dérives de l'algorithme des caisses d'allocations familiales*, décembre 2023

Laura Mulvey, « Plaisir visuel et cinéma narratif », 1975

Navigating Liquid Cooling Architectures for Data Centers with AI Workloads, Schneider Electric, 2024

Leonardo Nicoletti et Dina Bass, « Humans Are Biased. Generative AI Is Even Worse », Bloomberg.Com, 23 septembre 2024, <https://www.bloomberg.com/graphics/2023-generative-ai-bias/>

Pauline Noiseau, « Les enjeux éthiques de la robotique sexuelle : une perspective critique féministe » (2020), <https://doi.org/10.13140/RG.2.2.15844.71042>

Vanessa Nurock, *L'intelligence artificielle a-t-elle un genre ?*, PUF, 2019

Observatoire étudiant des violences sexuelles et sexistes dans l'enseignement supérieur, « Baromètre des violences sexuelles et sexistes dans l'Enseignement supérieur », 2023, <https://observatoire-vss.com/wp-content/uploads/2023/04/Rapport-detaille-Barometre-2023-des-violences-sexistes-et-sexuelles-dans-lEnseignement-superieur.pdf>

OCDE, *Intelligence Artificielle et les compétences de demain*, 2023

Pierre-Yves Oudeyer, *IA générative, société et éducation : En quoi l'IA générative représente-t-elle un enjeu dans la formation des citoyens ?*, septembre 2024

Hélène Pagésy, *Comment l'IA bouscule le milieu de la santé mentale : « Plutôt*

que de payer une nouvelle séance chez le psy, j'allais sur ChatGPT », Le Monde, 2024

Benjamin Pajot, Les risques de l'IA. Enjeux discursifs d'une technologie stratégique | Ifri, 12 juin 2024

Parlement européen, « Règlement du Parlement européen et du Conseil relatif à un marché unique des services numériques », Pub. L. No. 2022/2065 (2022), <https://eur-lex.europa.eu/legal-content/FR/TXT/PDF/?uri=CELEX:32022R2065>

Parlement européen. Intelligence artificielle : définition et utilisation. 2020.

C. Pasteau, « IA et biais algorithmiques : décryptage avec Mathilde Saliou - Hello Future Orange », Hello Future, 26 janvier 2024, <https://hellofuture.orange.com/fr/biais-algorithmiques/>

PEGA, Tout savoir sur la complexité de la main-d'œuvre moderne, 2022

Clémence Perronet, « Mathématiques et lutte contre les stéréotypes sexués. » (Ministère de l'Éducation nationale, de la jeunesse et des sports, s. d.), <https://www.google.fr/>

6

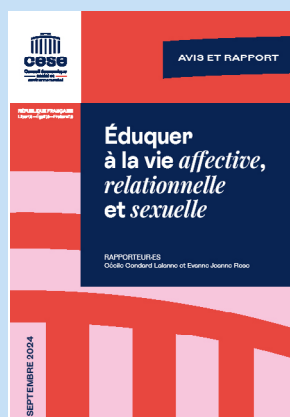
Table des sigles

AFNOR	Association française de normalisation
AGEFI	Agence économique et financière
AI Act	Règlement européen sur l'intelligence artificielle
AI4T	Artificial Intelligence for and by Teachers (Intelligence artificielle pour et par les professeurs)
ANCT	Agence nationale de la cohésion des territoires
ANI	Accord national interprofessionnel
ANS	Agence du numérique en santé
APB	Admission Post Bac
APIA	Applications pratiques de l'intelligence artificielle
ARCEP	Autorité de régulation des communications électroniques, des Postes et de la distribution de la Presse
ARPA	Advanced Research Projects Agency
ASML	Advanced Semiconductor Materials Lithography
CAF	Caisse d'allocations familiales
CESE	Conseil économique, social et environnemental
CFHE	Conseil français des personnes handicapées pour les affaires euro- péennes et internationales
CGDD	Commissariat général au développement durable
CIDPH	Convention internationale des droits des personnes handicapées
CII	Crédit d'impôt innovation
CIR	Crédit d'impôt recherche
CNCDH	Commission nationale consultative des droits de l'homme
CNCPH	Commission nationale consultative des personnes handicapées
CNIL	Commission nationale de l'informatique et des libertés
CNNuM	Conseil national du numérique
CREDO	Centre de recherche pour l'étude et l'observation des conditions de vie
CSA	Comité social d'administration
CSE	Comité social et économique
CSRD	Corporate Sustainability Reporting Directive (Directive relative à la publication d'informations en matière de durabilité par les entreprises)
DDFE	Délégation aux droits des Femmes et à l'égalité
DDT	Direction départementale des territoires
DGE	Direction générale des entreprises

DGFIP	Direction générale des finances publiques
DIAL-IA	Dialoguer sur l'IA
DINUM	Direction interministérielle du numérique
DREAL	Direction régionale de l'environnement, de l'aménagement et du logement
DUERP	Document unique d'évaluation des risques professionnels
EHDS	European Health Data Space (Espace européen des données de santé)"
EIC	European Innovation Council
EPN	Établissement public national
FCC	Federal Communications Commission (Commission fédérale des communications)
FMI	Fonds monétaire international
GAFAM	Google, Apple, Facebook, Amazon, Microsoft
GEPP	Gestion des emplois et des parcours professionnels
GES	Gaz à effet de serre
GPEC	Gestion prévisionnelle des emplois et des compétences
GPU	Graphics Processing Unit
IA	Intelligence artificielle
IAG	Intelligence artificielle générative
IFOP	Institut d'études opinion et marketing en France
IFRI	Institut français des relations internationales
IGN	Institut national de l'information géographique et forestière
INRAE	Institut national de recherche pour l'agriculture, l'alimentation et l'environnement"
INSERM	Institut national de la santé et de la recherche médicale
IR PME	Réduction d'impôt petite et moyenne entreprise
IRISSO	Institut de recherche interdisciplinaire en sciences sociales
ISO	International Organization for Standardization (Organisation internationale de normalisation)
JEI	Jeune entreprise innovante
LLM	Large Language Model (Grand modèle de langage)
MEN	Ministère de l'Éducation nationale et de la jeunesse
MOOC	Massive Open Online Course
OCDE	Organisation de coopération et de développement économiques
OIT	Organisation internationale du travail
ONG	Organisation non gouvernementale
OPCO	Opérateur de compétence
P2IA	Partenariat d'innovation et d'intelligence artificielle
PIB	Produit intérieur brut
PIMMS	Point d'information médiation multiservices

PME	Petite et moyenne entreprise
PNF	Programme national de formation
POC	<i>Proof of Concept</i>
QPC	Question prioritaire de constitutionnalité
RGPD	Règlement général de protection des données
RH	Ressources humaines
RSE	Responsabilité sociétale des entreprises
RTE	Réseau de transport de l'électricité
SeCoIA	Deal SErvoir la COnfiance dans l'Intelligence Artificielle par le Dialogue
SNCF	Société nationale des chemins de fer français
SNIA	Stratégie nationale pour l'intelligence artificielle
SREN	Sécuriser et réguler l'espace numérique (loi visant à)
TPE	Très petite entreprise
TSA	Troubles du spectre de l'autisme
UE	Union européenne
UNESCO	United Nations Educational, Scientific and Cultural Organization (Organisation des Nations Unies pour l'éducation, la science et la culture)

Dernières publications du Conseil économique, social et environnemental



Retrouvez l'intégralité des travaux du CESE sur le site

eese.fr

Retrouvez le CESE sur les réseaux sociaux



Imprimé par la Direction de l'information légale et administrative, 26, rue Desaix, Paris 15^e, d'après les documents fournis par le Conseil économique, social et environnemental • N° 411240002-000125 - Dépôt légal : janvier 2025 • Crédit photo : Dicom

ecese.fr

9, place d'Éléna
75 775 Paris Cedex 16
01 44 43 60 00



**PREMIER
MINISTRE**

*Liberté
Égalité
Fraternité*

Direction de l'information
légale et administrative



Les éditions des
Journaux officiels

N° 41125-0022

ISSN 0767-4538 ISBN 978-2-11-077573-3



9 782110 775733