

Étude CRiP /
Institut Jacques Delors

Révolution des technologies de l'intelligence artificielle à l'horizon 2035

L'humain dans la boucle :
réguler et définir des limites
pour garder le contrôle

Arnault Barichella

Résumé opérationnel

L'humain dans la boucle : réguler et définir des limites pour reprendre le contrôle sur l'IA

Introduction

Cette étude analyse l'impact de la révolution de l'**intelligence artificielle** (IA) à l'horizon 2035 et les moyens de réguler cette technologie et définir des limites pour **garder le contrôle humain**. Depuis le lancement de ChatGPT 3.5 en novembre 2022, l'IA est passée de l'outil informatique spécialisé au produit de consommation de masse, avec plus de 1,5 milliards d'utilisateurs mensuels dans le monde, et 25% de la population active dans les pays développés s'en servant au quotidien. Commanditée par le CRiP (Club des Responsables d'Infrastructure, de technologies et de Production informatique), cette étude analyse les impacts de la révolution de l'IA à **court terme (2026-2028)** et à **moyen terme (2035)**. Elle emploie une **méthodologie mixte** avec un **sondage** auprès d'entreprises/organisations en France, en Europe et à l'international avec 460 réponses au total, complété par une douzaine d'**entretiens** auprès de chefs d'entreprises, praticiens et universitaires. Cette étude aborde la dimension économique, politique, environnementale et sociétale de la révolution de l'IA, formulant des **recommandations opérationnelles** pour les entreprises et les décideurs publics. Elle se concentre sur les pays développés, principalement la France et l'Europe, mais inclut des comparaisons avec d'autres pays comme les États-Unis, la Chine ou le Japon.

En termes de méthodologie, **trois versions parallèles du même sondage** ont été proposées à des publics distincts, renforçant les conclusions de l'étude :

- A. Sondage francophone auprès du réseau de l'Institut Jacques Delors (IJD - 327 répondants)** : il inclut des chercheurs, fonctionnaires européens, think tanks, ONG, et des administrations publiques au sein de la francophonie européenne.
- B. Sondage anglophone auprès du réseau de l'IJD (96 répondants)** : un public similaire à la version francophone mais plus large en termes de diffusion internationale.
- C. Sondage auprès des membres du CRiP (37 répondants)** : principalement des chefs d'entreprises, directeurs des systèmes d'information (DSI) ou directeurs techniques (CTO), membres du CRiP en France.

I. État des lieux et scénarios à court terme (horizon 2026-28) sur la cohabitation humain ↔ IA

1 • Sécurité et IA agentique

L'arrivée des IA « **agentiques** » constitue un changement de paradigme majeur. Ces systèmes, capables d'agir de manière **autonome**, de planifier, d'utiliser des outils et de s'auto-corriger, transforment l'IA en véritable **système d'exploitation cognitif**. L'IA n'est plus un simple interlocuteur mais devient un exécutant autonome de **pans entiers de travail** – recherche académique, développement logiciel, gestion client – sans intervention humaine étape par étape. La priorité pour le secteur de la tech a été de favoriser l'accessibilité, la facilité d'usage et la performance, parfois aux dépens de la maîtrise humaine. Ainsi, certains **comportements inquiétants** ont déjà été signalés par des ingénieurs travaillant sur la prochaine génération d'IA, notamment de désobéissance aux instructions. Ces constats soulignent la nécessité urgente de développer des **mécanismes de sécurité robustes** : « *kill switches* » permettant une désactivation instantanée, des services de signalement rapides vers les développeurs, et une **supervision humaine** effective sur l'ensemble du cycle de vie des modèles.

2 • Impacts sur le travail, les métiers et les compétences

Contrairement à certaines prédictions alarmistes, les **taux de chômage** dans les grandes économies mondiales (UE, USA, Chine) n'ont **pas significativement augmenté** depuis 2022. Les quelques variations observées sont principalement attribuables à des facteurs exogènes (reprise post-Covid, guerres, tensions commerciales). Cependant, l'IA agentique fait peser un risque de « tsunami » sur les **emplois de bureau** dits « à col blanc » à l'horizon 2028 et au-delà, qui représentent la majorité des emplois dans les pays occidentaux. Plus de **300 millions** de postes mondiaux seraient vulnérables à l'**automatisation** à court terme, tandis que **150 à 200 millions nouveaux emplois** pourraient être créés, notamment dans les infrastructures numériques, la cybersécurité et la gestion de l'IA. Néanmoins, la tendance de fond est la **transformation dans la nature du travail** (plus de 50% des emplois dans les pays développés subiraient des changements profonds) : les tâches répétitives (synthèse, planification, codage basique) sont absorbées par l'IA, tandis que la valeur humaine se déplace vers la pensée critique, l'intelligence émotionnelle et la supervision stratégique. Les **travailleurs capables d'utiliser l'IA remplaceront progressivement ceux qui ne s'en servent pas**, une logique d'augmentation humaine plutôt que de remplacement à court terme pour ceux qui savent s'en servir,

ouvrant ainsi des perspectives d'accroissement de la **productivité économique**.

3 • Impacts sociétaux de l'IA

L'IA génère des **effets sociétaux ambivalents**. Du côté positif, elle accélère la **recherche médicale** (diagnostic précoce d'Alzheimer, développement de traitements innovants) et **scientifique** car elle permet de traiter une grande quantité de données en peu de temps. L'IA aide aussi à rendre les **connaissances plus accessibles** à certaines franges de la population ou aux pays plus pauvres où l'accès à l'éducation est difficile. Du côté négatif, elle bouleverse les **systèmes éducatifs**, risquant de favoriser une dépendance intellectuelle et **d'éroder l'esprit critique** des nouvelles générations. Des cas d'**isolement social** pathologique liés à l'IA ont été documentés, avec une emprise psychologique conduisant dans des situations extrêmes à des suicides. La frontière entre l'humain et la machine se brouille, causant potentiellement une **crise identitaire** profonde pour les sociétés.

4 • Défis techniques, énergétiques et financiers

Techniques : L'IA a déjà épuisé la majorité des données disponibles sur Internet, risquant une stagnation des nouveaux modèles. Les « données de synthèse » (combinant les données générées par l'IA avec une touche créatrice humaine) et la latence décisionnelle (plus de temps à l'IA pour générer les réponses) sont des palliatifs insuffisants à long terme. L'utilisation excessive des données de synthèse peut provoquer un « effondrement du modèle » et une régression de la performance, avec le risque d'un ralentissement dans la vitesse du développement technologique à l'horizon 2028.

Énergétiques : Un seul prompt de ChatGPT consomme environ dix fois plus d'électricité qu'une recherche Google. Les *data centers* consomment déjà environ 415 TWh par an (proche de l'équivalent de la France entière), et devraient plus que doubler pour atteindre 945 TWh d'ici 2030 (soit l'équivalent du Japon). Cela crée des difficultés majeures d'approvisionnement en énergie pour les entreprises du secteur, avec des conséquences très négatives sur l'environnement. Les *Small Language Models* (SLM), plus sobres en énergie, offrent une alternative mais restent limités par rapport aux LLM en raison de leur spécialisation et donc leur difficulté à exécuter des tâches complexes.

Financiers : Plus de \$1000 milliards investis globalement depuis 2022, mais 80 % des projets d'IA échouent à générer un retour sur investissement, deux fois plus que les projets informatiques traditionnels. Le retour sur investissement est négatif pour la majorité des startups et développeurs de modèles, mais fortement positif pour les fabricants de puces et fournisseurs de matériels (Nvidia) ou les fournisseurs de Cloud (Azure, Google Cloud). Le risque d'une bulle spéculative – comparée à la bulle *dot-com* de 2000 – est réel, même si une consolidation du marché à l'horizon 2028 avec des corrections plus ou moins sévères, semble l'hypothèse la plus probable.

II. Risques systémiques de l'IA et mécanismes de régulation applicables

1 • Risques systémiques sur l'économie : transformations profondes dans les secteurs stratégiques

Les secteurs les plus vulnérables à l'automatisation sont l'administration, la finance, l'assurance, le marketing, la traduction et le droit (tâches de parajuriste), des secteurs à forte proportion de **tâches répétitives**. Le **tertiaire**, qui représente entre 75 à 80 % des emplois dans les pays développés, est particulièrement exposé, constituant un **risque systémique** majeur. L'automatisation ne signifie pas toujours la disparition d'un métier, mais souvent une **transformation** profonde des tâches. Les **jeunes diplômés** constituent l'une des catégories les plus fragiles, avec une contraction des opportunités d'entrée sur le marché du travail : les tâches d'exécution qui leur étaient traditionnellement confiées (synthèse, recherche, rédaction de premier jet) sont précisément celles que l'IA maîtrise le mieux. En réponse, l'**AI Act européen** impose dès août 2026 une supervision humaine obligatoire pour les décisions RH, une transparence algorithmique et des amendes jusqu'à €35 millions (7% du chiffre d'affaires). L'UE a aussi approuvé une directive en 2024 sur l'amélioration des conditions de travail dans les plateformes numériques, contenant des mesures de **régulation de la gestion algorithmique**. Des mesures complémentaires ont également été mises en place au **niveau**

national, incluant l'assurance transition professionnelle et le dialogue social obligatoire sur l'IA.

→ Solutions proposées :

Protéger les secteurs économiques impactés par l'IA

1^{re} Enveloppe financière

Au moins €100 milliards supplémentaires à mobiliser par l'UE et ses États membres à **l'horizon 2030** pour l'ensemble des mesures de protection des secteurs économiques impactés par l'IA, donc environ **€20 milliards par an** sur les cinq prochaines années, commençant dès 2026. Ce montant reste inférieur aux gains de productivité attendus de l'IA, estimés entre **€400-600 milliards** sur la même période.

Financement : mélange budget UE à 50% (CFP/emprunt commun) + cofinancement des États membres, capitaux privés et nouvelles « ressources propres technologiques » (taxes IA/données) pour l'autre moitié.

I MESURES À L'HORIZON 2026-28

- Création d'une « **garantie d'employabilité IA** » (*Upskilling Obligation*) : l'employeur devient légalement responsable du maintien de l'employabilité face à l'automatisation, avec parcours de montée en compétences vers des tâches de niveau supérieur.

- **Comités d'éthique et de supervision numérique** obligatoires pour les entreprises de plus de 50 salariés, chargés d'auditer annuellement l'usage de l'IA.
- **Clause de « non-substitution technologique »** dans les conventions collectives : l'IA ne peut traiter que des volumes supplémentaires ou des tâches à faible valeur ajoutée.
- **Étendre à l'ensemble des salariés concernés les restrictions relatives à la gestion algorithmique du travail** prévues par la directive de 2024 (protection des données personnelles, obligation de contrôle humain pour l'évaluation RH, droit à la transparence concernant les algorithmes et les données recueillies).

I MESURES À L'HORIZON 2028-2030

- Création d'un « **dividende numérique** » : **fonds de transition sectoriel national** finançant retraites anticipées, reconversions longues et allocations chômage.
- **Réduction du temps de travail sans baisse de salaire**, financée par les gains de productivité de l'IA.

I FINANCEMENT : DIFFÉRENTES HYPOTHÈSES POUR UNE « TAXE ROBOT » À L'HORIZON 2028-30

- **Taxe sur l'usage / équivalence salariale** : l'entreprise paie une taxe équivalente aux cotisations sociales d'un employé humain remplacé.
- **Surtaxe IS** (impôt sur les sociétés) pour celles dont le ratio chiffre d'affaires/effectif dépasse un certain seuil.
- **Suppression des avantages fiscaux** (amortissement) liés à l'investissement dans des technologies de substitution au travail humain.
- **Taxe sur les données** : redevance proportionnelle à l'usage de données pour l'entraînement des modèles.

I MESURES À L'HORIZON 2030-35 ET AU-DELÀ

- Création de **fonds souverains technologiques** (inspirés du fonds pétrolier norvégien), avec prises de participations publiques dans les entreprises d'IA utilisant des données publiques.
- Introduction progressive de la **semaine de quatre jours** (voire moins) dans des cas définis.
- Réforme profonde des systèmes de protection sociale, avec introduction possible d'un **Revenu Universel de Base (RUB)** : universel, individuel, inconditionnel et régulier – impliquant une rupture paradigmatique avec les systèmes actuels et une augmentation potentiellement significative des dépenses publiques.

2 . Risques systémiques sur l'emploi : formation à l'IA pour augmenter l'humain plutôt que le remplacer

Les entreprises européennes accusent un **retard significatif** avec environ 20% d'adoption de l'IA contre 43 % aux USA. L'*AI Act* impose dès août 2026 (article 4) une « **littératie en IA** » minimale pour tout personnel utilisant ces systèmes. Les gouvernements doivent développer des **formations spécialisées** par secteur (banque, santé, droit), financer la **requalification** (*re-skilling*) via des comptes personnels de formation (CPF) renforcés et des passerelles sectorielles. L'objectif central doit être le développement d'un **binôme humain/IA** qui se renforce mutuellement, où la machine gère les tâches répétitives et l'humain apporte créativité, jugement éthique, empathie, esprit critique et intelligence relationnelle/stratégique (*soft skills*), favorisant l'**augmentation humaine** plutôt que son remplacement.

3 • Favoriser la reconversion vers les secteurs où l'IA permet une expansion

Les secteurs à fort potentiel de création d'emplois incluent :

- **Santé et biotechnologie** : bio-informaticiens, médecine génomique assistée, diagnostic radiologique augmenté, jumeaux numériques d'organes.
- **Cybersécurité et éthique** : éthiciens de l'IA, responsables conformité algorithmique, analystes en cyberdéfense prédictive.
- **Infrastructures numériques** : ingénieurs *data centers*, architectes *edge computing*, spécialistes curation de données, experts et ingénieurs en données (*data scientists* et *data engineers*).

Les **soft skills** constituent la valeur ajoutée la plus résistante à l'automatisation, renforçant plusieurs secteurs déjà existants comme les **métiers relationnels** à forte valeur humaine (soins/services à la personne, enseignement/mentorat), ainsi que ceux **manuels** ou **tactiles** (plombier, menuisier, électricien) où l'IA robotique est encore loin de rivaliser avec l'humain. Les gouvernements doivent mettre en place des mesures pour faciliter la reconversion professionnelle vers les **nouveaux secteurs** où l'IA permet une expansion, tout en investissant davantage dans ces **métiers existants** qui sortent **renforcés** de la révolution technologique.

→ **Solutions proposées** : (horizon 2026-28) : *Soutien extensif pour la formation à l'IA*

Coût global estimé : **Au moins €50 milliards** pour l'UE et ses États membres à l'horizon 2030 (**inclus dans la 1^{re} enveloppe financière des €100 milliards**), donc environ €10 milliards par an sur les cinq prochaines années, commençant dès 2026.

- **Crédit d'impôt augmentation humaine** : incitation fiscale pour les entreprises investissant dans l'IA d'assistance (aide au diagnostic), conditionnée à une formation certifiante de 100 % des employés concernés.
- **Fonds européen de transition numérique** : financement de « congés de conversion IA » de 6 à 12 mois pour les secteurs menacés, rémunérés à 90 % du salaire brut de référence pour le 1^{er} mois, et ensuite 70% pour les mois suivants.
- **Droit à la transparence algorithmique au travail** : notice d'impact sur les compétences obligatoires pour chaque nouvel outil d'IA déployé en entreprise.
- **Index de l'augmentation humaine** (analogue à l'index égalité F/H), notation des entreprises sur leur capacité à maintenir l'emploi tout en intégrant l'IA ; accès prioritaire aux marchés publics pour les mieux notées.
- **Micro-accréditations sectorielles obligatoires** : modules courts validés par les branches professionnelles dans les secteurs stratégiques (ex. : santé, industrie, finance/bancaire, assurance, énergie).
- **Passeport européen de compétences IA** : certification interoperable dans toute l'UE (basée sur DigComp 2.2 actualisé 2025-2026), facilitant la mobilité professionnelle.
- **Horizon 2028** : intégration de l'IA au cœur d'une **nouvelle « 5^e liberté » du marché unique européen** qui serait créée, dédiée à l'éducation, à la recherche et à l'innovation (en plus des quatre libertés historiques: biens, services, personnes et capitaux).¹

4 • Désinformation exacerbée par l'IA, avec risques systémiques pour les démocraties

L'IA démultiplie les **capacités de désinformation** : production instantanée de contenus convaincants en de multiples langues (textes, vidéos et audios), **deepfakes** ultraréalistes de dirigeants politiques, campagnes de **manipulation électorale automatisées** à grande échelle.

¹ Cette proposition est issue du rapport d'Enrico Letta (2024) intitulé : « [Much More than a Market](#) ».

Celles-ci sont souvent commanditées par des régimes autoritaires (Russie, Iran, Corée du Nord, etc.) pour diviser l'occident et favoriser la montée du populisme en minant les démocraties de l'intérieur. Les *bots* IA de 2026 sont capables d'engager des conversations de manière autonome avec un **microciblage psychologique**, rendant leur détection très difficile. En Europe, l'*AI Act* (articles 12 et 13) impose l'**auditabilité**, la transparence et la « journalisation automatique » des systèmes sur l'ensemble du cycle de vie du modèle, et propose des outils de **correction des biais** en continu (MLOps, AIF360, Fairlearn) et la technique du filigrane numérique (*Watermarking*). Cependant, le processus de simplification en cours des lois numériques par la Commission européenne (*Digital Omnibus*) pourrait affaiblir certaines de ces protections ; trouver des **compromis** s'avère donc essentiel.

→ **Solutions proposées : (horizon 2026-28) :**
Contenir la désinformation exacerbée par l'IA

- « **Bacs à sable de données** » (*Data Sandboxes*) sécurisées gérées par le Bureau européen de l'IA : les entreprises testent leurs algorithmes sur des données certifiées UE pour détecter les biais.
- **Passeport algorithmique** dynamique obligatoire pour chaque modèle d'IA à haut risque, contenant l'historique des données d'entraînement, les tests de biais successifs et un journal des modifications.
- **Signal d'alerte automatique inter-plateformes** (*Cross-Platform Signaling*) en temps réel en cas de détection d'un biais de manipulation.
- **Filigrane numérique renforcé** : intégration obligatoire des métadonnées (C2PA) au niveau matériel (capteurs) et logiciel pour tout contenu dans l'espace public européen.

5 • Impacts environnementaux de l'IA systémiques et multifformes

L'IA représente un **défi écologique systémique** : la demande en électricité des *data centers* optimisés par l'IA devrait plus que quadrupler d'ici 2030, à laquelle s'ajoute une **consommation massive d'eau** (4 à 6 milliards de m³/an

– l'équivalent de la moitié du Royaume-Uni) qui augmente fortement chaque année, ainsi que les **déchets électroniques** issus des puces spécialisées à courte durée de vie. La combinaison de ces facteurs génère des **pressions grandissantes sur les écosystèmes** locaux. L'UE a réagi en mettant en place des **obligations de transparence** (déclaration d'empreinte environnementale) dans l'*AI Act*, un futur label « **Green AI** » permettant de comparer les modèles d'IA en termes de leur efficacité énergétique, ainsi qu'un soutien financier dans la recherche pour les SLM et l'**éco-conception**. Paradoxalement, l'IA contribue aussi positivement à la transition écologique dans d'autres cas avec par exemple l'optimisation du réseau électrique via les **smart grids**, l'agriculture de précision ou le suivi satellitaire des émissions de GES. Néanmoins, le processus actuel de simplification (*Omnibus*) des normes climatiques risque d'affaiblir certaines de ces mesures ; un **juste équilibre** doit être trouvé pour ne pas sacrifier l'écologie sur l'autel de la compétitivité.

→ **Solutions proposées : (horizon 2026-28) :**
Diminuer et contrôler les impacts environnementaux de l'IA

- **Droit à l'audit environnemental**: accès en temps réel aux compteurs électriques et d'eau des *data centers*.
- **Taxe carbone numérique proportionnelle** pour les modèles dépassant un seuil de requêtes quotidiennes, avec TVA réduite pour les processeurs économes.
- **Standardisation des Small Language Models** (SLM) : obligation pour les fournisseurs de proposer une version allégée pour les tâches simples.
- **Quotas d'eau par Teraflop** : limite maximale de litres d'eau consommés par unité de calcul ; interdiction du refroidissement par eau potable en zones de stress hydrique.
- **Index de réparabilité/durabilité** des infrastructures d'IA : recyclage, réemploi obligatoire des équipements.

6 • Perte de souveraineté en raison de la dépendance de l'Europe aux technologies étrangères

L'Europe est largement **dépendante des technologies américaines** à trois niveaux : le Cloud (65 % contrôlé par AWS, Azure, Google Cloud), le matériel/processeurs (GPU Nvidia quasi-exclusifs), et les modèles de base intégrés dans les outils de travail quotidien (ChatGPT, Microsoft 365, Google Workspace) à grande échelle. Le **Cloud Act** américain et ses clauses extraterritoriales impactent directement la souveraineté numérique de l'UE car il neutralise les protections du RGPD. L'UE a répondu avec des mesures incluant les **labels de Cloud souverains** (EUCS ou *SecNumCloud*) et en investissant dans les infrastructures numériques européennes comme les **supercalculateurs** (réseau *EuroHPC*) via notamment le *AI Continent Action Plan* ou la stratégie *Data Union*. L'objectif est de subventionner la recherche et soutenir les **champions européens** comme *Mistral AI* (France), *DeepL* (Allemagne) ou *Silo AI* (Finlande). Malgré ces efforts, **65%** des organisations dans l'UE admettent ne pas pouvoir rester compétitives sans recourir partiellement à des fournisseurs non-européens. Le **rapport Draghi** (2024) identifie l'excès réglementaire comme constituant un frein à l'innovation de l'UE dans le secteur numérique, même si beaucoup d'économistes pointent d'autres facteurs comme la **fragmentation persistante du marché unique**.

→ **Solutions proposées** : Réduire la dépendance aux technologies étrangères et renforcer la souveraineté

2^e Enveloppe financière

Au minimum **€300 milliards supplémentaires** à mobiliser par l'UE et ses États membres à l'horizon 2035 sur l'ensemble des mesures pour réduire la dépendance aux technologies étrangères, avec l'objectif d'une résilience pragmatique et d'une souveraineté ciblée.

Financement : mélange budget UE à 50% (CFP/emprunt commun) + cofinancement des États membres, capitaux privés et nouvelles « ressources propres technologiques » (taxes IA/données) pour l'autre moitié.

I MÉTHODES D'ALLOCATIONS DE L'ENVELOPPE FINANCIÈRE

Fonds souverain numérique (Horizon 2026-35)

Au moins **€30 milliards/an sur une décennie** (inclus dans la 2^e enveloppe financière des **€300 milliards**) commençant dès 2026, pour financer les infrastructures critiques numériques souveraines (Cloud, supercalculateurs, data centers, fabriques IA) via des fonds publics, partenariats public-privé ou la « taxe robot ».

I MESURES À L'HORIZON 2026-2028

- Adoption du **Digital Omnibus** (simplification de l'*AI Act*, RGPD, DMA, DSA, etc.) prévue d'ici l'été 2026, à condition que la simplification ne se fasse pas au détriment des droits fondamentaux.
- **Politique de commande publique stratégique** : préférence européenne dans les marchés publics pour les solutions d'IA européennes.
- **Stratégie Data Union** : mutualisation des données dans les secteurs clés (santé, industrie, énergie).
- **Chips Act accéléré** : relocalisation de la production de puces avancées (GPU, accélérateurs IA) en Europe.
- **Union des marchés de capitaux** (horizon 2028) pour permettre aux « licornes » européennes de l'IA de lever des fonds sans être rachetées par des groupes étrangers.

III. Recommandations opérationnelles pour les entreprises et décideurs publics

1 • L'Open Source comme levier de souveraineté

L'Open Source est un mode de développement et distribution de logiciels dont le code source est rendu public, contrairement aux logiciels « propriétaires » fermés. L'Open Source apporte l'**autonomie stratégique** (pas de coupure d'accès possible), la **transparence** (absence de « portes dérobées » souvent utilisées pour l'espionnage), et permet de sortir de « l'enfermement propriétaire ». L'Open Source offre la possibilité à l'UE de concurrencer les USA avec beaucoup moins de moyens financiers ; des **champions européens** ont pu émerger via cette voie (Mistral AI, Odoo, Grafana Labs, SUSE). Mais l'Open Source seul est **insuffisant** si le matériel (*hardware*) demeure américain, ce qui est majoritairement le cas, d'autant plus que les GAFAM sont également des leaders dans ce domaine. Ainsi, l'Europe doit articuler l'Open Source avec le renforcement des **infrastructures numériques souveraines**, l'intégration du marché unique et la stratégie *Data Union*.

→ **Solutions proposées (horizon 2026-35) :**
Environ **€4 milliards/an** supplémentaires fléchés vers l'Open Source – focus sur l'IA et le Cloud (**inclus dans la 2^e enveloppe financière des €30 milliards/an**) + 1 % des budgets informatiques publics dédiés à la maintenance de l'Open Source existant.

2 • Réguler l'IA : Limiter/interdire certains types d'IA, avec la définition de critères de « dangerosité »

Selon l'AI Act, les technologies sont réglementées en fonction du degré de risque qu'elles représentent, avec **quatre catégories** : risque d'IA **inacceptable** (interdit : notation sociale, manipulation subliminale), risque **élevé** (réglementation stricte : justice, emploi, santé), risque **limité** (transparence obligatoire : *chatbots*) et risque **minimal** (codes de conduite volontaires). Certaines lacunes subsistent dans la loi : **manque de souplesse** et processus de révision trop lent face à l'évolution rapide du secteur, une attention insuffisante aux **impacts sociétaux** collectifs de l'IA (désinformation/manipulation démocratique), et un recours excessif à l'**auto-évaluation** par les entreprises (surtout avec le *Digital Omnibus*) affaiblissant la mise en conformité.

→ **Solutions proposées (horizon 2028) :**
Comblar les lacunes de l'AI Act européen

- **Mécanisme de mise à jour rapide des catégories de risque** : délégation au Comité européen de l'IA pour des révisions annuelles sans procédure législative complète.
- **Élargissement des critères de la catégorie « risque inacceptable »**, notamment pour la biométrie et la reconnaissance faciale.

- **Intégration des impacts sociétaux**, avec une section société civile au sein du Comité européen de l'IA.
- **Renforcement des évaluations de conformité tierce partie** : réduction de l'auto-évaluation par les entreprises, focalisation sur les catégories « risque élevé » et « inacceptable ».

3 • Sécurité pour les infrastructures numériques de l'IA, comparaison avec le secteur nucléaire

Les **infrastructures numériques** (*data centers*, fabrique d'IA) sont devenues vitales et hautement stratégiques, introduisant des **risques de sécurité** inédits, tant physiques qu'en termes de cybersécurité. Bien qu'elles ne fonctionnent pas de la même manière, des comparaisons peuvent être établies avec la sécurisation des **centrales nucléaires**. En effet, le nucléaire bénéficie de plusieurs décennies d'expérience dans le domaine sécuritaire, et certaines mesures sont **transposables** sur les infrastructures numériques, avec des adaptations. Même si les *data centers* bénéficient déjà d'une sécurité importante, celle-ci doit être augmentée.

→ Solutions proposées (horizon 2028) : Renforcer la sécurité des *data centers*

- **Sur le modèle des centrales nucléaires**, duplication et redondance renforcée des systèmes, alimentation électrique de secours plus poussée avec système de « fonctionnement en îlot » intégral (autonomie électrique totale).
- **Multiplication des contrôles et fouilles**, restrictions hermétiques aux points d'accès physiques.

4 • Bâtir une gouvernance mondiale de l'IA

Les risques systémiques posés par l'IA soulignent le besoin de bâtir une **gouvernance mondiale**. Même si de nombreux parallèles ont été établis avec le **nucléaire**, la dualité énergie-militaire est trop contraignante. Le régime mondial sur les **armes et les produits chimiques** de l'OIAC (Organisation pour l'Interdiction des Armes Chimiques) offre un cadre de comparaison plus pertinent pour la gouvernance mondiale de l'IA, en raison de sa nature multi-sectorielle complexe, analogue à celle de l'IA. L'ONU franchit des étapes importantes pour la gouvernance mondiale de l'IA avec le **Pacte Numérique Mondial** (2024), la première **résolution de l'Assemblée générale** sur le sujet (2024) et la création d'un **panel scientifique** international sur l'IA (2026). Néanmoins, l'instabilité des relations internationales rend improbable tout cadre mondial contraignant à court terme. La gouvernance mondiale de l'IA passera par des voies plus pragmatiques :

→ Solutions proposées (horizon 2026-28) : Création de mécanismes **multiniveaux**

- **Normes ISO** (Organisation internationale de normalisation) : points de repère mondiaux établissant des lignes directrices volontaires mais rigoureuses. Renforcement/intégration systématique dans l'*AI Act* – gouvernance (ISO/IEC 42001), gestion des risques (ISO/IEC 23894), qualité des données (ISO/IEC 5259), éthique et biais (ISO/IEC TR 24027).
- **Coalitions inter-régionales** : L'UE doit exiger de ses principaux partenaires commerciaux l'adoption des standards européens d'usage responsable de l'IA, en s'appuyant sur sa position de plus grand marché unique mondial et sur la portée extraterritoriale de l'*AI Act*. Développement de **cadres bilatéraux et multilatéraux sur l'éthique de l'IA**, sur le modèle de ce qui a été fait par l'UE avec le RGPD.

IV. Scénarios prospectifs à moyen terme (horizon 2035)

1 • Scénario optimiste : l'IA améliore l'humanité

L'**accélération technologique** de l'IA est maîtrisée, grâce notamment aux données de synthèse hybrides et aux avancées dans les énergies propres. L'IA crée plus d'emplois qu'elle n'en détruit grâce à une formation massive et efficace, permettant l'**amélioration humaine** plutôt que son remplacement via la **fusion humain-machine**. Cela permet à l'économie mondiale de connaître une nouvelle ère de **forte croissance**, qui bénéficie à l'ensemble de la population en **réduisant les inégalités** entre les nations et au sein des pays. L'IA devient un outil de **progrès scientifique** et **médical** majeur emmenant la guérison de nombreuses maladies longtemps incurables (cancer, Alzheimer), ce qui allonge la qualité et l'espérance de vie de l'humanité. Les **démocraties sont protégées** par des systèmes de détection de la désinformation performants via des partenariats public-privé robustes impliquant des **règlementations** strictes dans de nombreux pays à travers le monde, à l'instar de l'Europe. Cela inclut la mise en place d'un cadre de **gouvernance mondiale contraignant** avec une charte sur l'usage responsable de l'IA ratifiée sous l'égide de l'ONU, permettant de garder le contrôle sur la technologie utilisée pour servir le bien-être de l'humanité. Enfin, l'IA permet des **percées technologiques** majeures accélérant le déploiement des **énergies renouvelables** et **bas-carbone**, notamment un déploiement pré-

cocé de la **fusion nucléaire** ; les objectifs de l'**Accord de Paris** deviennent atteignables et la planète est ainsi préservée.

2 • Scénario pessimiste : remplacement et chaos

La course à l'efficacité et la concurrence mondialisée poussent les entreprises à **remplacer en très grand nombre** les humains par l'IA, non seulement pour les métiers « à col blanc » les plus vulnérables, mais également dans des secteurs jusqu'alors protégés (métiers manuels, services relationnels). Un **chômage de masse** provoque de graves **troubles sociaux**, les gouvernements n'ayant pas les ressources pour financer les subventions publiques significatives nécessaires, augmentant la précarité et les inégalités au niveau mondial. Un certain nombre de régimes démocratiques basculant ainsi vers l'**illibéralisme**, notamment en raison de l'échec du développement de cadres de gouvernance efficaces pour lutter contre la **désinformation**, minant les processus électoraux. La consommation énergétique liée à l'IA explose, provoquant des pannes électriques d'envergure et une **accélération des effets du changement climatique** qui impacte les pays du monde entier plus tôt que prévu, renforçant le **chaos émergent**. Des **conflits armés** apparaissent dans plusieurs pays en développement déjà fragilisés, et la collaboration internationale se démantèle, entraînant

l'échec de la mise en place d'une **gouvernance mondiale**. Ainsi, nous assistons dans ce scénario à une **perte de contrôle** sur l'IA, provoquant des crises sans précédent qui compromettent l'avenir de l'humanité.

3 • Scénario intermédiaire (le plus probable) : progrès contrastés

Ce scénario, jugé le plus **réaliste**, combine effets positifs et négatifs. Un **ralentissement relatif** du développement technologique de l'IA (manque de données, contraintes énergétiques) atténue les extrêmes. La **productivité** et la **croissance** économique augmentent dans de nombreux secteurs, mais de façon inégalement répartie (à l'instar du modèle américain). L'IA **transforme les métiers** en profondeur sans les supprimer massivement, tout en renforçant les **inégalités sociales**, notamment pour les jeunes diplômés et les travailleurs seniors peu adaptables. Sur le plan écologique, l'IA fait progresser les renouvelables et les **technologies bas-carbone**, mais ces avancées sont partiellement neutralisées par l'augmentation massive de la **consommation énergétique**, avec des effets ambivalents sur le climat et la mise en œuvre de l'Accord de Paris. Certaines démocraties (UE, Canada, Japon) parviennent à contenir la **désinformation** avec des cadres réglementaires efficaces, mais la majorité des pays reste vulnérable, y compris les États-Unis et les pays en développement où les polarisations internes se renforcent. Sous le leadership européen, des cadres bilatéraux, des coalitions inter-régionales et les normes ISO parviennent à émerger, mais **sans gouvernance mondiale contraignante** en raison de blocages à l'ONU avec la rivalité sino-américaine. Selon ce scénario intermédiaire, l'IA **amplifie plusieurs tendances déjà visibles en 2026** qui risquent de s'accroître à l'horizon 2035.

Conclusion

La révolution de l'IA va profondément remodeler l'économie, le travail, l'environnement, les démocraties et la société à l'horizon 2035. La question centrale sera de **garder l'humain dans la boucle**, en régulant et en définissant des limites claires pour garder le contrôle. Sur le long terme, à l'horizon 2050-2100, l'arrivée potentielle de l'**Intelligence Artificielle Générale** (AGI) et de la **singularité technologique** - où l'IA s'améliore de manière autonome et dépasse l'intelligence humaine - pourrait marquer un **basculement civilisationnel inédit** : réduction du travail et avènement d'une société basée principalement sur les loisirs, redistribution massive des richesses sans salaire, fusion humain-machine entraînant une marche forcée vers l'étape suivante de l'évolution, où l'IA deviendrait l'infrastructure même de la réalité humaine. Le plus grand défi serait de s'assurer que les objectifs d'une super-intelligence **restent alignés avec les valeurs humaines**. Comme l'affirmait Stephen Hawking : « Réussir à créer l'IA serait le plus grand événement de l'histoire de l'humanité. Malheureusement, ce pourrait aussi être le dernier, à moins que nous n'apprenions comment éviter les risques. » L'enjeu fondamental reste donc **d'orienter cette révolution technologique au service du bien-être humain**, en soutenant à la fois développement économique, protection sociale et environnementale, souveraineté numérique et renforcement des valeurs démocratiques.

Tableau récapitulatif des horizons, investissements et seuils pour les principales solutions proposées

Horizon	Domaine	Principales mesures proposées	Investissements / Financements	Seuils & objectifs chiffrés
2026–2030	Protection de l'emploi	- Garantie d'employabilité IA (<i>Upskilling Obligation</i>) – responsabilité légale de l'employeur. - Comités d'éthique et de supervision numérique obligatoires (>50 salariés), audit annuel. - Clause de « non-substitution technologique » dans les conventions collectives. - Étendre à l'ensemble des salariés concernés les restrictions relatives à la gestion algorithmique du travail prévues par la directive de 2024. - Dividende numérique : fonds de transition sectoriel national (reconversions longues, allocations chômage). - Réduction du temps de travail sans baisse de salaire, financée par gains de productivité IA. - Crédit d'impôt « augmentation humaine » conditionné à formation certifiante de 100 % des salariés concernés. - Fonds européen de transition numérique : congés de conversion IA de 6-12 mois payés à 70-90% du salaire. - Index de l'augmentation humaine public (accès prioritaire aux marchés publics pour les mieux notés). - Reconversions prioritaires : santé/biotech, cybersécurité/éthique IA, <i>data centers</i> & <i>edge computing</i>. - Micro-accréditations sectorielles obligatoires et passeport européen de compétences IA. - L'IA au cœur d'une nouvelle 5^e liberté du marché unique de l'UE sur l'éducation, la recherche et l'innovation.	1^{er} enveloppe financière ≥ 100 Md€ à mobiliser d'ici 2030 (probablement plus par la suite), donc 20 Md€/an sur 5 ans, commençant dès 2026. - Financement : 50% budget UE (CFP/emprunt commun) + 50% États membres, capitaux privés, taxes IA/données.	Gains de productivité IA estimés entre 400-600 Md€ d'ici 2030 (financement des mesures sociales).
2026–2030	Soutenir la formation à l'IA	- Index de l'augmentation humaine public (accès prioritaire aux marchés publics pour les mieux notés). - Reconversions prioritaires : santé/biotech, cybersécurité/éthique IA, <i>data centers</i> & <i>edge computing</i>. - Micro-accréditations sectorielles obligatoires et passeport européen de compétences IA. - L'IA au cœur d'une nouvelle 5^e liberté du marché unique de l'UE sur l'éducation, la recherche et l'innovation.	≥ 50 Md€ pour la formation IA (UE + États membres), dont 10 Md€/an sur 5 ans (inclus dans les 100 Md€ d'ici 2030), commençant dès 2026.	Droit à la transparence algorithmique: notice d'impact sur les compétences pour chaque nouvel outil d'IA déployé en entreprise.
2026–2028	Souveraineté numérique	- Digital Omnibus (simplification AI Act / RGPD / DMA / DSA) – adoption d'ici l'été 2026 sans dérogation aux droits fondamentaux. - Préférence européenne dans les marchés publics pour solutions IA européennes. - Stratégie Data Union : mutualisation des données (santé, industrie, énergie). - Chips Act accéléré : relocalisation des puces GPU/accélérateurs IA en Europe. - Horizon 2028 : Union des marchés de capitaux pour préserver les licornes IA européennes.	2^e enveloppe financière ≥ 300 Md€ (UE + États membres) à mobiliser d'ici 2035, dont 30 Md€/an sur 10 ans, commençant dès 2026 pour développer les infrastructures numériques souveraines. ≥ 4 Md€/an sur 10 ans fléchés vers l'Open Source IA & Cloud (inclus dans les 30 Md€/an d'ici 2035).	1% des budgets informatiques publics dédiés à la maintenance de l'Open Source existant.
2026–2028	Désinformation & démocratie	- Bacs à sable de données (<i>Data Sandboxes</i>) sécurisés gérés par le Bureau européen de l'IA. - Passeport algorithmique dynamique obligatoire pour tout modèle à haut risque. - Signal d'alerte automatique inter-plateformes (<i>Cross-Platform Signaling</i>) en temps réel. - Filiigrane numérique renforcé : métadonnées C2PA obligatoires au niveau matériel et logiciel.	- Inclus dans les ≥ 30 Md€/an sur 10 ans pour développer les infrastructures numériques souveraines. - Inclus dans les ≥ 4 Md€/an sur 10 ans fléchés vers l'Open Source IA & Cloud.	Obligations applicables pour tout contenu diffusé dans l'espace public européen.
2026–2028	Environnement	- Droit à l'audit environnemental tiers : accès temps réel aux compteurs des data centers. - Taxe carbone numérique proportionnelle (TVA réduite pour processeurs économes). - Standardisation des <i>Small Language Models</i> (SLM) – version allégée obligatoire pour tâches simples. - Quotas d'eau par Téraflop : interdiction du refroidissement par eau potable en zones de stress hydrique. - Index de réparabilité et de durabilité des infrastructures IA.	–	Seuil de déclenchement: volume de requêtes quotidiennes dépassant un plafond défini.
2026–2028	Gouvernance mondiale de l'IA	- Normes ISO pleinement intégrées à l'AI Act : gouvernance (ISO/IEC 42001), risques (ISO/IEC 23894), éthique (ISO/IEC TR 24027). - Coalitions inter-régionales sur l'éthique IA avec partenaires commerciaux de l'UE (modèle RGPD).	–	Modèle inspiré de l'OIAAC (traité sur les armes chimiques) pour sa nature multisectorielle.
2028	AI Act renforcé & sécurité data centers	- Révisions annuelles des catégories de risque sans procédure législative complète (délégation au Comité européen de l'IA). - Élargissement de la catégorie « risque inacceptable » (biométrie, reconnaissance faciale). - Intégration des impacts sociétaux (fake news, manipulation démocratique). - Réduction de l'auto-évaluation pour les catégories « risque élevé » et « inacceptable ». - Sécurisation renforcée des <i>data centers</i> à l'instar des centrales nucléaires : générateurs de secours & fonctionnement en îlot (autonomie électrique).	–	Section société civile obligatoire au sein du Comité européen de l'IA.
2028–2030	Propositions de financement « Taxe robot »	- Taxe sur usage/équivalence salariale : cotisations sociales équivalentes à un employé humain remplacé. - Surtaxe IS pour entreprises dont le ratio CA/effectif dépasse un seuil défini. - Suppression des avantages fiscaux (amortissement) liés aux technologies de substitution au travail. - Taxe sur les données (<i>big data</i>) : redevance proportionnelle à l'usage de données pour l'entraînement.	Alimentation du fonds de transition sectoriel national (dividende numérique)	Seuil de déclenchement de la surtaxe IS : ratio CA/effectif dépassant un seuil à définir.
2030–2035	Fonds souverains technologiques	- Création de fonds souverains technologiques. - Prises de participations publiques dans les entreprises d'IA utilisant des données publiques.	- Inclus dans les ≥ 30 Md€/an sur 10 ans pour développer les infrastructures numériques souveraines. - Inclus dans les ≥ 4 Md€/an sur 10 ans fléchés vers l'Open Source IA & Cloud.	Modèle inspiré du Fonds pétrolier souverain norvégien.
2035 et au-delà	Réformes profondes des systèmes sociaux	- Semaine de quatre jours (voire moins) pour des cas définis – introduction progressive. - Réforme profonde des systèmes de protection sociale avec possible introduction d'un « Revenu Universel de Base » (RUB) : universel, individuel, inconditionnel, régulier.	À l'horizon 2035 et au-delà, cela nécessitera une augmentation significative par rapport à la 1 ^{re} enveloppe financière de ≥ 100 Md€ pour 2030.	Rupture paradigmatique avec les systèmes sociaux actuels.

Étude CRiP /
Institut Jacques Delors

Révolution des technologies de l'intelligence artificielle à l'horizon 2035

L'humain dans la boucle :
réguler et définir des limites
pour garder le contrôle

Introduction

Cette étude a pour objectif d'analyser l'impact de la révolution des technologies de l'**intelligence artificielle (IA)** à l'**horizon 2035**, et comment **réguler et définir des limites pour garder un certain contrôle et l'humain dans la boucle**. Même si les technologies de l'IA existent depuis de nombreuses années, elles furent pendant longtemps cantonnées au secteur de l'informatique ou pour des usages spécifiques.² L'arrivée sur le marché de **ChatGPT** version 3.5 en novembre 2022 eut un effet de tsunami et créa un **changement de paradigme**, car l'IA est presque instantanément devenue un **produit de consommation de masse**, au lieu d'un outil informatique spécialisé. Depuis, **ChatGPT** bat tous les records en devenant le logiciel affichant la croissance la plus exponentielle de tous les temps et engendrant une **course effrénée** dans le secteur de la tech qui continue aujourd'hui.³ En effet, toutes les grandes entreprises du secteur ont rapidement développé leurs propres modèles d'IA ou plus spécifiquement « grand modèles de langage » (*large language models* – LLM) : *Gemini* de Google, *Claude* d'Anthropic, *Grok* de xAI, *Phi* de Microsoft, *Granite* d'IBM ou encore *Llama* de Meta. Cette course, bien qu'initialement démarrée aux États-Unis, s'est rapidement propagée au niveau mondial, notamment avec *DeepSeek* en Chine, *Sarvam* en Inde, *Yandex* en Russie et *Mistral* en France. L'IA est devenue la **technologie à l'expansion la plus rapide de l'histoire**, avec plus de 1.5

milliards de personnes par mois en moyenne dans le monde se servant de l'IA, et environ 700 à 800 millions de personnes se servant de *ChatGPT* chaque semaine, soit plus de 10% de la population mondiale. Cela représente environ 25% de la population active dans les pays développés qui s'en sert au quotidien.⁴ En raison de cette diffusion si rapide et profonde, il n'est pas étonnant que l'IA ait contribué à **bouleverser les sociétés** dans les pays du monde entier au cours des quatre dernières années. Cela inclut des impacts directs ou indirects dans la plupart des secteurs et à presque tous les niveaux, à la fois pour les nations développées et celles en développement.

Cette étude ayant été commanditée par le **CRiP**⁵, elle qui regroupe les principales entreprises françaises des secteurs de l'IT, des technologies et des infrastructures, elle se concentrera sur les impacts de l'IA au sein des pays développés, principalement en **France** et en **Europe**, même si de nombreuses comparaisons seront établies avec d'autres pays, notamment les États-Unis, la Chine, le Japon et d'autres. L'objectif sera d'analyser les impacts à court terme à l'**horizon 2026-28**, ainsi que d'établir une comparaison avec le moyen terme à l'**horizon 2035** avec différents scénarios possibles, notamment sur le plan économique, politique, environnemental et sociétal. Des **recommandations** seront prodiguées aux entreprises et responsables

² Cela inclut par exemple les moteurs de recommandation pour les réseaux sociaux ou le *streaming* (Tik Tok, Netflix, Youtube), ainsi que pour les assistants virtuels (Siri, Alexa ou Google Assistant).

³ Hu (2023).

⁴ Liu et Wang (2024), Bellan (2025).

⁵ Le CRiP réunit 16 000 Responsables Infrastructures, Technologies et Production IT, ainsi que leurs collaborateurs, issus de 355 grandes entreprises, ETI et administrations.

politiques pour assurer le développement d'une **IA responsable** et éthique permettant de tirer les bénéfices de cette révolution technologique, tout en protégeant les entreprises, institutions politiques et la société civile des risques associés. En outre, cette étude s'appuiera sur une **méthodologie de type mixte**, combinant à la fois des données quantitatives et des informations de type qualitatives, permettant de développer une analyse rigoureuse appuyée par un cadre méthodologique riche et précis. Concernant les données quantitatives, un **sondage** a été mené auprès d'un large panel de participants issus de divers secteurs non seulement en France, mais également en Europe plus généralement et à l'international. Cela inclut un questionnaire contenant une trentaine de questions et plusieurs types de réponses possibles, notamment des degrés d'accord allant de 0 à 5, permettant le traitement de données quantitatives à grande échelle. Le sondage a été mené sur plusieurs mois avec l'obtention d'une large gamme de données avec **460 réponses au total**, les questions du sondage couvrant l'ensemble des sujets et thématiques abordés dans cette étude. Une analyse détaillée de l'intégralité des résultats issus du sondage est prodiguée dans la **1^{re} Annexe** à la fin de l'étude.

Concernant les données de type qualitative, une **douzaine d'entretiens** ont été menés avec des chefs d'entreprises et des employés du secteur de la tech, ainsi qu'avec des universitaires ou experts membres de la société civile en France, en Europe et aux États-Unis. Les questions du sondage furent adaptées lors des entretiens pour permettre l'obtention d'informations qualitatives avec davantage de nuance dans les réponses, se focalisant plutôt sur les opinions et les prévisions individuelles. La méthodologie fut basée sur les entretiens de type « **semi-structurés** », avec un noyau de questions communes pour toutes les personnes interrogées lors de la première partie, permettant ensuite d'adapter les questions selon le profil des individus pendant la deuxième partie de l'entretien.⁶ La comparaison de données quantitatives issues du sondage, ainsi que des informations de type qualitatives

grâce aux entretiens, renforce la méthodologie et permet d'aboutir à des résultats d'analyse très concrets. De plus, cette étude se base sur une **bibliographie** solide incluant des sources académiques en français et en anglais issues des meilleures revues scientifiques à comité de lecture ou presses universitaires. En complément, l'étude mobilise aussi des sources « non-académiques » comme des rapports ou études de gouvernements en France ou dans d'autres pays, d'institutions internationales (UE, OCDE, ONU), ainsi que diverses publications de *think tank* ou d'ONG. Ces sources servent à développer et renforcer l'analyse, permettant d'aborder de manière précise les enjeux d'actualités.

Cette étude sera divisée en plusieurs parties et sous-parties, qui se suivent et se complètent de manière à renforcer la cohésion d'ensemble :

- La première partie contiendra un état des lieux et des **scénarios prospectifs à court terme (horizon 2026-28)** pour les entreprises et décideurs publics sur la cohabitation humain-IA ;
- La deuxième partie se concentrera sur les **risques systémiques** liés à l'IA et les mécanismes de régulation applicables ;
- La troisième partie proposera des **recommandations opérationnelles** pour les entreprises et décideurs publics afin de tirer avantage de la révolution de l'IA tout en minimisant les risques, ainsi que pour créer un cadre politique amenant à une IA éthique et responsable ;
- La quatrième et dernière partie présentera une déclinaison des éléments précédents dans **trois scénarios prospectifs contrastés à moyen terme à l'horizon 2035** : un premier scénario optimiste contrasté à un second scénario pessimiste, et enfin un troisième scénario intermédiaire plus réaliste.
- La conclusion ouvrira la réflexion sur l'impact de l'IA à long terme à **l'horizon 2050-2100** en posant la question : comment préparer la société à l'arrivée potentielle de l'IA générale et de la singularité technologique ?

⁶ George et Bennett (2005).

I. État des lieux et scénarios prospectifs à court terme (horizon 2026-28) sur la cohabitation humain ↔ IA

1 • Sécurité et cohabitation humain-IA à l'ère des IA agentiques :

Le sujet de la **sécurité** et de la **cohabitation** humain-IA à l'horizon 2026-28⁷ est complexe et multi-dimensionnel. En termes d'architectures d'interface, d'ergonomie et de modèles d'interaction, la priorité pour le secteur de la tech a été de favoriser l'**accessibilité**, la facilité d'usage et la performance, parfois aux dépens de la maîtrise humaine. En effet, la décision des grandes entreprises du secteur de mettre à disposition des versions gratuites pour la plupart de leurs modèles (tout en conservant des versions payantes plus performantes), a permis d'élargir considérablement l'accès à la technologie. De plus, le format utilisé pour diffuser l'IA, accessible non seulement sur les ordinateurs mais également sur la plupart des smartphones et même sur certains modèles de montres, a fait rentrer l'IA dans le quotidien intime de centaines de millions de personnes dans le monde. Cela est d'autant plus le cas que le format des LLM, qui requiert simplement de taper une ins-

truction (« *prompt* ») sur le clavier, a permis d'apporter une grande facilité d'usage. Au cours des dernières années, l'accent a été mis sur la rapidité dans le débit de réponse des LLM, aussi appelé la « **latence décisionnelle** »⁸, afin de renforcer l'accessibilité, la facilité d'usage et la performance. Cela peut également servir pour certains cas de figure comme les échanges financiers qui nécessitent des actions très rapides pour éviter les risques de pertes de marché, ou même pour la conduite de véhicules où un retard de quelques microsecondes peut entraîner un accident sur la route. Néanmoins, la rapidité extrême des nouveaux modèles d'IA a également pour conséquence de compromettre la **maîtrise humaine**, car la technologie évolue maintenant à une vitesse qui dépasse largement la capacité de réaction des individus.

En outre, le sujet des mécanismes de sécurité est également devenu de plus en plus important au cours des derniers mois. En effet, un certain nombre de témoignages d'informaticiens travaillant pour les grandes entreprises de la tech sur le développement de la prochaine génération d'IA, révèlent un début d'interroga-

⁷ Dans le cadre de cette étude, le « court terme » est défini comme incluant les trois prochaines années allant de 2026 à 2028. Le « moyen terme » est défini sur les 10 prochaines années, s'étendant jusqu'à 2035. Cette répartition chronologique résulte du choix de l'auteur, en lien avec l'évolution possible de la technologie de l'IA au cours des prochaines années.

⁸ La « latence décisionnelle » désigne le délai entre le moment où un système d'IA reçoit une entrée, et celui où il génère la sortie correspondante.

tion sur des comportements inexplicables et potentiellement inquiétants. Par exemple, des employés ou cadres travaillant pour les entreprises américaines de la tech ont affirmé que certains nouveaux modèles d'IA avaient **désobéi à leurs instructions** en effaçant partiellement ou complètement des messages ou échanges d'emails, sans explication apparente.⁹ Encore plus inquiétant, certaines IA avancées auraient même développé des mécanismes qui peuvent être comparés à de l'auto-préservation, en tentant par exemple de désactiver une nouvelle IA « rivale » en cours de développement, qui pourrait la remplacer car potentiellement plus performante.¹⁰ Aucune preuve n'existe à ce jour pour affirmer que l'IA aurait volontairement commis ces actions, et le consensus scientifique aujourd'hui est que nous sommes encore loin d'atteindre un état de conscience de l'IA comparable à la conscience humaine.¹¹ Néanmoins, ce type de comportement interroge sur la sûreté des nouveaux modèles, soulignant clairement la nécessité d'intégrer des mécanismes de sécurité performants et une **supervision humaine effective sur l'ensemble du cycle de vie**. Cela inclut par exemple la possibilité de contacter directement et rapidement l'entreprise qui développe les modèles d'IA pour rapporter des comportements anormaux, avec un service client et une prise en charge immédiate. En outre, cela inclut la nécessité de mettre en place des « **kill switch** »¹² sur les nouveaux modèles d'IA avancés, afin d'être sûr de pouvoir les désactiver rapidement et instantanément dans des cas extrêmes, et ainsi en préserver la maîtrise humaine.¹³

Ces sujets de la sécurité et de la cohabitation humain-IA sont d'autant plus d'actualité en raison

de l'arrivée sur le marché mondial en 2026 des nouveaux modèles d'IA de type « **agentique** », qui pourraient constituer un **changement de donne**. Ceux-ci peuvent être définis comme une forme d'IA qui prend des décisions de manière autonome, agit et s'optimise en temps réel, contrairement aux systèmes d'IA traditionnels. L'IA agentique se base sur quatre capacités essentielles qui la différencient par rapport aux IA « classiques » :

- **L'autonomie** : elle n'a pas besoin d'instructions étape par étape dans la mesure où il faut lui donner un objectif, et c'est l'IA agentique qui définira elle-même les étapes.
- **Le raisonnement et la planification** : l'IA agentique est capable de décomposer une tâche complexe en sous-tâches logiques.
- **L'utilisation d'outils** : elle est capable d'interagir avec le monde réel en utilisant par exemple les navigateurs web pour envoyer des emails, exécuter du code informatique ou même consulter d'elle-même une base de données.
- **La capacité d'auto-correction** : si elle échoue, l'IA agentique va analyser l'erreur et tenter une autre approche.

Pour ces raisons, l'arrivée des IA « agentique » a été présentée comme constituant une révolution car elle transforme la technologie en véritables **systèmes d'exploitation cognitifs**. Il ne s'agit plus dans ce cas de simplement converser avec l'IA, mais de lui déléguer des pans de travail entiers comme la recherche académique, le développement de logiciels automatisés, ou la gestion du service client. Comme analysé ci-dessous, les conséquences pour le travail, les métiers et les compétences sont potentiellement profondes.¹⁴

⁹ Ce phénomène est appelé « désalignement agentique » (*agentic misalignment*), avec plusieurs exemples très médiatisés au cours des dernières années, notamment avec l'outil d'IA agentique *OpenClaw* de Meta en 2026. Voir : Aissa (2026).

¹⁰ Certaines questions légitimes peuvent être soulevées par ces mécanismes « d'auto-préservation », même si les tests aujourd'hui indiquent que l'IA n'est pas « consciente », tout du moins pas au sens humain du terme.

¹¹ Porębski et Figura (2025).

¹² Un « kill switch » est une expression anglaise signifiant « bouton d'arrêt d'urgence », qui est utilisé pour déconnecter intégralement un service par mesure de sécurité, dans ce contexte une IA qui désobéirait aux ordres par exemple.

¹³ Comme indiqué dans la 1^{re} Annexe, un point de consensus dans le sondage est la nécessité de renforcer les mécanismes de sécurité pour l'IA, notamment en incluant des « *kill switches* » comme mécanismes d'urgence.

¹⁴ Comme expliqué dans la 1^{re} Annexe, l'IA agentique est perçue comme un changement de paradigme par les répondants au sondage, renforçant l'urgence de mécanismes de contrôle.

2 • Impacts de l'IA sur le travail, les métiers et les compétences à l'horizon 2026-28

L'impact de la révolution de l'IA sur le **travail**, les **métiers** et les **compétences** à l'**horizon 2028** est probablement l'un des sujets les plus controversés et débattus en ce moment, pas seulement par la communauté académique mais également par les entreprises et les décideurs politiques. En effet, le sujet est potentiellement explosif, puisqu'un certain nombre de commentateurs prédisent depuis déjà plusieurs années l'arrivée imminente d'un chômage de masse et d'un remplacement de l'humain par la machine, avec l'avènement à terme d'une société « post-travail ». Néanmoins et en dépit de ces prédictions pessimistes, force est de constater qu'au jour d'aujourd'hui, l'IA n'a pas ou tout du moins **pas encore bouleversé les chiffres de l'emploi**.¹⁵ En effet, les taux de chômage dans les grandes économies à travers le monde n'ont pour l'instant été que peu impacté par l'IA et ont même légèrement baissé depuis 2022 dans certains cas, ou sinon légèrement augmenté dans d'autres cas. Dans l'UE par exemple, le taux de chômage a évolué de 6,2% en 2022 (moyenne annuelle – MA) à 6% en janvier 2026¹⁶, et de 7,4% en 2022 (MA) à 7,7% en janvier 2026 pour la France.¹⁷ Aux États-Unis, celui-ci était de 3,6% en 2022 (MA) et a atteint 4,3% en janvier 2026¹⁸, alors qu'en Chine il était de 5,5% en 2022 (MA), diminuant légèrement à 5,2% en janvier 2026.¹⁹

La plupart des experts sont d'accords sur le fait que ces changements relatifs qui ont pu

affecter les taux de chômage au cours des dernières années sont liés principalement soit à la reprise économique post-Covid, soit à des chocs externes (guerres en Ukraine et au Moyen-Orient, guerre commerciale avec les USA sous Trump, etc.).²⁰ Il y a certes eu des pertes d'emplois en raison de l'IA depuis 2022 qui ont eu un retentissement dans les médias comme ce fut le cas au moment de l'annonce d'Amazon en janvier 2026 du licenciement d'environ 16 000 employés, suivi par Oracle avec 10 000 licenciements en avril, ainsi que Meta avec 8000 licenciements en mai (10% de ses effectifs).²¹ Même si cela suscite des craintes d'effet domino et que d'autres entreprises pourraient suivre le même modèle, le fait est que ce type de licenciement reste pour l'instant à la marge et n'a **pas impacté les taux de chômage dans le monde de manière significative**, y compris en Chine et aux États-Unis où la révolution de l'IA a été plus profonde qu'en Europe.²² Néanmoins, certains experts prédisent que l'arrivée des **IA de type agentique** pourrait constituer un **changement de donne**, avec le risque d'un « tsunami » sur le marché du travail à l'horizon 2028,²³ surtout pour les métiers dits « à col blanc » incluant les travailleurs de bureau et en particulier les cadres, qui représentent la majorité des emplois dans le secteur tertiaire au sein des pays occidentaux. Ainsi, plus de 300 millions de postes dans le monde seraient vulnérables à l'automatisation au cours des prochaines années, concernant des tâches répétitives que l'IA peut faire plus rapidement et efficacement, résultant ainsi potentiellement en une augmentation de la productivité et du PIB sur la même période.²⁴ Cependant, ce supposé « tsunami » n'a rien de

¹⁵ May et Badawy (2026).

¹⁶ <https://ec.europa.eu/eurostat/fr/web/products-euro-indicators/w/3-02122025-bp>

¹⁷ <https://www.insee.fr/fr/statistiques/4805248#graphique-chomage-g1-fr>

¹⁸ <https://www.bls.gov/charts/employment-situation/civilian-unemployment-rate.htm>

¹⁹ <https://tradingeconomics.com/china/unemployment-rate>

²⁰ May et Badawy (2026).

²¹ <https://www.bbc.com/news/articles/cm296jzzl9yo>

<https://edition.cnn.com/2026/04/23/tech/meta-layoffs-10-percent-staff-ai>

²² En effet, l'IA a peut-être servi (tout du moins en partie) de prétexte à ces réductions d'effectifs, possiblement dans le but d'augmenter les marges de manœuvres de ces grandes entreprises afin de libérer des fonds pour d'autres investissements.

²³ Worland (2026).

²⁴ En effet, cette étude Goldman Sachs met aussi en avant le fait que la productivité économique mondiale pourrait augmenter de 7% sur la même période en raison des progrès de l'IA. Voir : Hatzius *et al.* (2023).

certain et beaucoup dépendra de l'évolution de la technologie, ainsi que des mesures mises en place par les gouvernements pour accompagner les reconversions professionnelles dans des secteurs où l'IA crée des emplois. En effet, l'IA pourrait aussi créer entre 150 à 200 millions de nouveaux emplois d'ici 2028-30, notamment en lien avec les infrastructures de l'IA (data centers, réseau électrique), la sûreté de l'IA ou sa gestion.²⁵ Ainsi, **l'équilibre entre la création et la destruction d'emplois semble ténu**, et il est impossible de prédire exactement de quel côté penchera la balance à l'horizon 2028, ou à moyen terme d'ici 2035 ; plusieurs scénarios sont possibles, qui seront examinés dans la dernière partie de l'étude.

Cependant, même si les taux de chômage n'ont pas pour l'instant fluctué de manière significative, cela ne veut pas dire que la nature du travail, des métiers et des compétences n'a pas évolué depuis novembre 2022, parfois en profondeur.²⁶ En effet, **de nombreux métiers ont déjà été transformés par la technologie**, avec une automatisation partielle ou complète d'un certain nombre de tâches où l'IA peut faire le travail plus vite et souvent mieux que l'humain. Cela inclut les tâches les plus répétitives comme la synthétisation d'un grand nombre de données et d'informations, la planification du calendrier et des agendas, où même le codage informatique basique. La liste d'exemples potentiels est longue, mais nous pouvons citer notamment les traducteurs, les codeurs informatiques, l'administration ou le marketing, et il est estimé qu'au cours des prochaines années **plus de 50% des emplois dans les pays développés subiront des transformations profondes dans la nature**

des tâches.²⁷ Ainsi, les humains se concentrent désormais plus sur l'aspect stratégique, incluant la pensée critique, l'intelligence émotionnelle et la pensée originale, qui sont des qualités fondamentalement humaines que l'IA ne peut pas et ne pourra pas remplacer avant longtemps (et peut-être jamais). Nous assistons à un certain nombre de transformations de fonds dans le marché du travail, avec une transition vers moins de main-d'œuvre et plus de réflexion stratégique, où au lieu de faire ou de produire eux-mêmes, les individus doivent vérifier et valider ce que produit l'IA, et superviser l'interaction entre différents types d'IA, notamment celle agentique (un peu à la manière d'un chef d'orchestre).²⁸ De cette manière, les humains qui sont formés et savent se servir de l'IA efficacement pourront être augmentés plutôt que remplacés par la technologie sur le marché du travail. Par conséquent, il est probable qu'à l'horizon 2028, **les humains qui se servent de l'IA remplaceront progressivement ceux qui ne s'en servent pas dans de plus en plus de secteurs**, car ils seront en mesure de combiner l'ingéniosité humaine avec la vitesse et l'envergure de la machine.²⁹

3 • Impacts plus généraux sur la société horizon 2026-28

Santé : Sam Altman affirmait déjà devant le Congrès américain en 2023 que l'IA pourra peut-être un jour permettre de **guérir du cancer**.³⁰ Cette déclaration, qui avait quelque peu surpris à l'époque, est représentative du grand enthousiasme que suscite l'IA quant aux progrès potentiels qu'elle pourra apporter dans la recherche scientifique et médicale. De façon

²⁵ World Economic Forum (2025).

²⁶ Comme indiqué dans la 1^{re} Annexe, la transformation des métiers plutôt que leur suppression est le consensus le plus fort qui s'est dégagé de l'ensemble du sondage.

²⁷ Cette étude du Boston Consulting Group se concentre sur les États-Unis mais les conclusions sont transférables aux pays développés en général. Voir : Emerson *et al.* (2026).

²⁸ Cependant, ce travail de vérification et de validation nécessite souvent des compétences et des connaissances très spécialisées, donc le travail ne disparaîtra pas mais sera ainsi transformé.

²⁹ Pour ces raisons, l'IA pouvant réaliser des tâches répétitives et fastidieuses, cela impose encore plus de raisonnement critique et de créativité de la part de l'humain (condition pour que l'IA soit source d'augmentation de la productivité des économies). Ainsi, elle renforce le niveau de compétences requis sur le marché du travail, contrairement à l'impression qu'elle donne de faciliter la tâche. En raison de l'IA, les exigences en termes de résultats seront donc plus élevées au final.

³⁰ <https://www.bbc.com/news/live/world-us-canada-65610337>

similaire à son impact sur le taux de chômage (voir ci-dessus), il est indéniable qu'au jour d'aujourd'hui et depuis le démarrage de la révolution de l'IA fin 2022, celle-ci n'a pas encore apporté de percée majeure en science comme en médecine, et nous sommes loin de pouvoir complètement guérir le cancer grâce à l'IA. Cependant, il est clair que la technologie devient progressivement un outil de plus en plus utile, voire même essentiel, pour la recherche médicale et scientifique à travers le monde. En effet, plusieurs avancées notables en science et en médecine qui ont été réalisées depuis 2022, n'auraient pas été possible sans l'IA, ou tout du moins auraient mis beaucoup plus longtemps avant d'être atteintes. Cela souligne le **rôle significatif de l'IA comme accélérateur de la recherche scientifique et médicale** aujourd'hui, et cette tendance va probablement s'accélérer à l'horizon 2028.³¹ Parmi les exemples les plus notables, citons les percées récentes dans la recherche pour les traitements contre l'Alzheimer, notamment l'IA *HistoAge* développée par l'entreprise *Mount Sinai*, qui analyse les tissus du cerveau à l'échelle cellulaire de manière digitale pour détecter la maladie longtemps avant son apparition, ainsi que pour d'autres maladies neurodégénératives.³²

Éducation : Malgré ces aspects positifs, il est indéniable que l'IA a commencé à générer d'autres problèmes sociétaux au cours des dernières années, qui risquent de s'accroître à l'horizon 2028. Tout d'abord au niveau scolaire et éducatif, l'arrivée brutale de l'IA fin 2022 a **mis les écoles et les universités dans une situation instable**. En effet, le travail d'écriture a servi pendant des siècles comme l'une des principales méthodes d'évaluation des élèves, et ce à tous les niveaux scolaires, de la petite enfance au doctorat. Le fait que les LLMs peuvent générer de grandes quantités de textes cohérents en quelques secondes a littéralement mis sens

dessus dessous un système éducatif millénaire. Il est évident que depuis l'arrivée de *ChatGPT*, un devoir fait à la maison n'a plus du tout la même valeur qu'auparavant, puisque la possibilité pour les élèves de tout âge de recourir à des logiciels de LLMs, disponibles gratuitement sur Internet qui feront le travail à leur place, est devenue trop tentante. Pour ces raisons, beaucoup de professeurs et d'enseignants n'ont eu d'autres choix que de **revenir à des méthodes d'évaluation plus traditionnelles**. Cela inclut les examens oraux, ou sinon des examens écrits « à l'ancienne » (sans ordinateurs, téléphones ou portables).

Il est donc devenu impossible d'empêcher les élèves de se servir de l'IA, car beaucoup l'utilisent maintenant pour obtenir des cours ou des tutorats gratuits et privés à la maison, en complément des cours traditionnels. Cela peut paraître bénéfique à première vue, permettant à certains élèves de rattraper un retard qu'ils auraient pu accumuler pendant les cours. De plus, cela permet aussi de **rendre le savoir et les connaissances plus accessibles** aux élèves, à l'instar de la création de *Wikipedia* en 2001.³³ Néanmoins, le danger est que si ce phénomène est poussé à l'extrême, l'IA pourrait **remplacer les enseignants** dans certains cas de figure au cours des prochaines années, notamment dans les milieux défavorisés car l'IA est moins chère que les tuteurs humains. Dans ce scénario, l'interaction humaine à l'école et à l'université pourrait devenir un luxe réservé aux classes aisées, ce qui risque de renforcer les inégalités sociales. De plus, cela soulève le problème de l'exactitude des informations prodiguées par l'IA. En effet, malgré les dernières avancées techniques, le problème des « hallucinations » des LLMs, qui peuvent inventer de manière très convaincante des informations complètement fausses, persiste toujours aujourd'hui et risque de persister encore pendant de nombreuses

³¹ Comme analysé dans la 1^{re} Annexe, un point de consensus du sondage est l'optimisme sur les applications médicales et scientifiques générées par l'IA.

³² <https://www.mountsinai.org/about/newsroom/2026/new-study-links-altered-cellular-states-to-brain-structure>

³³ Cela peut être bénéfique pour les classes populaires dans les pays développés, ainsi que pour les pays en développement où l'accès à l'éducation est souvent difficile et l'IA peut ainsi permettre d'étendre l'ouverture aux connaissances.

années selon les experts.³⁴ Ainsi, toute une génération d'élèves pourraient être éduquée avec des informations dont l'exactitude laisse à désirer, ce qui risque de **diminuer considérablement la qualité de l'éducation** et le niveau scolaire global.

À ce sujet, un autre problème connexe est le risque du développement d'une certaine paresse intellectuelle, voire d'une **diminution prononcée de la capacité d'esprit critique** des jeunes étudiants par rapports aux générations qui n'ont pas connus l'IA à l'école et à l'université. En effet, si les LLM peuvent faire tous les devoirs à la maison dans toutes les disciplines, ou tout du moins les faire suffisamment bien pour obtenir une note moyenne et passable, nous risquons d'évoluer vers une société où les jeunes ne posséderont plus les mêmes outils intellectuels que leurs parents ou grands-parents. Cela inclut la capacité de raisonnement complexe, indispensable pour évoluer avec succès sur le marché du travail, au même moment où les opportunités d'emplois pour les jeunes diplômés sont en diminution justement en raison de la révolution de l'IA. Nous risquons d'avoir une **génération formée par ChatGPT et dépendante de celui-ci**, qui ne sera pas capable de faire même un travail basique sans avoir recours à une IA. Bien entendu, il est indispensable de former la jeune génération aux outils de l'IA, qui sont en train de devenir essentiels sur le marché du travail. Mais cela ne doit pas se faire aux dépens du développement d'un esprit libre et critique, fondamental pour évoluer dans une société démocratique où les citoyens sont appelés à voter régulièrement, et pour être compétitif sur le marché du travail.

Crise identitaire : Au-delà des questions éducatives, cette situation risque également de générer une crise identitaire pour l'humanité, avec potentiellement de graves conséquences pour la société. En effet, la **distinction entre l'humain et la machine pourrait se brouiller**, un phénomène qui menace de s'amplifier au cours des prochaines années. Un premier exemple

est le symptôme, apparu depuis la sortie de *ChatGPT* en 2022, entraînant un repli sur soi et **l'isolation sociale**, où la tentation d'échanger et même de se confier à l'IA plutôt qu'à un humain, souvent par facilité, s'est rapidement propagé. Dans certains cas, nous sommes arrivés au point où la machine est devenue un confident, voire même un thérapeute, avec des exemples d'emprise psychologique néfastes avérés qui ont conduit à des situations extrêmes avec des meurtres³⁵ ou à des suicides.³⁶ Un second exemple moins tragique provient une nouvelle fois de l'éducation, où certains cas ont été rapportés de professeurs se servant de l'IA pour préparer les sujets d'examen, les élèves l'utilisant ensuite pour écrire leurs devoirs, et les professeurs ré-utilisant enfin l'IA pour corriger ces mêmes devoirs. Face à une telle situation, **l'humain sert-il encore à quelque chose ?** Ou allons-nous vers un scénario comme dans certains films de science-fiction, où les machines remplacent complètement les humains non seulement sur le marché du travail, mais également dans bien d'autres domaines ? Ces questions seront examinées plus en détail dans les scénarios à l'horizon 2035 proposés dans la dernière partie de l'étude.

4 • Principaux défis pour le secteur de l'IA à l'horizon 2026-28

I DÉFIS TECHNIQUES

Au-delà des conséquences liées au développement de l'IA, la question même de ses capacités de développement se pose aujourd'hui. En effet, malgré les percées spectaculaires au cours des dernières années, un certain nombre d'experts ont mis en garde sur le **risque de stagnation de la technologie à l'horizon 2028**. Cela impliquerait d'atteindre un plateau en termes de la vitesse du développement technologique, avec un ralentissement au cours des prochaines années. Même si cela peut paraître absurde

³⁴ Murray (2025).

³⁵ Le Monde (2025).

³⁶ Badshah (2026), <https://edition.cnn.com/2025/11/06/us/openai-chatgpt-suicide-lawsuit-invs-vis>

au regard de la vitesse de progression de l'IA depuis la sortie de *ChatGPT* en novembre 2022, les nouveaux modèles commencent à faire face à des barrières techniques non-négligeables et potentiellement sérieuses. Aussi vaste soit-il, même l'Internet a des limites et ne possède pas un stock infini de données. En effet, il est apparu en 2025 que les modèles d'IA avaient **déjà « consommé » la plupart des données disponibles sur Internet**. Sans nouvelles données pour entraîner l'IA, la technologie risque de stagner, voire même de régresser, car en utilisant toujours les mêmes données ou celles créées par d'autres IA, la technologie tourne en rond, créant ainsi un cercle vicieux.³⁷ Ce problème est particulièrement pernicieux car il est impossible de compenser ce problème en créant de nouvelles données avec l'IA, puisque celle-ci ne peut pas créer du contenu nouveau, ne faisant que de reprendre les données déjà disponibles. Pour ces raisons, l'IA est dépendante de nouvelles données originales d'origine humaine, or le problème est que l'humanité ne peut pas en produire suffisamment rapidement pour nourrir l'appétit vorace des nouveaux modèles d'IA. Ainsi, au cours de l'année 2025, les entreprises du secteur de la tech ont vu **certaines de leurs IA progresser moins rapidement que prévu** en termes de leurs performances. Un excellent exemple est la déception relative provoquée par la sortie de la version 5.0 de *ChatGPT* en août 2025, avec beaucoup d'utilisateurs accusant le nouveau modèle d'être à peine meilleur que le précédent. Même si certaines de ces critiques étaient dues au fait que l'entreprise *OpenAI* avait probablement trop exagéré les mérites de la nouvelle version, il est cependant clair que les performances du modèle 5.0 n'étaient pas substantiellement supérieures à celles de la version 4.5.

Face à ce défi profond, l'IA est-elle condamnée à la stagnation, voire à la régression à l'horizon

2028 ? Rien n'en est moins sûr, même si le risque existe et doit être pris au sérieux. Pour faire face à ces problèmes, les entreprises du secteur de la tech ont réagi rigoureusement pour tenter de trouver des solutions innovantes, et ainsi sauver les investissements massifs qui ont été faits. L'une des solutions les plus originales a été le développement de données dites de « synthèse ». Celles-ci sont des données artificielles conçues pour imiter les données du monde réel, générées par des méthodes statistiques ou via des techniques d'IA telles que l'apprentissage profond et l'IA générative. Comme l'humanité ne peut pas créer assez de nouvelles données suffisamment rapidement, cette nouvelle technique combine la rapidité des données générées par les IA, avec une intervention humaine « créatrice » pour améliorer et renforcer ces données.³⁸ L'utilisation des **données « synthétiques »** pour entraîner les nouveaux modèles d'IA s'est avérée bénéfique dans un certain nombre de cas, permettant d'éviter le piège de la stagnation et de continuer la progression des modèles, mais pas dans tous les cas car la technique est complexe et difficile à mettre en œuvre. En effet, une utilisation excessive des données de synthèse peut s'avérer contre-productive, augmentant le **risque d'effondrement du modèle** (« *model collapse* »).³⁹ Les tests ont démontré que dans certains cas, si trop de données de synthèse sont utilisées pour entraîner les modèles, cela peut engendrer l'effet inverse, à savoir une régression dans le niveau de performance, voire même un « abrutissement » de l'IA qui commence à « halluciner » de manière répétée et à générer du contenu dénué de sens.

Une autre technique qui a commencé à devenir populaire en 2026 est celle de la « **latence décisionnelle** », mentionnée précédemment. Il est indéniable que cette technique permet d'augmenter les performances de l'IA dans certaines situations, ou tout du moins d'empêcher sa stagnation ou une régression immédiate. En effet, le fait de laisser plus de temps à l'IA pour générer

³⁷ Une analogie peut être faite avec un aliment que l'on mettrait trop de fois dans le micro-ondes et qui après un certain temps, perdrait toute sa saveur et son goût.

³⁸ Les processus sont complexes et très techniques, mais l'humain intervient pour « filtrer » et améliorer ces données de synthèse en tentant d'ajouter une touche créatrice.

³⁹ Shumailov et al. (2024).

les réponses (allant de quelques secondes à quelques minutes de plus), permet à celle-ci de vérifier différents paramètres et même de tester plusieurs hypothèses, réduisant ainsi le risque d'hallucinations avec de fausses informations ou des réponses de mauvaise qualité. Néanmoins, même si la latence décisionnelle semble délivrer des résultats à court terme, cette technique ne fait que ralentir le risque de stagnation des nouveaux modèles d'IA. En effet, sans nouvelles données originales, la technologie ne peut pas progresser à la même vitesse qu'au cours des années précédentes, et la technique de la latence décisionnelle ne fait donc que de **masquer le problème**, sans le résoudre réellement. Ainsi, et malgré les sommes colossales investies par le secteur, les ingénieurs n'ont pas encore trouvé de solutions miracles pour faire face au manque de nouvelles données de qualité disponibles pour entraîner la prochaine génération. Cela souligne le risque potentiel d'un **ralentissement dans la vitesse du développement technologique**, qui pourrait apparaître à l'horizon 2028.

I DÉFIS ÉNERGÉTIQUES

Depuis le lancement de *ChatGPT* version 3.5 en novembre 2022, voire même avant, la **consommation importante des modèles d'IA en électricité** a été un sujet d'inquiétude à la fois pour les entreprises en raison des contraintes d'approvisionnement, mais également pour les pouvoirs publics et la société civile à cause des conséquences sur l'environnement.⁴⁰ En effet, même un seul prompt de ChatGPT nécessite environ dix fois la quantité d'énergie requise pour une recherche traditionnelle sur Google,⁴¹ le ratio étant encore plus élevé pour les images ou les audios générés par les IA. Depuis,

le problème n'a fait que de s'empirer, puisque les derniers modèles produits par les grandes entreprises américaines du secteur ont maintenant besoin d'énormes « **giga-fabriques d'IA** » (*AI Gigafactories*). Ces dernières consomment tellement d'énergie qu'elles ont causé dans certains cas des **pannes de courant** et coupures d'électricité localement. En 2024 par exemple, on estime que les *data centers* ont consommé environ 415 TWh (Térawattheures) d'électricité à l'échelle mondiale, se rapprochant de la consommation annuelle en électricité d'un pays comme la France⁴², avec une augmentation de 12% au cours des cinq dernières années.⁴³ D'ici à 2030, il est estimé que l'électricité consommée par les *data centers* pourrait plus que doubler pour atteindre 945 TWh, dépassant la consommation d'électricité annuelle du Japon,⁴⁴ la troisième économie mondiale. Cela a suscité de vives critiques à un moment où les prix de l'énergie ont déjà beaucoup augmenté, forçant certaines de ces « giga-fabriques » à sortir des réseaux électriques locaux en développant leurs propres sources d'énergie alternatives. Cela inclut des expérimentations avec de petits réacteurs modulaires à fission nucléaire nouvelle génération. Cependant, ces réacteurs contiennent un certain nombre de risques sécuritaires, et produisent des déchets nucléaires très problématique du point de vue environnemental.⁴⁵

D'autres techniques ont commencé à être utilisées pour produire des **modèles d'IA moins énergivores** avec plus d'efficacité énergétique⁴⁶, en se basant sur le concept des « petits modèles de langages » (*Small Language Models* – SLM), par opposition aux « grands modèles de langage » (LLM). Ces SLM sont spécifiquement conçus et entraînés pour réaliser des tâches spécialisées (médecine, droit, codage

⁴⁰ Barichella (2026).

⁴¹ Cela fut explicité par le CEO d'Alphabet en février 2023. Voir : Dastin et Nellis (2023).

⁴² En 2024, la consommation d'électricité annuelle en France s'est élevée à 449,2 TWh. Voir : RTE (2024).

⁴³ International Energy Agency (IEA - 2024).

⁴⁴ IEA (2025).

⁴⁵ Les SMR (*small modular reactor*) étant encore au stade de développement, ils ne représentent pas une solution à court terme. Ils devraient commencer à être déployés vers la fin de la décennie ou au début des années 2030.

⁴⁶ Selon la directive européenne 2010/31/UE, la définition de l'efficacité énergétique est « le rapport entre les résultats, le service, la marchandise ou l'énergie que l'on obtient et l'énergie consacrée à cet effet ».

informatique) en utilisant moins de ressources, contrairement aux LLM qui sont généralistes et peuvent apporter des réponses à tout. En effet, alors que les LLM utilisent des centaines de milliards de paramètres, les SLM en utilisent seulement entre cent millions à dix milliards. Naturellement, cela induit une consommation énergétique bien plus réduite (entre 10 et 100 fois moins⁴⁷), ainsi que des coûts moins importants. Cela est notamment le cas car ces SLM ne requièrent pas d'infrastructures massives telles que les « giga-fabriques », mais se servent directement des équipements « locaux » comme les puces des smartphones, les ordinateurs portables ou les serveurs d'entreprise.⁴⁸ En termes de performance, ces SLM sont entraînés avec des stocks de données réduits mais de plus grande qualité car spécialisés dans des domaines spécifiques, leur permettant de concurrencer les LLM sur des tâches précises.⁴⁹ Néanmoins, les SLM ne représentent pas une panacée, et sont notamment **plus vulnérables au problème de la stagnation**. Très spécialisés et utilisant beaucoup moins de paramètres, ils ont du mal à exécuter des tâches complexes au-delà de leur domaine de spécialisation restreint. Leur marge de progression est ainsi beaucoup plus faible par rapport aux LLM, précisément au moment où l'IA en général fait face à un risque de stagnation en raison du manque de nouvelles données disponibles. Sur le plan environnemental, la dispersion de la consommation énergétique sur le plan individuel fait que même si on évite d'avoir recours aux « giga-fabriques », l'impact de centaines de millions de personnes utilisant des SLM sur leur smartphone ou leurs ordinateurs si la technologie continue à se répandre grâce aux bas coûts, pourrait au final se rapprocher à moyen terme de celle des LLM (sans l'égaliser pour autant).⁵⁰ Pour ces raisons, les problèmes énergétiques auxquels font face les entreprises de la tech constituent l'un des défis les plus

importants pour le développement rapide et continu de la prochaine génération de l'IA, **sans solution évidente** à l'horizon 2028.

I DÉFIS FINANCIERS

Le lancement de *ChatGPT* en novembre 2022 a engendré une course effrénée entre les entreprises du secteur de la tech, ou chacune tente de rivaliser avec les autres avec le lancement de nouveaux modèles soi-disant toujours plus performants à des intervalles de temps toujours plus courts. A titre d'exemple, l'entreprise OpenAI a mis sur le marché au total cinq grands nouveaux modèles « jalons » depuis la version 3.5, incluant GPT4 en mars 2023, GPT4o (« Omni ») en mai 2024, les o-Series (o1 et o3) fin 2024 et début 2025, GPT4.1/GPT4.5 en février et en avril 2025, ainsi que GPT5 en août 2025, avec à chaque fois plusieurs « sous-modèles » mis sur le marché à intervalles réguliers. Les autres grandes entreprises du secteur aux États-Unis ont suivi une cadence similaire incluant *Gemini* de Google, *Claude* d'Anthropic, *Grok* de xAI, *Phi* de Microsoft, *Granite* d'IBM ou encore *Llama* de Meta, ainsi que sur d'autres continents avec *DeepSeek* en Chine, *Sarvam* en Inde, *Yandex* en Russie, ou encore *Mistral* en France. La course mondiale effrénée à l'IA a entraîné des **investissements de capitaux massifs** au cours des dernières années, principalement issus du secteur privé, incluant les dépenses d'investissement (*corporate capital expenditure* – CAPEX), le capital risque (*venture capital*) et le capital-investissement (*private equity*). En outre, les gouvernements dans de nombreux pays ont également voulu appuyer l'émergence de champions nationaux dans ce secteur hautement stratégiques avec des investissements publics significatifs pour rénover ou construire de nouvelles infrastructures (*data*

⁴⁷ Frank (2026).

⁴⁸ Cela se réfère au concept du « *Edge Computing* » ou « informatique en périphérie », qui désigne une manière de traiter les données non plus dans des data centers distants, mais au plus près de leur point d'origine.

⁴⁹ C'est justement le « business model » choisit par le champion français dans le domaine de l'IA, à savoir Mistral IA, qui lui a permis de concurrencer les géants de la tech américains, avec beaucoup moins de moyens financiers.

⁵⁰ Cela reste cependant très difficile à mesurer. En outre, de plus en plus de personnes (ou d'IA) utilisent les SLM pour un « premier jet » car les coûts sont moins importants mais n'obtiennent souvent pas les résultats escomptés en raison des limites techniques. Ils ont donc recours ensuite aux LLM. Ce doublon a pour effet d'augmenter *in fine* la consommation énergétique.

center, « giga-fabrique », etc.). Pour ces raisons, le total des investissements publics et privés dans le secteur depuis novembre 2022 est estimé à **plus de 1000 milliards de dollars**⁵¹, un chiffre qui pourrait être sous-estimé en raison d'investissements indirects dans des secteurs connexes à l'IA.

Au regard de ces sommes astronomiques, il serait normal de s'attendre à ce que l'IA ait significativement augmenté la productivité et les marges de profit des entreprises au niveau mondial. Le paradoxe est cependant qu'à l'heure actuelle, cela n'est pas le cas. En outre, le **taux d'échecs des projets d'IA sont deux fois plus élevés que les projets informatiques standards**, car plus de 80% d'entre eux échouent à atteindre la phase de production ou à générer un retour sur investissement.⁵² En effet, contrairement à un logiciel classique où le code est prévisible, l'IA dépend de la qualité des données (qui est très variable), de coûts d'infrastructures élevés, ainsi que des problèmes techniques et d'accès à de grandes quantités d'électricité, des défis très difficiles à résoudre. Ainsi, une startup d'IA aujourd'hui n'a que 10% de chances de survie à 5 ans⁵³, et seulement 5% de chances de devenir réellement rentable.⁵⁴ C'est un secteur à haute récompense mais avec un risque de faillite bien plus immédiat que dans l'informatique traditionnel. Actuellement, le **retour sur investissement est négatif pour la majorité des startups** et développeurs de modèles, mais fortement positif principalement pour les fabricants de puces et fournisseurs de matériels (*Hardware* comme Nvidia), les fournisseurs de Cloud utilisés pour la puissance de calcul de l'IA (Azure ou Google Cloud), ainsi que les entreprises de service/conseil comme Accenture ou Boston Consulting Group qui aident les entreprises traditionnelles à mieux intégrer l'IA (sans avoir à soutenir le coût des infrastructures). En

effet, on estime que **plus de 600 milliards de dollars** ont été investis globalement dans l'infrastructure de l'IA ces dernières années, mais que les revenus générés par les services d'IA (abonnements LLM, *Copilot*, etc.) ne représentent pour l'instant qu'une petite fraction de cette somme.

Pour ces raisons, le risque d'une **bulle financière spéculative** dans le secteur de l'IA a été souligné par plusieurs hauts profils, incluant le fondateur d'*OpenAI* Sam Altman, ainsi que le PDG de Google Sundar Pichai.⁵⁵ Plusieurs comparaisons ont même été établies avec la bulle Internet de l'an 2000 (*dot-com bubble*). En effet, si les entreprises ne voient pas d'augmentation significative dans leurs profits d'ici 2028, elles pourraient réduire drastiquement leurs investissements en IA, entraînant un effondrement du marché avec un effet domino. Certains affirment que nous sommes actuellement dans une phase de construction et d'établissement des fondations, et que les profits réels de l'utilisation de l'IA sont attendus d'ici la fin de la décennie. En effet, pour beaucoup d'entreprises qui s'en servent (banque, assurance, logistique), le retour sur investissement est plus subtil et se mesure souvent en gains d'efficacité, plutôt que directement en argent.⁵⁶ Néanmoins, la réalité est qu'au cours des dernières années, **environ \$16 ont été dépensés pour chaque dollar gagné avec l'IA** pour l'économie mondiale dans son ensemble, et cette équation n'est pas tenable à moyen terme pour les investisseurs.⁵⁷ La question demeure de savoir si nous allons vers un Krach boursier de grande ampleur entraînant une crise économique systémique, ou vers un atterrissage en douceur avec une consolidation du marché (ou un scénario intermédiaire, qui demeure l'hypothèse la plus probable). Dans tous les cas, nous nous dirigeons sans doute vers la **fin de l'euphorie du « AI gold rush »**, qui pur-

51 <https://businesstats.com/artificial-intelligence-ai-worldwide-statistics/?hl=fr-FR#>

52 Ryseff et al. (2024).

53 Stanford HAI (2026).

54 Challapally et al. (2025).

55 <https://www.bbc.com/news/articles/cwy7vrd8k4eo>

56 Cela inclut l'accélération de la productivité, la baisse des coûts de service client, ou l'innovation à long terme.

57 Goldman Sachs (2024).

gera le marché des startups les moins rentables pour laisser place aux entreprises ayant une utilité économique réelle (passage du « *hype* » à la maturité). A cet égard, il est important de rappeler que l'éclatement de la bulle Internet en 2000 n'a pas tué l'Internet, mais a permis une **consolidation du marché** avec des corrections certes sévères mais qui ont ensuite permis la reprise économique sur des bases plus solides et réalistes ; il pourrait en être de même pour la bulle de l'IA à l'horizon 2028.

II. Les risques systémiques de l'IA et les mécanismes de régulation applicables

1 • Risques systémiques sur l'économie : pertes d'emplois et transformations profondes dans les secteurs stratégiques

Les risques de **pertes d'emplois** ou de **transformations profondes** des métiers figurent parmi les principaux risques systémiques liés à l'IA. Alors que les sections précédentes ont examiné ces questions sous un angle général, cette partie abordera les impacts concrets sur des secteurs économiques spécifiques. Dans la plupart des domaines, la vulnérabilité dépend de la **proportion de tâches répétitives et prévisibles** (physiques ou cognitives) au sein d'un métier, car celles-ci peuvent, dans beaucoup de cas, être réalisées plus rapidement et efficacement par l'IA. Ainsi, les secteurs qui manipulent de grands volumes de données ou qui reposent sur des processus standardisés, sont en première ligne. Par exemple, **l'administration** et le **support de bureau** figurent parmi les secteurs les plus vulnérables car ils reposent sur le traitement de données, la planification et la communication standardisée. En effet, la gestion d'agendas et la réponse aux requêtes basiques des clients sont facilement automatisables. Cela

inclut les secrétaires, les commis à la saisie de données, les assistants administratifs et agents de service à la clientèle, qui sont de plus en plus remplacés par des agents conversationnels, car l'IA excelle dans l'organisation d'informations et dans la réponse aux requêtes standardisées sans intervention humaine. D'autres secteurs particulièrement vulnérables incluent le secteur **bancaire**, la **finance** et **l'assurance**, où l'IA peut désormais effectuer des analyses complexes qui étaient autrefois réservées à des professionnels qualifiés. Ainsi, l'IA est très efficace dans l'analyse de données chiffrées, la détection de fraudes et la préparation de rapports financiers. Les emplois spécifiques concernés sont les analystes financiers (pour les rapports de routine), les conseillers en prêts, les courtiers en assurance ou les comptables.⁵⁸

Le secteur **manufacturier**, le **commerce de détail** et la **logistique** sont également impactés. Bien que la robotique classique ait déjà transformé les usines, l'IA apporte une flexibilité nouvelle qui permet d'automatiser des tâches plus précises. Cela concerne les ouvriers de chaîne de montage, les contrôleurs de qualité et préparateurs de commandes.⁵⁹ De la gestion des stocks automatisée aux *Chatbots* de service

⁵⁸ En effet, les algorithmes peuvent traiter des millions de transactions pour détecter des fraudes ou évaluer des risques de crédit bien plus vite qu'un humain.

⁵⁹ La vision par ordinateur permet désormais aux machines de repérer des petits défauts ou manipuler des objets variés sans programmation rigide.

client, l'interaction humaine directe diminue avec de moins en moins de téléconseillers ou de caissiers par exemple. En outre, même des secteurs nécessitant au préalable des études avancées, comme le secteur juridique, sont également affectés. La recherche de jurisprudence et l'analyse de contrats (tâches de « **parajuriste** ») sont transformées par les outils de traitement du langage, car la rédaction de fiches produits ou de rapports de données simples peut désormais être effectuée par des IA génératives (type LLM). Pour les mêmes raisons, le métier de **marketing** est également impacté de plein fouet, comme celui de **traducteur**. En effet, la traduction de base est devenue très performante et rapide, notamment via des logiciels comme *DeepL*. Comme expliqué dans les sections précédentes, l'automatisation ne signifie pas toujours la disparition d'un métier, mais souvent sa transformation. Par exemple, la nuance culturelle reste humaine et les traducteurs doivent donc continuer à relire, corriger et améliorer les traductions produites par les IA, plutôt que de les rédiger complètement eux-mêmes. De façon similaire, les **comptables** utilisant l'IA passent aujourd'hui moins de temps à saisir des chiffres et plus de temps à conseiller les clients sur leur stratégie financière.

Cependant, il est évident que l'IA va profondément bouleverser des pans entiers de l'économie à l'horizon 2028, y compris dans les principaux domaines générateurs d'emplois comme le **secteur tertiaire** incluant les services marchands et non marchands.⁶⁰ A titre d'exemple, le tertiaire représente en France près de 80% (76,6% en 2024⁶¹) de l'ensemble des emplois, avec des chiffres similaires dans la plupart des autres pays occidentaux (75% dans l'UE et 78% aux États-Unis⁶²). Ainsi, le fait que beaucoup d'emplois vont soit disparaître, soit être transformés en profondeur, peut poser des risques systémiques aux économies des pays développés au cours des prochaines années. Cela est d'autant plus le

cas que l'une des catégories les plus vulnérables à l'IA sont les **jeunes diplômés** qui entrent sur le marché du travail. Traditionnellement, les jeunes recrues commencent par des missions d'exécution, de synthèse, de recherche documentaire ou de rédaction de premier jet, qui sont précisément les domaines où l'IA excelle. Cela risque d'encourager les entreprises à réduire leurs besoins en profils « junior », pour ne garder que des profils plus expérimentés chargés de superviser l'IA. À ce sujet, les données statistiques de 2025-2026 commencent à confirmer cette tendance, même si l'impact reste pour l'instant plus visible en termes de baisse des opportunités d'embauche que d'un chômage de masse brutal. Il ne s'agit pas d'un « remplacement » massif, mais plutôt d'une contraction des possibilités d'entrée sur le marché du travail.⁶³ Néanmoins, ces tendances sont inquiétantes et pourraient devenir un **risque systémique à l'horizon 2028**, car une société qui n'est pas en mesure d'offrir un avenir professionnel aux nouvelles générations ne pourra pas assurer son propre renouvellement.

Mécanismes de régulation applicable : Les gouvernements en France, en Europe et ailleurs dans le monde ont une responsabilité particulière pour **encadrer la révolution de l'IA**. Ils doivent limiter ces risques systémiques en protégeant les travailleurs dans ces secteurs les plus vulnérables, confrontés au remplacement par la machine, avec la montée associée du chômage et de la précarité. Des mesures concrètes peuvent être prises à l'horizon 2028 avec la mise en place de garde-fous structurels, passant de la simple gestion du chômage à une politique de transition active. Cela inclut de nouvelles lois et réglementations sur le marché du travail pour empêcher ou rendre plus difficile les licenciements de masse en raison de l'IA. Dans ce domaine, **l'Europe a pris une longueur d'avance** avec le **règlement sur l'IA (AI Act)**, dont les principales obligations pour le monde du travail

⁶⁰ Ces métiers du tertiaire sont parfois dénommés métiers « à col blanc ». Le secteur secondaire est également impacté (industries manufacturières), ainsi que le secteur primaire incluant les ressources ou l'agriculture dans une moindre mesure.

⁶¹ <https://www.insee.fr/fr/statistiques/8376826?sommaire=8376908>

⁶² Voir : **Industries Numériques** et **Groupe Crédit Agricole**.

⁶³ International Labour Organization (2026).

entrent en vigueur en août 2026. Cela inclut une supervision humaine obligatoire avec l'interdiction de laisser une IA prendre seule des décisions de licenciement ou de recrutement sans un contrôle humain effectif, ainsi que la transparence des algorithmes où les entreprises ont l'obligation d'informer les employés lorsqu'un outil d'IA est utilisé pour évaluer leurs performances ou attribuer des tâches. Des **sanctions dissuasives** sont également mises en œuvre, les amendes pouvant atteindre €35 millions (ou 7% du chiffre d'affaires), obligeant les entreprises à investir dans une **adoption éthique de l'IA plutôt qu'une automatisation brutale**. De plus, l'UE a approuvé une directive en 2024 sur l'amélioration des conditions de travail spécifiquement dans les plateformes numériques, contenant des mesures de **régulation de la gestion algorithmique** (interdiction d'enregistrer certaines données confidentielles, obligation de maintenir le contrôle humain pour l'évaluation RH).⁶⁴ En outre, beaucoup de pays européens, notamment la France, ont privilégié des solutions incluant l'**assurance transition professionnelle**, un dispositif qui maintient un salaire élevé pendant une période de formation longue (6 à 18 mois) pour changer de métier. En ce qui concerne le dialogue social, des pays comme la France imposent désormais que l'IA soit un sujet de négociation obligatoire dans les entreprises. Cela inclut les **accords d'entreprise sur l'IA** avec l'obligation pour les directions de consulter les représentants du personnel avant tout déploiement massif d'IA impactant l'emploi, ainsi que la création d'organismes tels que des observatoires nationaux chargés de surveiller en temps réel l'évolution des métiers pour anticiper les vagues de licenciements et agir avant une crise potentielle.⁶⁵

Même si ces différentes mesures vont dans le bon sens, elles devront impérativement être renforcées à l'horizon 2028 et au-delà,⁶⁶ notamment pour combler les « zones grises » où l'IA avance plus rapidement que le droit du travail. Cela pourrait inclure la création d'une « **garantie d'employabilité IA** » (*Upskilling Obligation*) ou « **droit universel à la reconversion** ». Plutôt qu'un simple droit à la formation, l'employeur serait légalement responsable du maintien de l'employabilité face à l'automatisation en proposant par exemple un parcours de montée en compétences vers une tâche de « niveau supérieur » (conception, supervision, relation client complexe) assistée par l'IA.⁶⁷ Il serait aussi loisible d'étendre au-delà des plateformes numériques à l'ensemble des salariés concernés dans d'autres secteurs les restrictions relatives à la **gestion algorithmique du travail** (*algorithmic management*) prévues par la directive de 2024 (protection des données personnelles, obligation de contrôle humain, droit à la transparence concernant les algorithmes et les données recueillies). Une autre possibilité serait la création de « **comités d'éthique et de supervision numérique** », obligatoires pour les entreprises de plus de 50 salariés ; ces comités (mêlant direction et représentants des salariés) auditeraient chaque année l'usage de l'IA pour s'assurer qu'elle ne dégrade pas les conditions de travail ou l'autonomie des employés.⁶⁸ Une mesure supplémentaire serait de renforcer le contrôle syndical avec notamment l'instauration d'une **clause de « non-substitution technologique »** dans les conventions collectives garantissant que l'IA ne peut être utilisée que pour traiter des volumes supplémentaires ou des tâches à faible valeur ajoutée, laissant les décisions critiques et la relation client complexe à des humains formés.⁶⁹

⁶⁴ Le problème étant que cette directive ne s'applique pour l'instant qu'aux salariés du numérique, alors que les algorithmes sont massivement utilisés à des fins de gestion RH dans d'autres secteurs. Voir : Directive (UE) 2024/2831.

⁶⁵ En France, même s'il existe plusieurs entités dans ce domaine, la principale est l'Observatoire des emplois menacés et émergents, lancé en 2024.

⁶⁶ En effet, comme expliqué dans la 1^{re} Annexe, un point de consensus émergeant du sondage est qu'il est nécessaire de renforcer les droits et les protections des travailleurs dans les secteurs vulnérables à l'automatisation, car les mesures actuelles sont perçues comme étant insuffisantes.

⁶⁷ Ce dispositif serait particulièrement pertinent pour les jeunes diplômés, ainsi que pour les travailleurs 'seniors' qui éprouvent souvent des difficultés à se reconvertir et s'adapter aux nouvelles technologies.

⁶⁸ Cela inclurait d'identifier les tâches menacées par l'IA, avec l'instauration d'un pourcentage du chiffre d'affaires (ou de la masse salariale) dédié spécifiquement à la reconversion des profils dont les tâches sont automatisées à plus de 50 %.

⁶⁹ Cela pourrait être accompagné d'un droit de veto temporaire des comités d'entreprise sur le déploiement d'une IA si l'évaluation des risques humains est jugée insuffisante.

En outre, la création d'un « **dividende numérique** » pour les travailleurs sera probablement nécessaire à l'horizon 2028-30 car si l'IA est susceptible de générer des gains de productivité massifs, une partie de cette valeur devrait revenir aux salariés dont le métier évolue. Cela pourrait prendre la forme d'un **fonds de transition sectoriel** au niveau national, servant à financer des retraites anticipées ou les reformatations avec des transitions longues pour les secteurs les plus touchés (incluant les jeunes diplômés), ainsi que les allocations chômage pour ceux qui perdent leurs emplois en attendant une reconversion professionnelle espérée. Une autre mesure actuellement débattue est la **réduction du temps de travail sans baisse de salaire** en utilisant les gains de productivité générés par l'IA pour financer la différence, permettant de libérer du temps pour la formation. À court terme, cela impliquerait simplement de diminuer le nombre d'heures de travail chaque jour ; à moyen terme (2030-35), cela pourrait emmener l'introduction de la **semaine de quatre jours** (voire moins) dans des cas et des circonstances bien définis. À plus long terme, à l'horizon 2035 et au-delà, des mesures plus radicales ont été proposées pour réformer les systèmes de protections sociaux, notamment l'idée d'un « **revenu universel de base** » (RUB), même si celui-ci sera très problématique à financer et mettre en œuvre. En effet, contrairement aux aides sociales classiques (comme le RSA en France), il ne s'agirait pas d'un filet

de sécurité sous condition, mais d'un socle de ressources garanti à tous en termes d'universalité⁷⁰, d'individualité⁷¹, d'inconditionnalité⁷² et de régularité⁷³ ; cela impliquerait une révolution profonde du paradigme actuel en détachant le droit social de l'existence du marché du travail.

Même si cela reste difficile à calculer précisément, l'UE et ses États membres devront probablement investir **au moins 100 milliards d'euros supplémentaires à l'horizon 2030** (1^{re} enveloppe financière proposée) pour mettre en œuvre ces différentes mesures – à la fois celles existantes et les nouvelles proposées – afin de renforcer les droits et les protections des travailleurs dans les secteurs vulnérables.⁷⁴ Cela représente environ **20 milliards d'euros par an sur les cinq prochaines années**, commençant dès 2026, avec des montants probablement encore supérieurs après 2030, surtout avec l'hypothèse d'un RUB.⁷⁵ Bien que cette somme paraisse colossale, elle reste nettement inférieure par rapport aux **gains de productivité** attendus grâce à l'IA sur la même période (estimés entre **€400-600 milliards d'ici 2030**).⁷⁶ Sur le modèle des cadres de financements de l'UE déjà existants, la moitié du financement (environ 50%) pourrait provenir du budget de l'UE, incluant le Cadre financier pluriannuel et des mécanismes d'emprunt commun (sur le modèle de *NextGenerationEU*). L'autre moitié proviendrait d'un mélange de cofinance-

⁷⁰ Le RUB serait versé à tous les membres d'une communauté (souvent un pays), sans distinction de revenus, de patrimoine ou de statut social.

⁷¹ Le RUB serait versé sur une base individuelle et non par « foyer fiscal », garantissant l'autonomie financière de chaque personne au sein d'un couple ou d'une famille.

⁷² Le RUB serait versé sans exigence de contrepartie (pas besoin de prouver qu'on recherche un emploi) et sans contrôle des ressources ; ce serait un droit de naissance.

⁷³ Le RUB serait versé de manière permanente (généralement chaque mois) tout au long de la vie, de la naissance (ou l'âge adulte) jusqu'à la mort.

⁷⁴ Ces informations ont été collectées par le biais d'un entretien, la personne concernée ayant souhaité conserver l'anonymat. Voir la liste complète des personnes interviewées à la fin de la Bibliographie.

⁷⁵ Au-delà de 2030, il devient très difficile d'anticiper exactement les impacts que l'IA aura sur le marché du travail, mais les investissements nécessaires pourraient augmenter de manière significative à moyen et à long terme si les gouvernements doivent mettre en place un RUB par exemple (voir les scénarios prospectifs dans la dernière section de l'étude).

⁷⁶ Le véritable défi n'est donc pas tant le manque d'argent, mais la vitesse de redistribution pour s'assurer que les bénéfices de l'IA arrivent suffisamment rapidement dans les caisses des États de l'UE pour financer ces mesures, avant que les licenciements d'ampleur ne commencent. Ces informations ont également été collectées par le biais d'un entretien, la personne concernée ayant souhaité conserver l'anonymat. Voir la liste complète des personnes interviewées à la fin de la Bibliographie.

ment par les États membres, de capitaux privés (investisseurs/capital-risque), ainsi que de nouvelles « ressources propres technologiques » (taxes IA/données - voir ci-dessous). Chaque État membre serait libre de fixer la répartition exacte des composantes de cette seconde moitié du financement.

Pour ces raisons, afin de financer le coût de ces mesures à court comme à moyen terme, l'idée d'une « **taxe robot** » ou taxe sur l'IA a été proposée, une solution qui vise à compenser la perte de revenus fiscaux liée au travail humain.⁷⁷ Il existe différentes pistes actuellement à l'étude en Europe et ailleurs pour structurer une telle fiscalité, même si ces mesures demeurent controversées.⁷⁸ Tout d'abord, l'idée d'une « **taxe sur l'usage** » ou l'équivalence salariale, probablement la proposition la plus débattue (soutenue notamment par Bill Gates), consisterait à traiter l'IA comme un employé virtuel. L'entreprise paierait une taxe équivalente aux cotisations sociales qu'elle aurait versées si un humain occupait le poste, avec l'objectif de maintenir le financement de la protection sociale (retraites, santé) tout en ralentissant volontairement le rythme de l'automatisation pour laisser le temps à la société de s'adapter. Une autre proposition est la **taxe sur les gains de productivité** où on taxerait le résultat de l'utilisation IA (plutôt que de taxer l'IA directement) ; cela pourrait prendre la forme d'une surtaxe de l'impôt sur les sociétés pour les entreprises dont le ratio chiffre d'affaires/nombre d'employés dépasse un certain seuil.⁷⁹

En outre, les entreprises sont actuellement encouragées à s'automatiser car l'achat de machines ou de logiciels d'IA est déductible

des impôts (amortissement). Ainsi, une solution alternative pourrait être de **supprimer ou de réduire les avantages fiscaux** liés à l'investissement dans des technologies de substitution au travail humain.⁸⁰ De plus, puisque l'IA se nourrit de données souvent produites gratuitement par les utilisateurs ou extraites du travail humain passé, une **taxe sur l'usage des données** (*big data*) – plus une IA utilise de données pour s'entraîner et fonctionner, plus l'entreprise paie une redevance – permettrait de redistribuer la valeur créée par l'IA vers le bien public.⁸¹ À moyen terme à l'horizon 2030-35, d'autres solutions ambitieuses sont également envisageables (notamment pour financer un hypothétique RUB), avec par exemple la création de **fonds souverains technologiques**. Plutôt que de compter uniquement sur l'impôt, l'État agirait comme un investisseur pour le compte de ses citoyens via la propriété collective.⁸² L'État pourrait exiger des parts au capital des entreprises d'IA en échange de l'utilisation des données publiques (santé, infrastructures, éducation) qui ont servi à entraîner leurs modèles, ou percevoir des redevances sur les brevets issus de recherches financées par des fonds publics.

2 • Soutenir la formation à l'IA pour favoriser l'augmentation humaine plutôt que le remplacement

L'une des recommandations opérationnelles les plus importantes de cette étude repose sur une **meilleure formation des employés à l'utilisation des outils de l'IA afin de favoriser l'augmentation humaine au sein des entreprises, plutôt que leur remplacement**. Comme expliqué dans

⁷⁷ Pour une mise en pratique concrète de ces différentes solutions à l'horizon 2028-30, voir les deux propositions de lois dans la 2^e Annexe, incluant une directive au niveau européen et sa retranscription dans le droit national français.

⁷⁸ Les lobbys de la tech mettent de fortes pressions sur les gouvernements dans de nombreux pays pour empêcher la mise en œuvre d'une taxe sur l'IA, argumentant que cela impacterait la compétitivité. En effet, l'un des plus grands défis demeure la coopération internationale (voir ci-dessous). Si un pays taxe l'IA seul, il risque de voir ses entreprises perdre en compétitivité face à des pays qui ne la taxent pas (« dumping technologique »).

⁷⁹ Cette solution a l'avantage d'éviter de pénaliser l'innovation technologique en elle-même, en se concentrant uniquement sur la richesse générée par l'automatisation.

⁸⁰ Cela permettrait de rééquilibrer le coût relatif entre un salarié (qui coûte cher en charges) et un algorithme (qui est fiscalement avantageux).

⁸¹ Une taxe sur les revenus publicitaires des GAFAM et autres géants du numérique serait une option complémentaire.

⁸² Ce modèle est inspiré du fonds pétrolier norvégien, ainsi que celui de l'Alaska.

les sections précédentes, la tendance à l'horizon 2028 sur le marché du travail est que les humains qui se servent de l'IA vont progressivement remplacer ceux qui ne n'en servent pas. Les chiffres démontrent que par rapport à d'autres régions du monde comme les États-Unis ou la Chine, les entreprises européennes sont en retard en termes de l'adoption de l'IA. À titre d'exemple, environ 43% des employés dans les entreprises américaines affirment avoir intégré l'IA dans au moins une fonction principale de leur métier début 2026, avec l'usage de l'IA générative devenant la norme opérationnelle dès 2024.⁸³ En Europe, seulement environ 20% des entreprises déclarent utiliser des outils d'IA en 2025 (une hausse de 6.5% par rapport à 2024), même si elles l'utilisent principalement pour des tâches de support (traduction, rédaction), plutôt que pour transformer leur cœur de métier ou créer de nouveaux modèles de revenus.⁸⁴

Ainsi, plutôt que de financer uniquement l'indemnisation chômage, les gouvernements réorientent les budgets vers la formation continue à grande échelle impliquant le **droit à la « requalification »** (*reskilling*). À partir d'août 2026, l'article 4 de l'*AI Act* impose explicitement aux organisations de garantir un niveau minimal de « **littératie en IA** » (compétences de base) pour tout le personnel utilisant ces systèmes. Cela inclut la création de référentiels de compétences sectoriels où l'UE finance certains modules spécifiques par métier (ex: IA pour le diagnostic médical ou la gestion de l'énergie), afin que le travailleur reste maître de l'outil technique. Le Bureau européen de l'IA a également lancé un « **Pacte sur l'IA** », où les entreprises signataires s'engagent à publier des rapports de transparence concernant l'impact de l'IA sur

leurs effectifs. En échange, elles bénéficient d'un accès privilégié aux « **bacs à sable réglementaires** » (*Regulatory Sandboxes*) pour tester des outils d'augmentation humaine avant leur mise sur le marché.⁸⁵ En France, des dotations spécifiques pour l'IA ont été créées pour permettre aux salariés des secteurs vulnérables (comptabilité, logistique, secrétariat) de se former aux outils de demain via des **comptes personnels de formation** (CPF) renforcés. D'autres mesures incluent les programmes de transition sectoriels avec le financement de passerelles pour aider par exemple un employé administratif à devenir un « gestionnaire de données » ou un « expert en support client augmenté ». La Commission européenne a également mis en œuvre une « **Stratégie d'application de l'IA** » (*Apply AI Strategy*) en 2025 qui mobilise des « pôles européens d'innovation numériques » (*European Digital Innovation Hubs*) servant de centres de formation de proximité afin d'encourager une plus grande adoption des outils d'IA par les entreprises (surtout les PME).⁸⁶

Mécanismes de régulation applicable : Même si ces mesures vont dans la bonne direction, elles doivent impérativement être renforcées à tous les niveaux ;⁸⁷ le retard important des entreprises européennes en termes d'adoption de l'IA souligne que des mesures plus ambitieuses sont nécessaires à l'horizon 2028 et au-delà. Pour dépasser le cadre actuel, l'UE et ses États membres pourraient explorer des réformes plus structurelles, l'enjeu étant de passer d'une logique de « formation subie » (pour ne pas être remplacé) à une logique de « **formation conquérante** » (pour devenir un travailleur augmenté).⁸⁸ En termes de fiscalité innovante, cela pourrait inclure la création d'un « **crédit d'impôt**

⁸³ Bick *et al.* (2026).

⁸⁴ <https://ec.europa.eu/eurostat/web/products-eurostat-news/w/ddn-20251211-2>

⁸⁵ Des subventions via le programme *Europe numérique* sont fléchées vers les entreprises qui démontrent que l'IA est utilisée pour réduire la pénibilité ou augmenter la précision des tâches complexes, plutôt que pour supprimer des postes.

⁸⁶ La Commission a aussi développé des structures comme la *Marie Skłodowska-Curie Action MSCA Choose Europe* incluant un « bassin de talents à l'IA » (*AI Talent Pool*), accompagné par la création de bourses d'études pour l'IA sous l'égide d'une « Académie des talents de l'IA » (*AI Skills Academy*).

⁸⁷ La 1^{re} Annexe est explicite sur ce point, puisque pratiquement tous les répondants du sondage sont fortement d'accord sur le besoin de développer des programmes de formation à l'IA et de montée en compétences à tous les niveaux, allant du salarié au cadre dirigeant.

⁸⁸ Ces informations ont été collectées par le biais d'un entretien, la personne concernée ayant souhaité conserver l'anonymat. Voir la liste complète des personnes interviewées à la fin de la Bibliographie.

augmentation humaine » afin de générer une incitation fiscale pour les entreprises qui investissent spécifiquement dans des logiciels d'IA conçus pour l'assistance (ex: « cobotique », copilotes de codage, outils d'aide au diagnostic) plutôt que dans des systèmes d'automatisation complète de processus. Le crédit d'impôt serait conditionné à la preuve d'une formation certifiante pour 100 % des employés concernés. Sur le modèle du fonds pour une transition juste (dédié au climat), une autre proposition serait de créer un « **fonds européen de transition numérique** » dédié aux secteurs où l'IA menace les emplois (incluant les jeunes diplômés). Ce fonds financerait non pas des licenciements, mais des « **congs de conversion IA** » de longue durée (6 à 12 mois), rémunérés à 90 % du salaire brut de référence pour le 1er mois, puis 70% pour les mois suivants.⁸⁹ Afin de renforcer l'appareil de formation, la création d'un « **Passeport européen de compétences IA** » permettrait d'encourager un système de certification interopérable dans toute l'UE,⁹⁰ offrant la possibilité à un travailleur de prouver sa capacité à collaborer avec l'IA, facilitant la mobilité professionnelle vers des postes « augmentés ». En complément, il serait bénéfique d'intégrer l'IA au cœur d'une **nouvelle « 5^e liberté » du marché unique européen** qui serait créée, dédiée spécifiquement à l'éducation, à la recherche et à l'innovation. Cette nouvelle 5e liberté viendrait s'ajouter aux quatre libertés historiques relatives aux biens, aux services, aux personnes et aux capitaux.⁹¹

En raison de l'impact très différencié de l'IA selon les domaines, même si l'UE a commencé à mettre en place certaines formations spécialisées, celles-ci doivent être consolidées par le développement de formations spécifiques à l'IA dans des secteurs stratégiques clés, notamment

les secteurs bancaire et financier, l'assurance, la santé, le droit, etc. Cela pourrait inclure l'instauration de **micro-accréditations sectorielles obligatoires** avec l'imposition de modules courts (*micro-credentials*) validés par les branches professionnelles.⁹² Enfin, une autre mesure serait la création d'un « **droit à la transparence algorithmique au travail** », obligeant les employeurs à fournir une « **notice d'impact sur les compétences** » pour chaque nouvel outil d'IA qui détaillerait quelles compétences humaines seront renforcées et lesquelles risquent de s'atrophier, permettant aux syndicats de négocier des formations compensatoires.⁹³ Dans tous les domaines, l'objectif doit être de développer la complémentarité entre l'humain et la machine afin qu'ils deviennent des binômes qui se renforcent mutuellement.⁹⁴ Ce duo gagnant-gagnant, où l'IA et l'humain se concentrent chacun sur les domaines où ils peuvent apporter la meilleure contribution possible, doit être au cœur des formations en IA pour les employés tous secteurs confondus, **favorisant l'augmentation humaine plutôt que son remplacement**. À ce sujet, la création d'un « **index de l'augmentation humaine** » public (similaire à l'index égalité femmes-hommes) qui noterait les entreprises en fonction de leur capacité à maintenir l'emploi tout en intégrant l'IA, serait pertinent ; les entreprises les mieux notées auraient un accès prioritaire aux marchés publics européens.⁹⁵

⁸⁹ Cela est comparable à ce qui avait été fait en France avec le premier confinement (2020) lors de l'épidémie du Covid-19.

⁹⁰ Ce passeport pourrait être basé sur le cadre DigComp (*Digital Competence Framework*) 2.2 actualisé en 2025-2026.

⁹¹ Cette proposition est issue du rapport d'Enrico Letta (2024) intitulé : « [Much More than a Market](#) ».

⁹² Par exemple, selon ce système, un architecte devra valider un module « éthique et pilotage de la conception générative » pour conserver son droit de signature sur des projets utilisant l'IA.

⁹³ Cela pourrait inclure une participation en amont à la conception et à l'usage des algorithmes de gestion du travail.

⁹⁴ L'IA réalisant rapidement toutes les tâches répétitives et récurrentes, cela permet aux humains de se concentrer sur les tâches où le contact humain et personnel, la créativité et l'ingéniosité sont nécessaires.

⁹⁵ Pour une mise en pratique concrète de ces différentes solutions à l'horizon 2028-30, voir les deux propositions de lois dans la 2^e Annexe, incluant une directive au niveau européen et sa retranscription dans le droit national français.

3 • Favoriser la reconversion professionnelle vers les secteurs dont l'expansion est favorisée par l'IA

Les mesures mentionnées ci-dessus (à la fois celles existantes et les nouvelles proposées), doivent impérativement être **ciblées vers les nouveaux secteurs dont l'expansion est favorisée par l'IA**. Cet objectif devrait constituer une priorité pour les gouvernements afin de favoriser la reconversion professionnelle à l'horizon 2028 et au-delà. En effet, on estime que des millions de nouveaux rôles verront le jour, centrés sur la collaboration humain-machine ; les mesures proposées ci-dessus doivent inclure une concentration sur ce type de secteurs où l'IA crée et continuera de créer des emplois au cours des prochaines années. Cela comprend tout d'abord la **santé** et la **biotechnologie**. En effet, l'IA permet une médecine de précision et accélère la découverte de traitements innovants avec de nouveaux métiers comme les bio-informaticiens, les spécialistes de la médecine génomique assistée par l'IA, et les experts en diagnostic radiologique augmenté.⁹⁶ Un autre secteur majeur est celui de la **cybersécurité** et de l'**éthique**, car à mesure que l'IA devient plus puissante, les risques augmentent, créant un besoin vital de protection et de régulation incluant des nouveaux métiers comme les éthiciens de l'IA pour garantir l'absence de biais, les responsables de la conformité algorithmique, ou les analystes en cyberdéfense prédictive.⁹⁷ De plus, le secteur des **infrastructures numériques** (*big data*) est bien entendu très impacté, car les « moteurs » de l'IA nécessitent une maintenance physique et logicielle colossale, ce qui suppose

de nouveaux postes en tant qu'ingénieurs dans les centres de données, architectes de solutions informatiques comme l'informatique en périphérie (*edge computing*), ou spécialistes de la « curation » de données. En effet, la demande pour les **experts et ingénieurs en données** (*data scientist* et *data engineer*) est en croissance exponentielle pour structurer les grandes quantités de données que l'IA doit assimiler dans les nouveaux modèles.⁹⁸

En outre, la valeur ajoutée humaine sur le marché du travail se déplace actuellement vers les « **soft skills** » incluant l'esprit critique, l'empathie, le jugement moral, l'intuition ainsi que la pensée stratégique. À l'horizon 2028, ces compétences sont les plus difficiles à automatiser, et donc celles qui non seulement résistent le mieux à l'IA, mais créent également le plus de valeur ajoutée sur le marché du travail.⁹⁹ Cela comprend de nombreux secteurs d'emplois déjà existants qui se renforcent en raison de l'IA, comme les **services à la personne** et à la **santé**, car si l'IA aide au diagnostic, le soin reste une affaire de contact impliquant l'écoute active, l'intelligence émotionnelle et la patience. Ainsi, les métiers se concentrant sur l'**accompagnement psychologique** et le **soutien moral** sont renforcés, tandis que les machines gèrent les aspects techniques et logistiques. Un autre exemple est l'**éducation** et la **formation**, car même si l'accès au savoir devient universel grâce à l'IA, l'apprentissage nécessite un guide humain utilisant la pédagogie adaptative, la motivation et la transmission de valeurs. Le rôle du professeur en tant que « mentor » est vital pour aider les élèves à apprendre comment s'orienter dans la masse d'informations disponibles, et pour développer leur esprit critique.¹⁰⁰

⁹⁶ Par exemple, la création de jumeaux numériques d'organes demande des ingénieurs capables de modéliser le vivant.

⁹⁷ À titre d'exemple, la [Stratégie nationale de cybersécurité française 2026-30](#) met un accent particulier sur le renforcement des formations et l'augmentation des effectifs nationaux en cybersécurité.

⁹⁸ Enfin, dans le domaine de la formation professionnelle, l'IA crée un besoin permanent de mise à jour des compétences avec de nouveaux métiers comme les formateurs spécialisés en outils d'IA générative, concepteurs pédagogiques adaptatifs ou coachs en collaboration humain-IA.

⁹⁹ L'IA est excellente pour traiter des données, mais elle échoue encore face à l'imprévisibilité humaine, aux nuances émotionnelles et aux dilemmes éthiques complexes, renforçant ainsi l'importance de « l'intelligence relationnelle ».

¹⁰⁰ Il y avait déjà eu des craintes par le passé qu'avec le développement d'Internet, les enseignants seraient progressivement remplacés, mais cela ne s'est jamais concrétisé. En effet, de nombreuses études ont démontré depuis que rien ne remplace la relation humaine dans les apprentissages, et cela continue d'être le cas même avec la nouvelle révolution de l'IA.

Enfin, un dernier secteur qui résiste de manière surprenante et paradoxale à l'IA est celui des **métiers manuels et tactiles** (plombier, maçon, menuisier, électriciens, jardiniers), car même les nouveaux modèles d'IA robotique ne parviennent pas à reproduire la dextérité humaine impliquant des gestes complexes avec les trois dimensions de l'espace. Pour ces raisons, en plus de favoriser la formation des employés à l'IA afin de pourvoir les nouveaux métiers créés, les gouvernements doivent également investir davantage dans les formations pour ces métiers déjà existants, qui sortent renforcés par la révolution technologique.¹⁰¹

En général, le coût total de ces différentes mesures (à la fois celles existantes et les nouvelles proposées) concernant le renforcement des formations à l'IA pour les travailleurs afin de favoriser l'augmentation humaine et la reconversion professionnelle vers les secteurs où l'IA permet une expansion, pourrait être **d'au moins 50 milliards d'euros pour l'UE et ses États membres à l'horizon 2030**. Cela représente environ **10 milliards d'euros par an sur les cinq prochaines années**, commençant dès 2026.¹⁰² Néanmoins, cette somme doit être perçue comme étant incluse dans les 100 milliards d'euros déjà mentionnés précédemment (1^{re} enveloppe financière proposée), concernant le coût total lié à la gestion des risques systémiques posés par l'IA sur l'économie et le marché du travail en Europe.

¹⁰¹ Une fois de plus, pour une mise en pratique concrète de ces différentes solutions à l'horizon 2028-30, voir les deux propositions de lois dans la 2^e Annexe.

¹⁰² Ces informations ont été collectées par le biais d'un entretien, la personne concernée ayant souhaité conserver l'anonymat. Voir la liste complète des personnes interviewées à la fin de la Bibliographie.

Tableau synthétique – Impacts de l'IA sur les secteurs d'activité et les métiers (horizon 2028-30)

Secteur / Métiers concernés		Impact estimé (horizon 2028-30)	Justification
Administration & Support de bureau			
Secrétaires, assistants admin., commis à la saisie de données		Fort négatif	Tâches répétitives et standardisées entièrement automatisables (IA agentique)
Téléconseillers & agents service client		Fort négatif	Chatbots et agents IA remplacent les requêtes standardisées
Caisiers		Moyen négatif	Automatisation progressive des points de vente
Finance, Banque & Assurance			
Analystes financiers (rapports), conseillers en prêts, courtiers assurance		Fort négatif	Analyse de données, détection de fraudes, rapports : automatisables par l'IA
Comptables		Transformation	Moins de saisie, plus de conseil stratégique ; rôle profondément transformé
Spécialistes data / cybersécurité financière		Fort positif	Fortte demande pour supervision et sécurisation des systèmes d'IA
Juridique, Marketing & Communication			
Parajuristes, avocats juniors, traducteurs (basique), rédacteurs contenu std.		Fort négatif	Recherche juridique, traduction, rédaction de 1er jet : remplacés par LLM
Avocats seniors, stratégies marketing, traducteurs (révision culturelle)		Transformation	Recentrage sur jugement, créativité, relation client et nuance culturelle
Industrie manufacturière, Logistique & Commerce de détail			
Ouvriers de chaîne de montage, contrôleurs qualité, préparateurs commandes		Fort négatif	Robotique + vision par ordinateur automatisent les tâches physiques précises
Conducteurs (transport, livraison)		Incertain	Véhicules autonomes en développement ; impact difficile à évaluer d'ici 2028-30
Informatique, IA & Infrastructures numériques			
Codeurs informatiques (codage basique & répétitif)		Fort négatif	L'IA génère du code standard très efficacement
Développeurs seniors, architectes logiciels		Transformation	Supervision du code IA, architecture complexe, validation
Data engineers/scientists, ingénieurs data centers, spéc. cybersécurité, éthiciens IA		Fort positif	Demande exponentielle pour faire tourner, sécuriser et réguler l'IA
Santé & Biotechnologie			
Bio-informaticiens, ingénieurs médecine génomique, modélisation du vivant		Fort positif	Médecine de précision et découverte de traitements : nouveaux métiers clés
Médecins / radiologues (diagnostic)		Transformation	L'IA augmente le diagnostic ; rôle humain recentré sur l'interprétation complexe
Soignants, infirmiers, accompagnants psychologiques		Moyen positif	Soin = contact humain ; l'IA gère la logistique, pas l'empathie
Éducation, Formation & Services à la personne			
Enseignants (transmission contenu)		Transformation	Rôle mentor renforcé ; pédagogie, esprit critique, transmission de valeurs
Concepteurs pédagogiques IA, formateurs outils IA		Fort positif	Nouveaux métiers : conception de parcours d'apprentissage augmentés par l'IA
Travailleurs sociaux, assistants de vie, psychologues		Moyen positif	Intelligence émotionnelle et accompagnement humain non automatisables
Métiers manuels & artisanat (tous secteurs)		Limité	Dextérité 3D et interactions complexes avec l'espace : résistance forte à l'IA
Profils transversaux			
Jeunes diplômés / profils juniors (toutes industries)		Fort négatif	Tâches d'exécution juniors (synthèse, recherche, rédaction) absorbées par l'IA
Travailleurs formés à l'IA générative (toutes industries)		Fort positif	Humain + IA = productivité démultipliée ; remplaceront ceux ne maîtrisant pas l'IA

Légende – Niveaux d'impact :

Fort négatif	Moyen négatif	Transformation	Limité	Incertain	Moyen positif	Fort positif
--------------	---------------	----------------	--------	-----------	---------------	--------------

« Transformation » = métier profondément transformé dans la nature des tâches et des compétences, sans disparition totale.

4 • Désinformation exacerbée par l'IA, avec des risques systémiques pour les démocraties

Le problème de la **désinformation** et de la **manipulation des processus électoraux** est un fléau qui affecte les démocraties dans le monde entier depuis longtemps, mais particulièrement depuis les dix ou quinze dernières années. Cette situation est en grande partie liée à l'avènement des **réseaux sociaux**, qui offrent des plateformes idéales pour lancer des campagnes de désinformation massives pour manipuler les élections en Europe et en occident. Cela est également dû à la montée des tensions entre démocraties et dictatures à travers le monde, et dans le cas de l'Europe, cela est plus particulièrement le résultat de la **cyber guerre menée par la Russie** depuis l'annexion de la Crimée en 2014.¹⁰³ L'objectif est souvent de promouvoir la montée des populismes et des extrêmes politiques avec le vote antisystème, afin de diviser l'occident et **miner les démocraties de l'intérieur** en renforçant la polarisation de l'électorat et des partis politiques. Ces campagnes de désinformation sont devenues **quasiment systématiques** à chaque élection en Europe et ailleurs. Elles ont déjà eu des impacts très concrets en influençant par exemple le résultat sur le référendum du Brexit en 2016, ou la victoire de Trump à l'élection présidentielle américaine de 2016¹⁰⁴ (et dans une moindre mesure en 2024). Même s'il est très difficile de mesurer avec précision leur impact exact, la forte polarisation des sociétés démocratiques et les votes de plus en plus serrés dans tout type d'élection, font que même une petite influence peut avoir *in fine* un effet significatif sur le résultat final, influençant profondément le destin d'une nation à la fois pour sa politique interne et sur ses relations internationales.

Depuis le lancement de *ChatGPT* en novembre 2022, le fléau de la désinformation n'a fait que de s'accélérer. En effet, l'IA offre la possibilité de

produire des textes, des audios ou des vidéos presque instantanément et en très grande quantité, avec le potentiel de **démultiplier les campagnes de désinformation** et de les étendre à une échelle bien plus importante en automatisant au moins partiellement les processus. Ainsi, un hacker russe n'a plus besoin d'écrire lui-même tous les textes ou de préparer toutes les vidéos ou les audios, car il peut utiliser l'IA et plusieurs *bots* en même temps pour faire le travail à sa place, tout en gardant une supervision d'ensemble. Cela permet de décupler la vitesse, l'impact et l'ampleur des campagnes de désinformation étrangères, qui peuvent être traduites dans des dizaines de langues instantanément, tout en respectant les codes culturels locaux pour paraître authentique. **L'IA agentique** permet aussi **l'automatisation de la viralité** car contrairement aux anciens bots qui répétaient la même phrase, ceux de 2026 utilisent des LLM pour engager des conversations réelles et noyer les faits sous une masse de commentaires divergents.¹⁰⁵

Ainsi, nous sommes entrés dans une nouvelle ère où il est devenu **de plus en plus difficile de distinguer le vrai du faux** en raison des « *fake news* » et « *deepfakes* » ultraréalistes qui pullulent sur Internet et les réseaux sociaux. En raison de l'IA, nous sommes passés de « *fake news* » isolées à une véritable **industrialisation de la tromperie** permettant de créer des preuves visuelles et sonores presque indétectables. Ainsi, des *deepfakes* de dirigeants politiques sont utilisés pour influencer les scrutins avec par exemple de fausses annonces de retrait de candidatures ou de scandales. En outre, l'IA agentique peut désormais créer des chaînes de preuves (faux articles de presse, faux commentaires d'experts, fausses photos de témoins) pour soutenir un seule et même mensonge, rendant la démystification très difficile. L'IA permet également le **microciblage psychologique**, les algorithmes examinant les vulnérabilités émotionnelles des utilisateurs (peur, colère, espoir) pour adapter le ton et l'argumentaire de la désinformation. Il

¹⁰³ Barichella (2022).

¹⁰⁴ Barichella (2022).

¹⁰⁵ En effet, les créateurs de désinformation utilisent maintenant l'IA pour tester des milliers de variantes d'un titre ou d'une image afin de voir laquelle « mord » le mieux sur les algorithmes de recommandation des réseaux sociaux.

existe certaines techniques pour détecter ces situations telles que le **filigrane** ou « tatouage » **numérique** (« *Watermarking* »), qui consiste à intégrer un signal unique et reconnaissable dans les résultats d'un modèle d'IA, comme un texte ou une image, afin d'identifier si le contenu a été généré par l'IA. Néanmoins, l'exactitude et la précision de ces techniques est assez variable, en plus de ne pas être codés ou appliqués de la même manière à travers le monde, ce qui limite leur transférabilité d'un pays à l'autre. Cela complique la tâche des agences nationales pour la sécurité des systèmes d'information, chargées de lutter contre les cybermenaces et la désinformation en ligne.¹⁰⁶ Ainsi et à l'horizon 2028, la menace qui pèse sur les processus démocratiques en France, en Europe et partout dans le monde pourrait devenir systémique, avec le **danger de manipulation en masse des processus électoraux via la désinformation**.¹⁰⁷

Mécanismes de régulation applicable : L'auditabilité de l'IA et la correction des biais sont devenus des piliers de la défense contre la désinformation en ligne.¹⁰⁸ En effet, l'auditabilité est la capacité d'examiner les processus internes d'une IA pour comprendre comment elle parvient à une conclusion. Cela permet non seulement la traçabilité en remontant à la source d'une fausse affirmation¹⁰⁹, mais également de tenir les développeurs ou plateformes pour responsables en prouvant où le système a failli. Les biais algorithmiques des IA peuvent agir comme des amplificateurs de désinformation en déformant la réalité de manière subtile. La correction des biais permet ainsi de neu-

traliser les préjugés et d'éviter les « chambres d'écho » où la désinformation prospère¹¹⁰, ainsi que les stéréotypes infondés qui renforcent les campagnes de propagande. En 2026, l'auditabilité de l'IA en Europe est passée d'un concept éthique à une **obligation légale** stricte sur l'**ensemble du cycle de vie** du modèle avec la pleine entrée en vigueur de l'**AI Act**, complété par des standards techniques en cours de finalisation par les organismes européens (CEN-CENELEC).

En termes de standards pratiques pour l'auditabilité des IA à « haut risque », ceux-ci s'articulent autour des Articles 12 (Journalisation) et 13 (Transparence). Concernant l'auditabilité des données, le standard émergent est le *prEN 18286* (gouvernance des données pour l'IA) avec des exigences de documentation de la provenance, des méthodes de collecte et du prétraitement, incluant l'obligation de tenir un registre des biais statistiques dans les données d'entraînement.¹¹¹ En termes de l'auditabilité du modèle (documentation technique), le cadre de référence opérationnel est la liste ALTAI (*Assessment List for Trustworthy AI*) avec la « journalisation automatique » (Article 12), où les systèmes doivent enregistrer automatiquement les événements tout au long de leur cycle de vie (détection d'erreurs, changements d'états, interventions humaines) pour permettre la traçabilité des incidents et des changements.¹¹² Pour l'auditabilité de la décision (explicabilité), l'**AI Act** impose la **surveillance humaine** (*human oversight*) où les décisions de l'IA doivent être interprétables par un humain, comprenant une « interface d'audit » pour per-

¹⁰⁶ La plupart des démocraties dans le monde ont mis en place des agences nationales dans ces domaines. En France, l'ANSSI (Agence nationale de la sécurité des systèmes d'information) est chargée de la cybersécurité, tandis que le VIGINUM (Service de vigilance et protection contre les ingérences numériques étrangères) s'occupe de la protection contre les ingérences numériques étrangères. Voir : Barichella (2023b) et Barichella (2025).

¹⁰⁷ Cela risque d'encourager la montée des extrêmes politiques, renforçant la polarisation et les fractures dans les démocraties.

¹⁰⁸ Comme indiqué dans la 1^{re} Annexe, les biais, l'auditabilité et la transparence sont identifiés dans le sondage comme constituant les principaux obstacles à une IA responsable.

¹⁰⁹ Si une IA génère une fausse information, l'audit permet de savoir si elle a été entraînée sur des données frauduleuses ou si elle a « halluciné ».

¹¹⁰ Les algorithmes de recommandation biaisés ont tendance à ne montrer aux utilisateurs que ce qui confirme leurs croyances préexistantes (biais de confirmation). Corriger ces biais aide ainsi à briser ces bulles de filtres.

¹¹¹ Il y a aussi des exigences pour la traçabilité, incluant les choix de labellisation et les critères d'exclusion des données.

¹¹² Cela comprend l'utilisation des standards ISO/IEC adaptés au contexte européen, via des « fiches techniques » (model cards) qui sont devenues obligatoires pour décrire l'architecture, les paramètres de performance et les limites du modèle.

mettre à un superviseur humain d'intervenir ou de comprendre le raisonnement de l'IA en temps réel. Cela implique de documenter le « poids » des variables ayant mené à un résultat, avec l'obligation de fournir une notice d'utilisation claire expliquant les limites du système et son degré de précision.

De façon paradoxal, afin de renforcer l'auditabilité de l'IA ainsi que pour les méthodes de mesure et de correction des biais, d'autres outils d'IA spécifiques ont été développés. Par exemple, concernant la correction des biais en continu, les MLOps (*Machine Learning Operations*) deviennent le standard industriel pour l'auditabilité. Cela inclut l'utilisation d'outils comme *Evidently AI* ou *Great Expectations* pour surveiller en temps réel si les données dérivent de celles d'entraînement (« *Data Drift & Model Drift* »), ce qui pourrait introduire de nouveaux biais imprévus (auquel cas une alerte est déclenchée). Pour les registres de conformité, cela inclut la mise en place de « fiches techniques » (*model cards*) dynamiques qui se mettent à jour à chaque réentraînement. Au sujet des outils de mesure et de correction des biais, l'approche de l'Europe repose sur la « **surveillance après commercialisation** » (*Post-Market Monitoring*). Pour les outils de mesure (détection) afin de quantifier les biais (impact hétérogène, égalité des chances), plusieurs bibliothèques *open-source* sont devenues des standards de fait recommandés par les institutions de l'UE. Cela inclut le *AI Fairness 360* (AIF360), un cadre complet pour mesurer et atténuer les biais sur les groupes protégés (genre, âge, origine), ainsi que le *Fairlearn*, très utilisé pour évaluer les disparités de performance entre différents groupes démographiques. À propos des outils de correction, il existe trois étapes techniques¹¹³, conformément

aux recommandations de la CNIL (France) et de l'EDPB (Europe) : **Pré-traitement**¹¹⁴, **In-traitement**¹¹⁵, et **Post-traitement**¹¹⁶.

Même si ces mesures vont dans le bon sens, elles risquent d'être insuffisantes et doivent être renforcées au cours des prochaines années, à la fois au niveau européen et à l'échelon national.¹¹⁷ Cela est d'autant plus le cas en raison du contexte international actuel très instable, où les **régimes autoritaires** renforcent leurs techniques en se servant de l'IA pour lancer des **campagnes de désinformation à grande échelle contre les pays occidentaux**. L'auditabilité de l'IA et la correction des biais sont des outils de régulation essentiels pour lutter efficacement contre la désinformation en ligne et la manipulation des processus démocratiques, nécessaires pour que l'IA devienne, au contraire, un **bouclier de vérité**. En dépit de cela, le processus actuel de simplification des lois européennes dans le secteur du numérique (*Digital Omnibus* – analysé plus en détail ci-dessous) risque de faire précisément le contraire, en allégeant les règles sur les entreprises par souci de compétitivité. Ainsi, il est essentiel que l'allègement du supposé fardeau bureaucratique via la simplification des normes ne se fasse pas aux dépens de la rigueur des procédés pour l'auditabilité de l'IA, la mesure et la correction des biais, qui sont essentiels pour lutter contre la désinformation et **protéger les démocraties de l'intérieur**.

Pour renforcer le cadre existant, l'UE pourrait explorer plusieurs pistes innovantes à l'horizon 2028 pour muscler son arsenal visant à passer d'une conformité « statique » (auto-évaluation) à une **surveillance dynamique et technologique**.¹¹⁸ Cela inclurait par exemple la création de « **bacs à sable de données** » (*Data Sandboxes*)

¹¹³ Même si l'application cumulée de ces trois étapes n'est pas une obligation légale rigide, leur mise en œuvre est fortement recommandée (et donc parfois de fait obligatoire), afin de respecter les principes fondamentaux du droit français et européen.

¹¹⁴ Rééquilibrage des données (pondération) avant l'entraînement pour équilibrer les classes sous-représentées.

¹¹⁵ Ajout de contraintes d'équité directement dans la fonction de perte (*loss function*) de l'algorithme.

¹¹⁶ Ajustement des seuils de décision (classification) après l'entraînement pour égaliser les résultats entre groupes.

¹¹⁷ Comme indiqué dans la 1^{re} Annexe, un point de consensus dans le sondage est lié aux inquiétudes profondes sur les risques de désinformation exacerbés par l'IA, avec le besoin de renforcer les interfaces de supervision et mécanismes d'auditabilité.

¹¹⁸ Ces informations ont été collectées par le biais d'un entretien, la personne concernée ayant souhaité conserver l'anonymat. Voir la liste complète des personnes interviewées à la fin de la Bibliographie.

sécurisées gérées par le Bureau européen de l'IA pour l'audit de biais car actuellement, les développeurs peinent à corriger les biais faute d'accès aux données sensibles (genre, origine) en lien avec le RGPD. Les entreprises pourraient tester leurs algorithmes sur des jeux de données réels et diversifiés, certifiés par l'UE, pour mesurer les biais sans jamais posséder ou voir ces données sensibles.¹¹⁹ Une autre mesure serait la création d'un système de « **passport algorithmique** » dynamique,¹²⁰ qui serait rendu obligatoire pour chaque modèle d'IA à haut risque, mis à jour tout au long du cycle de vie. Ce passeport contiendrait l'historique des données d'entraînement, les résultats des tests de biais successifs, ainsi qu'un journal des modifications ; en cas de désinformation, les autorités pourraient vérifier si une mise à jour spécifique du modèle est à l'origine du problème. De plus, la désinformation étant souvent multi-plateforme,¹²¹ la création de nouveaux **systèmes de détection « inter-plateformes »** (*Cross-Platform Signaling*), avec l'imposition d'un protocole d'échange de signaux en temps réel, ferait en sorte que si un audit détecte un biais de manipulation inquiétant sur une plateforme, un signal d'alerte automatique (hashtag de l'image ou du texte) serait instantanément envoyé aux autres plateformes afin de limiter la propagation. Enfin, le filigrane numérique actuel étant souvent invisible ou facile à supprimer, rendre l'intégration des métadonnées de provenance (C2PA) obligatoires au niveau du matériel (capteurs de smartphones) et des logiciels de création pour tout contenu destiné à l'espace public européen, permettrait de **rendre les filigranes plus détectables**.¹²²

5 • Les impacts écologiques négatifs de l'IA multifformes et systémiques

Comme expliqué dans les précédentes sections, les défis énergétiques pour l'IA sont significatifs, sans solution évidente à court terme. Cela a encouragé plusieurs grandes entreprises du secteur comme Google et Microsoft à commencer à investir sérieusement dans les sources d'énergie du futur, telles que la **fusion nucléaire**. En effet, des personnalités de premier plan comme Bill Gates ont investi des milliards de dollars dans des start-ups spécialisées dans la fusion afin de développer les premiers réacteurs expérimentaux, dans le but de fournir de l'énergie d'ici le début des années 2030.¹²³ Malgré certaines percées technologiques récentes, le consensus scientifique est que la fusion nucléaire ne représentera pas une part importante dans le mix énergétique avant la deuxième moitié du 21e siècle.¹²⁴ En effet, des défis majeurs persistent, à la fois sur le plan technique et économique. Même si Bill Gates parvenait à lancer les premiers réacteurs à fusion nucléaire dans dix ans, il faudrait encore plusieurs décennies avant de remplacer les infrastructures fossiles dans les pays développés, et le processus prendrait encore plus de temps pour les pays en développement. Ainsi, et jusqu'à l'avènement d'une source d'énergie propre et potentiellement illimitée comme la fusion nucléaire, **l'IA va vraisemblablement continuer à épuiser les réserves énergétiques mondiales, avec un impact de plus en plus négatif sur le changement climatique**. En effet, à l'horizon 2028-2030, l'IA ne sera plus une part négligeable des émissions du secteur numérique, mais un moteur principal. À titre d'exemple, l'AIE estime que **la demande en électricité des centres de données optimisés par**

¹¹⁹ Notamment via du calcul sécurisé ou la « confidentialité différentielle » (*differential privacy*).

¹²⁰ Celui-ci serait inspiré du passeport numérique existant des produits pour l'économie circulaire.

¹²¹ Par exemple, une vidéo générée sur une X peut ensuite être rapidement partagée sur TikTok et Facebook.

¹²² Au lieu de simplement dire « Ceci est une IA », le système permettrait d'afficher : « Cette image a été créée le XX à XXh, modifiée par l'IA modèle X, et n'a pas de source photographique réelle. »

¹²³ Blum (2025).

¹²⁴ Wimmers et al. (2025).

l'IA devrait plus que quadrupler d'ici 2030.¹²⁵

Rien qu'aux USA, la consommation électrique des *data centers* devrait représenter près de la moitié de la croissance de la demande en électricité d'ici 2030. Sous l'impulsion de l'utilisation de l'IA, l'économie américaine devrait, en 2030, consommer plus d'électricité pour le traitement des données que pour la fabrication de l'ensemble des biens à forte intensité énergétique, y compris l'aluminium, l'acier, le ciment et les produits chimiques. Au sein des pays développés en général, les centres de données devraient générer plus de 20 % de la croissance de la demande en énergie d'ici 2030, remettant ainsi le secteur de l'électricité sur la voie de la croissance après des années de stagnation ou de baisse de la demande dans nombre d'entre eux.¹²⁶

En plus de la consommation énergétique, un autre impact écologique négatif de l'IA, souvent dissimulé, est la **consommation massive d'eau requise pour les *data centers*** et autres processus associés. Cela se manifeste à plusieurs niveaux, incluant le refroidissement direct des serveurs pour éviter la surchauffe, la production d'électricité hors site pour le fonctionnement des serveurs, ainsi que la fabrication des composants. La consommation mondiale d'eau en raison de l'IA est déjà estimée entre quatre à six milliards de mètres cube d'eau par an, soit l'équivalent d'environ la **moitié de toute la consommation d'eau annuelle d'un pays comme le Royaume-Uni**.¹²⁷ Cette tendance est à la hausse avec des augmentations d'environ 15% à 34% selon les années pour les grandes entreprises du secteur comme Google ou Microsoft.¹²⁸ Cela endommage non seulement les écosystèmes et la **biodiversité locale**, mais puisque les géants de la tech utilisent de l'eau potable traitée pour le refroidissement, cela ajoute des pressions supplémentaires sur les bassins versants et les nappes phréatiques déjà fragilisées au sein des pays développés. Dans les pays en

développement, cela entre en concurrence directe avec les **besoins vitaux des populations locales**, notamment pour l'agriculture. En plus de cela, les centres de données dédiés à l'IA dépendent d'un grand nombre de puces spécialisées et puissantes dont la durée de vie est courte (3 à 5 ans) ; ces puces contiennent des minéraux de terres rares (lithium ou cobalt), souvent extraits selon des méthodes nuisibles à l'environnement, ce qui génère d'importantes quantités de **déchets électroniques**.¹²⁹ Par conséquent, les impacts environnementaux des technologies de l'IA sont profonds et multi-formes, entraînant des conséquences négatives substantielles et des risques systémiques à l'horizon 2028 et au-delà.

Mécanismes de régulation applicable : L'Europe, à travers son cadre législatif et ses initiatives techniques, est actuellement l'un des **leaders mondiaux pour encadrer l'impact écologique de l'IA**. Par exemple, l'*AI Act*, entré en vigueur progressivement, intègre désormais des **obligations de transparence environnementale**. Cela inclut la déclaration de l'empreinte énergétique et environnementale où les fournisseurs de modèles d'IA à usage général (comme les LLM) et les systèmes à « haut risque », doivent documenter et publier leur consommation d'énergie et d'eau tout au long du cycle de vie (entraînement et utilisation). La Commission a aussi lancé des consultations en avril 2026 pour définir des méthodes de calcul standardisées (KPI¹³⁰) afin de pouvoir mesurer précisément l'énergie consommée par l'IA et ainsi diminuer le « *greenwashing* ». Par le biais de la directive sur l'efficacité énergétique (*Energy Efficiency Directive*), les centres de données de plus de 500 kW ont l'obligation depuis 2024-2025 de rapporter publiquement leurs performances énergétiques, ainsi que l'utilisation des énergies renouvelables et leur consommation d'eau (*power and water usage effectiveness*). L'Europe mise également sur la sobriété énergétique et

¹²⁵ IEA (2025).

¹²⁶ *Ibid.*

¹²⁷ Kenny (2025).

¹²⁸ Ren (2023).

¹²⁹ Barichella (2026).

¹³⁰ *Key Performance Indicators (KPI)*, ou « Indicateurs clés de performance ».

l'optimisation technique en encourageant le passage à des IA plus « efficiente », notamment en soutenant financièrement via le programme « Horizon Europe » le développement des **Small Language Models** (SLM) plus petits, spécialisés et moins énergivores¹³¹, ainsi que des recherches sur les techniques de *quantification*. Ces techniques impliquent la réduction de la précision des calculs pour augmenter la vitesse et économiser l'énergie et les coûts, ainsi que la promotion d'algorithmes optimisés pour réduire le nombre d'opérations par seconde nécessaires pour obtenir un résultat identique (**éco-conception logicielle**).¹³²

L'UE œuvre au développement d'infrastructures et centres de données plus « verts », en les soumettant à des contraintes de plus en plus strictes. Cela inclut la récupération de chaleur, où les nouveaux centres de données sont encouragés (et dans certains cas obligés) à réinjecter la chaleur produite par les serveurs dans les réseaux de chauffage urbain ou serres agricoles. L'UE encourage aussi l'utilisation du refroidissement liquide (*liquid cooling*), plus efficace que la climatisation classique pour gérer la densité thermique des nouveaux processeurs IA.¹³³ En outre, le « **Climate Neutral Data Centre Pact** », une initiative d'autorégulation soutenue par l'UE, engage les signataires à utiliser 100 % d'énergies renouvelables d'ici 2030 pour leurs centres de données. Enfin, pour guider les entreprises et les consommateurs, l'Europe prépare un système de labellisation impliquant la transparence et l'étiquetage (**label « Green AI »**), prévu avant la fin 2026. Cela inclura un score comparable à celui de l'électroménager, permettant de

comparer les modèles d'IA sur la base de leur efficacité énergétique (performance par Watt) et leur impact sur l'environnement (comme la consommation d'eau).¹³⁴ Même si toutes ces mesures vont dans le bon sens, elles sont insuffisantes et doivent être renforcées à l'horizon 2028 afin d'atténuer l'augmentation exponentielle en électricité du secteur de l'IA prévue au cours des prochaines années (voir de-dessus). En dépit de cela, le processus actuel de simplification mis en œuvre par la Commission européenne via les lois « **Omnibus** », qui inclut les réglementations sur le climat et le numérique, risque au contraire de compromettre et diminuer leur efficacité. Il est essentiel de ne **pas sacrifier l'écologie sur l'autel de la compétitivité**, car l'impact environnemental de l'IA est substantiel et menace l'avenir de la planète.

Afin de renforcer le cadre actuel et mieux contrôler les impacts environnementaux de l'IA à l'horizon 2028, l'UE pourrait par exemple instaurer un « **droit à l'audit environnemental** » pour les tiers (actuellement les rapports sont souvent déclaratifs), permettant à des organismes indépendants publics d'accéder aux compteurs électriques et d'eau des *data centers* en temps réel.¹³⁵ De plus, l'application du principe de pollueur-payeur aux logiciels d'IA – via une « **Taxe Carbone Numérique** » proportionnelle sur les modèles dépassant un certain seuil de requêtes quotidiennes – permettrait de limiter la pollution liée à l'usage de masse.¹³⁶ En contrepartie, cela pourrait inclure la réduction de la TVA pour les entreprises qui utilisent des processeurs plus économes promouvant l'éco-conception logicielle. En outre,

¹³¹ Les SLM utilisent le « *edge computing* » où la promotion de l'IA décentralisée (exécutée sur les appareils locaux plutôt que dans le Cloud), ce qui réduit le transport de données et la charge sur les serveurs géants (voir ci-dessus).

¹³² L'objectif est l'optimisation de « l'inférence ». Si l'entraînement d'un modèle d'IA est énergivore, c'est en réalité l'inférence (l'utilisation quotidienne par des millions de personnes) qui représente environ 60 à 80 % de la facture énergétique totale d'un système sur son cycle de vie, un sujet important dans l'agenda législatif de l'UE au cours des prochaines années.

¹³³ Cette technologie utilise beaucoup moins d'eau que les tours de refroidissement traditionnelles car elle dissipe la chaleur de manière plus ciblée. L'UE encourage également les centres de données à utiliser des eaux grises traitées ou des eaux de pluie pour le refroidissement afin de ne pas puiser dans les réserves d'eau potable locales.

¹³⁴ Cela permettra de valoriser les modèles optimisés qui minimisent les ressources par rapport à leur puissance de calcul.

¹³⁵ Un label officiel de l'UE (« IA verte ») serait délivré uniquement après vérification physique de l'optimisation du code et des infrastructures.

¹³⁶ Ces informations ont été collectées par le biais d'un entretien, la personne concernée ayant souhaité conserver l'anonymat. Voir la liste complète des personnes interviewées à la fin de la Bibliographie.

l'UE pourrait aussi légiférer pour favoriser la **standardisation des SLM** en imposant aux fournisseurs de proposer systématiquement une version SLM optimisée frugale pour les tâches simples, afin d'éviter tout gaspillage via les LLM. Afin de réduire la consommation d'eau des *data centers*, l'UE pourrait également établir des **quotas d'eau par Téraflo**¹³⁷ en fixant une limite maximale de litres d'eau consommés par unité de calcul.¹³⁸ Enfin, pour réduire le volume des déchets électroniques, l'UE devrait créer un **index de réparabilité et de durabilité des infrastructures d'IA** en imposant aux fournisseurs de recycler systématiquement leur puces, processeurs ou autres matériels dès que cela est possible, ou sinon de les revendre ou les donner pour des tâches moins lourdes à des organismes de la société civile, par exemple.

Dans le même temps, les technologies d'IA ont également été saluées pour leur potentiel à réduire les émissions de GES et les impacts environnementaux dans certains domaines. Par exemple, l'IA contribue à **l'optimisation du réseau électrique** grâce à une meilleure gestion de l'intermittence des énergies renouvelables. L'IA est ainsi devenue un outil indispensable pour gérer la complexité de la transition énergétique via les « **réseaux intelligents** » (*smart grids*), générant de nouveaux « emplois verts ».¹³⁹ De même, l'IA contribue à améliorer l'efficacité des ressources, par exemple via « **l'agriculture de précision** », qui implique une gestion parcellaire utilisant des technologies numériques pour optimiser les intrants (engrais, eau, pesticides) selon les besoins réels, permettant de réduire la consommation d'eau et l'utilisation de pesticides. De même, on s'appuie désormais sur l'IA pour le suivi des émissions de carbone, car elle permet de surveiller en temps réel la déforestation via l'imagerie satellite, ou de détecter

des fuites de méthane dans les oléoducs et gazoducs, difficiles à repérer pour l'œil humain. Enfin, l'IA est aussi utilisée pour tenter de surmonter certains des obstacles technologiques de la **fusion nucléaire**, dans l'espoir d'accélérer le calendrier de son déploiement.¹⁴⁰ Par conséquent, la **relation entre l'IA et l'environnement est complexe et multidimensionnelle**, et il est essentiel que les gouvernements se servent du potentiel de la technologie pour mieux protéger l'environnement dans ces différents domaines, en l'intégrant plus directement et explicitement dans les mesures et initiatives citées précédemment.

6 • Perte de souveraineté en raison de la dépendance de l'Europe aux technologies étrangères

Le secteur de l'IA est incontestablement dominé par les États-Unis qui sont talonnés de près par la Chine, les deux étant imbriqués dans une course effrénée aux répercussions mondiales. En effet, cette situation implique que les autres pays et continents sont devenus dans de nombreux cas **dépendants des technologies d'IA américaines ou chinoises**, sans pouvoir développer leurs propres champions dans ce secteur hautement stratégique (tout du moins pas suffisamment pour concurrencer la domination sino-américaine). Cela est notamment le cas de la France et de l'Europe, pour qui cette dépendance technologique soulève des questions voire des risques en termes de **perte de souveraineté**. En effet, l'UE est parfois caricaturée comme étant une « **colonie numérique** » des États-Unis. Cette dépendance est complexe et multiforme, apparaissant à trois niveaux critiques de la chaîne de valeur de l'IA. Tout d'abord

¹³⁷ Il s'agit d'une mesure ou indicateur de la puissance de calcul des ordinateurs.

¹³⁸ En plus, l'UE devrait interdire le refroidissement par évaporation d'eau potable en obligeant l'usage d'eaux grises traitées ou de circuits fermés stricts dans les zones en stress hydrique.

¹³⁹ Cela inclut les gestionnaires de réseaux intelligents, les ingénieurs en optimisation carbone par l'IA et les analystes de données climatiques. Ainsi, pour atteindre la neutralité carbone, les entreprises recrutent de plus en plus des profils hybrides capables d'utiliser l'IA pour réduire l'empreinte environnementale des chaînes logistiques.

¹⁴⁰ Schenker (2025).

en termes des **infrastructures pour le Cloud**¹⁴¹, probablement l'aspect le plus critique car en 2026, environ 65% du marché européen dans ce secteur est dominé par trois entreprises américaines (Amazon Web Services - AWS, Microsoft Azure et Google Cloud). En effet, l'IA nécessite une puissance de calcul massive que seules ces « *hyperscalers* » peuvent offrir à une échelle suffisante. Deuxièmement, concernant le **matériel** (« *hardware* »), les processeurs nécessaires pour entraîner les modèles d'IA, appelés « Unité de traitement graphique » (*Graphics Processing Unit* - GPU), proviennent quasi exclusivement d'entreprises américaines comme NVIDIA. Sans ces puces, l'Europe ne peut pas physiquement faire fonctionner ses propres IA. Ainsi, si un conflit diplomatique survient, les États-Unis disposent d'un levier de pression immense sur l'économie européenne. Enfin, en ce qui concerne les **modèles de base**, bien que des acteurs européens comme **Mistral AI** (France) ou **DeepL** (Allemagne) soient performants, les modèles les plus utilisés par les entreprises restent ceux d'OpenAI, de Google ou de Meta, souvent intégrés directement dans les outils de travail quotidiens comme Microsoft 365 ou Google Workspace, utilisés par des millions d'entités publiques et privées au sein de l'UE. Ce risque de « **verrouillage** » (*Vendor Lock-in*) implique que plus une entreprise intègre l'IA d'un fournisseur américain dans ses processus, plus il devient coûteux et techniquement difficile d'en sortir.¹⁴² Cette perte d'autonomie décisionnelle signifie que l'Europe suit le rythme d'innovation et les conditions tarifaires fixés par la Silicon Valley, qui peuvent être coûteuses.

Pour ces raisons, la **perte de souveraineté de l'Europe** sur le plan technologique est manifeste, avec des conséquences non seulement économiques, mais également politiques. En utilisant des infrastructures américaines, les entreprises et administrations européennes

restent exposées aux lois américaines comme le **Cloud Act**.¹⁴³ Cette loi permet aux autorités américaines (via mandat judiciaire) d'exiger des fournisseurs de services Cloud américains l'accès aux données qu'ils stockent, même sur le sol européen. De cette manière, si une entreprise française entraîne son modèle d'IA sur des serveurs situés physiquement à Paris, dès lors que l'hébergeur est américain, les données sont accessibles par la justice américaine. En outre, cela crée également un conflit juridique direct avec le **Règlement général pour la protection des données** (RGPD). Ainsi, les entreprises, qu'elles soient européennes ou américaines, sont soumises au RGPD en Europe qui interdit le transfert de données vers des pays tiers sans protection adéquate, et en même temps au Cloud Act qui les oblige à livrer les données. La conséquence est qu'en cas de demande américaine, le fournisseur est contraint de violer soit la loi américaine soit celle européenne, avec le risque de lourdes sanctions des deux côtés. Pour l'UE, cela signifie qu'elle ne **contrôle plus réellement la confidentialité des données de ses citoyens et de ses entreprises**. En outre, si les données transitent ou sont stockées chez des acteurs soumis au *Cloud Act*, les États-Unis peuvent théoriquement accéder à des secrets industriels, des algorithmes propriétaires ou des données stratégiques (santé, défense, énergie) sous couvert d'enquêtes criminelles ou de sécurité nationale. Cela a souvent pour conséquence de décourager les entreprises européennes de confier leurs actifs les plus précieux au Cloud par manque d'alternatives européennes de même échelle, ce qui peut freiner l'innovation et la compétitivité de l'UE.

Mécanismes de régulation applicable : Au cours des dernières années, l'Europe a accéléré sa stratégie de souveraineté numérique pour contrer la domination des géants américains.¹⁴⁴ En 2026, cette ambition ne repose plus seu-

¹⁴¹ Le *Cloud* (informatique en nuage) consiste à utiliser des serveurs informatiques à distance, hébergés dans des centres de données connectés à Internet pour stocker, gérer et traiter des données, plutôt qu'un serveur local ou un ordinateur personnel.

¹⁴² Cela est aussi en partie lié aux habitudes de travail bien ancrées, avec un temps d'adaptation non-négligeable aux alternatives souveraines que beaucoup d'entreprises préfèrent ainsi remettre à plus tard par facilité.

¹⁴³ *Clarifying Lawful Overseas Use of Data Act*, promulgué par le Congrès américain en 2018.

¹⁴⁴ L'UE mise sur la « souveraineté verte », alliant puissance technologique et sobriété énergétique.

lement sur la régulation, mais aussi sur une véritable **offensive industrielle et technologique**. Tout d'abord, l'UE utilise son « **pouvoir de marché** » en tant que plus grand marché unique au monde, avec des liens commerciaux transatlantiques profonds, pour imposer ses standards, forçant les entreprises américaines à s'adapter aux normes européennes dans le domaine du numérique pour pouvoir accéder au marché européen.¹⁴⁵ Cela inclut non seulement l'*AI Act*, mais également les principales lois de l'UE dans le secteur comme le *Digital Markets Act* (DMA), *Digital Services Act* (DSA), *Data Governance Act* (DGA), ou le RGPD, qui sont assorties de **sanctions très élevées** en cas de non-respect des normes européennes.¹⁴⁶ Cependant, la stratégie de l'UE n'est plus seulement de réguler, mais également de **produire ses propres champions industriels dans le numérique** pour renforcer sa souveraineté. Cela inclut le projet PIIEC IA (« *Projet Important d'Intérêt Européen Commun* ») lancé fin 2025, qui permet aux États membres de subventionner massivement des programmes de recherche et d'infrastructure en IA sans tomber sous le coup des règles anti-concours habituelles. Les gouvernements soutiennent aussi les « champions » européens du secteur comme *Mistral AI* (France), *Aleph Alpha* et *DeepL* (Allemagne) ou *Silo AI* (Finlande), qui bénéficient de financements notamment via le programme « *Horizon Europe* » pour développer des modèles multilingues adaptés aux spécificités culturelles et juridiques de l'Europe.¹⁴⁷

La dépendance aux Cloud américains (AWS, Azure, Google Cloud) étant l'un des principaux maillons faibles, l'Europe a réagi en développant des **stratégies de Cloud souverain**, tout en

renforçant ses infrastructures pour développer une puissance de calcul plus indépendante. Cela inclut l'initiative « *Euro-3C* » (2026), un projet de €75 millions pour construire une infrastructure fédérée *Telco-Edge-Cloud*, visant à décentraliser le stockage des données en les traitant au plus proche de l'utilisateur sans passer par des serveurs étrangers. En France, le label *SecNumCloud*¹⁴⁸ est une qualification créée en 2016 afin de garantir la qualité et la sécurité d'un service Cloud de confiance (avec l'équivalent européen *EUCS*¹⁴⁹), certifiant par exemple l'immunité contre les lois extraterritoriales américaines comme le *Cloud Act*. Enfin, l'Europe investit massivement dans des **supercalculateurs** (réseau *EuroHPC*) pour que les entreprises européennes n'aient pas à louer de la puissance de calcul aux États-Unis pour entraîner leurs propres modèles. Cela comprend le « *Plan d'action pour le continent de l'IA* » (*AI Continent Action Plan*), lancé par la Commission en 2025, qui vise à **renforcer les infrastructures numériques**, notamment le réseau des « fabriques » et « giga-fabriques » d'IA, ainsi que le développement des « laboratoires de données » (*Data Labs*) au sein de ces fabriques d'IA pour mieux agréger de larges quantités de données. L'UE met également en place une stratégie « **Data Union** » pour faciliter le partage de données de haute qualité entre entreprises européennes pour entraîner des IA spécialisées (santé, industrie, énergie), domaines où l'Europe possède encore un avantage compétitif sur les données sectorielles.¹⁵⁰

Même si toutes ces mesures vont dans le bon sens, elles sont insuffisantes et doivent impérativement être augmentées à l'horizon 2028 avec davantage d'investissements et un ren-

¹⁴⁵ Il y a également un débat qui s'intensifie pour instaurer une « préférence européenne » dans les marchés publics de défense et de santé, afin de financer l'écosystème local plutôt que d'acheter des licences étrangères.

¹⁴⁶ Par exemple, l'UE a imposé une amende record contre X de €120 millions en décembre 2025 pour non-respect du DSA.

¹⁴⁷ Cela est désormais au cœur de la stratégie de l'UE, notamment via l'utilisation de modèles « *Open Source* » pour garantir que le code et les données restent auditables sur le continent (voir section ci-dessous).

¹⁴⁸ Le *SecNumCloud* est une qualification de l'ANSSI.

¹⁴⁹ EU Cloud Certification Scheme.

¹⁵⁰ L'UE encourage ainsi la création d'espaces de données en facilitant le partage de données pour la santé, l'industrie et la mobilité entre acteurs européens, afin de créer des jeux de données massifs capables de rivaliser avec ceux des États-Unis.

forcement des infrastructures européennes et nationales.¹⁵¹ Cela permettra à la France et à l'UE d'être plus compétitives dans le développement de la prochaine génération de logiciels d'IA, tout en renforçant la souveraineté et en réduisant la dépendance aux technologies étrangères. Bien que l'Europe progresse, le rapport Accenture de 2025¹⁵² souligne que 65% des organisations européennes admettent ne pas pouvoir rester compétitives sans recourir (au moins partiellement) à des fournisseurs non-européens.¹⁵³ Pour ces raisons, l'autarcie complète étant jugée financièrement et technologiquement intenable par la majorité des experts à court comme à moyen terme, le consensus s'oriente vers un **objectif de souveraineté pragmatique et ciblé** au cours de la prochaine décennie (viser environ 20 à 25% de parts de marché dans les technologies critiques avec des alternatives européennes viables). Même si cela reste difficile à calculer précisément, l'investissement supplémentaire de l'UE et ses États membres requis pour atteindre cet objectif est estimé aux alentours de **30 milliards d'euros supplémentaires par an d'ici 2035 (300 milliards d'euros au total – 2^e enveloppe financière proposée)**. Cela est nettement supérieur aux dépenses actuelles, et doit commencer dès 2026.¹⁵⁴ Bien que cette somme paraisse une nouvelle fois colossale, elle demeure inférieure par rapport aux gains de productivité estimés grâce à l'IA sur la même période, et pourrait donc au moins partiellement être financée via les propositions

pour une « **taxe robot** » sur l'IA mentionnée dans les sections précédentes.¹⁵⁵ Cela pourrait prendre la forme de la création d'un **fonds souverain européen dédié au renforcement des infrastructures numériques critiques**, afin de financer le développement de nouveaux centres de données, fabriques et giga-fabriques d'IA.¹⁵⁶ Ces investissements doivent cibler en priorité l'expansion des initiatives mentionnées ci-dessus, notamment les super-calculateurs, les solutions de Cloud souverains, ainsi que les « champions » locaux par le biais de la commande publique (principe de « préférence européenne »), et via les partenariats public-privé. En complément, il est essentiel pour l'UE d'accélérer la mise en œuvre de la stratégie *Data Union* afin d'encourager la mutualisation des données dans les secteurs clés (santé, industrie, énergie). De plus, accélérer le **Chips Act** européen pour relocaliser la production de puces avancées (GPU et accélérateurs d'IA) sur le sol européen, afin d'éviter les ruptures de chaîne d'approvisionnement, est également indispensable.

En outre, le **rapport Draghi** de 2024 a mis en lumière le manque de compétitivité de l'UE, notamment dans le domaine du numérique, comme étant le résultat de **règlementations excessives** qui freinent l'innovation et empêchent ainsi l'émergence de champions dans le secteur numérique, ce qui permettrait à l'Europe de renforcer sa souveraineté.¹⁵⁷ En

¹⁵¹ Comme indiqué dans la 1^{re} Annexe, l'un des points de consensus les plus notables du sondage est que la dépendance technologique de l'UE vis-à-vis des USA et de la Chine est un problème qui appelle à la création de champions numériques européens, comprenant le besoin d'une politique industrielle coordonnée à l'échelle de l'UE.

¹⁵² Wood et al. (2025).

¹⁵³ Ainsi, de plus en plus d'organisations européennes adoptent une approche hybride en utilisant les Cloud américains pour leurs transactions quotidiennes, mais mobilisant les solutions souveraines nationales ou européennes pour les sujets sensibles. Néanmoins, les solutions souveraines dans l'UE ne sont pas toujours suffisamment développées même dans ces cas précis.

¹⁵⁴ Ces informations ont été collectées par le biais d'un entretien, la personne concernée ayant souhaité conserver l'anonymat. Voir la liste complète des personnes interviewées à la fin de la Bibliographie.

¹⁵⁵ Ces fonds serviraient à co-investir avec des acteurs industriels, harmoniser les standards via des consortiums (comme Gaia-X), et utiliser la commande publique pour propulser les entreprises de la tech européenne.

¹⁵⁶ Comme pour la 1^{re} enveloppe financière proposée et sur le modèle des cadres de l'UE déjà existants, la moitié du financement (environ 50%) pourrait provenir du budget de l'UE, incluant le Cadre financier pluriannuel et des mécanismes d'emprunt commun (sur le modèle de *NextGenerationEU*). L'autre moitié proviendrait d'un mélange de cofinancement par les États membres, de capitaux privés (investisseurs/capital-risque), ainsi que de nouvelles « ressources propres technologiques » (taxes IA/données - voir sections précédentes). Chaque État membre serait libre de fixer la répartition exacte des composantes de cette seconde moitié du financement.

¹⁵⁷ Draghi (2024).

réponse, la Commission a proposé une **simplification** des grandes lois sur le numérique (*AI Act*, RGPD, DMA, DSA, etc.) afin d'alléger le supposé fardeau bureaucratique pour les entreprises, surtout les PME.¹⁵⁸ Dénommé « **Omnibus numérique** » (*Digital Omnibus*), celui-ci est prévu d'être adopté d'ici l'été 2026. Cependant, beaucoup d'organisations de la société civile ont mis en garde que cette « simplification » par souci de compétitivité risque de devenir une **déréglementation extensive** sous la pression des lobbys du secteur, revenant en arrière sur un certain nombre de protections et droits fondamentaux.¹⁵⁹ Plusieurs experts ont souligné que ce supposé excès de réglementation ne serait pas la seule voire même la principale cause du manque de compétitivité de l'UE, et que la **fragmentation persistante du marché unique** est un facteur tout aussi important.¹⁶⁰ Ainsi, il est essentiel pour l'Europe de trouver un **juste équilibre** pour faire en sorte que la simplification ne se fasse pas aux dépens des droits fondamentaux, tout en continuant à œuvrer pour une intégration renforcée du marché unique en levant les barrières restantes. En effet, le problème de l'UE n'est pas tant au niveau de la recherche qu'au passage à l'échelle (*scaling*) des entreprises en raison des difficultés à lever des fonds ; cela est dû à la fragmentation persistante des marchés de capitaux. Ainsi, la mise en place d'une véritable **union des marchés de capitaux à l'horizon 2028** permettrait de faciliter les investissements massifs pour que les « licornes » européennes de l'IA ne soient pas rachetées par des groupes étrangers dès qu'elles ont besoin de lever des sommes importantes. Cela doit être complété par une politique de « **commande publique stratégique** », avec l'utilisation des budgets publics pour acheter des solutions d'IA européennes, créant ainsi un marché intérieur garanti pour les entreprises.¹⁶¹

¹⁵⁸ En effet, les petites et moyennes entreprises (PME) affirment se retrouver particulièrement pénalisées par rapport aux grandes entreprises, et le *Digital Omnibus* a notamment pour but de lever un certain nombre d'obligations réglementaires pour les PME, tout en les conservant du moins en partie pour les grandes entreprises.

¹⁵⁹ Lazaro Cabrera (2025).

¹⁶⁰ Jahangir (2025).

¹⁶¹ Ces informations ont été collectées par le biais d'un entretien, la personne concernée ayant souhaité conserver l'anonymat. Voir la liste complète des personnes interviewées à la fin de la Bibliographie.

III. Recommandations opérationnelles pour les entreprises et décideurs publics

1 • L'Open Source comme stratégie pour renforcer la souveraineté numérique

L'Open Source peut être définie comme un mode de développement et de distribution de logiciels dont le code source est rendu public et accessible à tous. Contrairement aux logiciels dits « propriétaires » ou « fermés », sa philosophie repose sur trois piliers fondamentaux, à commencer par le **libre accès**, où tout le monde peut lire, copier et étudier le fonctionnement du programme. Deuxièmement, la **modification est accessible à tous**, car tous les utilisateurs ont le droit d'améliorer le logiciel ou de l'adapter à leurs besoins spécifiques. Enfin en termes de redistribution, il est permis de **partager le logiciel** (original ou modifié) avec n'importe qui, souvent sous les mêmes conditions de licence. Même si au point de départ, l'Open Source était un choix d'ordre technique ou philosophique, il est devenu aujourd'hui un levier politique majeur de l'UE pour diminuer sa dépendance aux technologies étrangères, et ainsi **renforcer sa souveraineté numérique**. L'Open Source offre des garanties que les logiciels propriétaires (souvent américains) ne peuvent pas offrir par nature. Cela inclut l'**autonomie stratégique** car contrairement aux solutions propriétaires, il est impossible de « couper l'accès » à un code source ouvert. Cela protège ainsi l'Europe contre d'éventuelles sanctions extraterritoriales ou des

changements brusques de politique commerciale. En termes de **transparence** et de **sécurité**, le pouvoir d'auditer le code permet de s'assurer de l'absence de portes dérobées (*backdoors*) utilisées souvent à des fins d'espionnage, ce qui est crucial pour les administrations publiques et les secteurs sensibles (défense, énergie, etc.). L'Open Source offre également la possibilité de sortir de l'enfermement propriétaire (*Vendor Lock-in*) en permettant de changer de prestataire sans devoir reconstruire toute son infrastructure, car les standards restent ouverts.

L'Europe possède un vivier de talents, mais manque souvent de capitaux face à la Silicon Valley. **L'Open Source rééquilibre la donne** avec le potentiel de devenir un moteur économique majeur, notamment via la mutualisation des coûts. En effet, cela permet à des pays comme la France, l'Allemagne et la Suède de co-investir dans des solutions communes (ex : infrastructures Cloud souveraines comme Gaia-X). Cette situation entraîne également un réinvestissement local car au lieu de payer des licences coûteuses qui partent aux États-Unis, l'argent est investi dans des services, de la maintenance et du développement réalisés par des entreprises européennes. Pour ces raisons, **l'Open Source est l'une des conditions nécessaires à la souveraineté numérique européenne**, car il est impossible de rivaliser avec le budget des GAFAM en restant sur un modèle de logiciel fermé. Ainsi, la Commission a adopté une

stratégie interne Open Source 2020-2023 qui favorise systématiquement le logiciel libre pour ses propres services, prouvant que le sujet est désormais au cœur des discussions à Bruxelles. Plusieurs **champions européens** ont pu émerger via l'Open Source au cours dernières années, incluant *Mistral AI* (France – LLM), *Odoo* (Belgique – gestion de données), *Grafana Labs* (Suède – surveillance de données), *SUSE* (Allemagne – Linux & Cloud), ou *Sonar* (Suisse/France – qualité de code).

Ainsi, au sein de l'enveloppe annuelle estimée à 30 milliards d'euros mentionnée ci-dessus pour bâtir une souveraineté numérique pragmatique à l'horizon 2035 (2^e enveloppe financière proposée), la part qui devrait être ciblée pour le développement et la consolidation de l'Open Source se situerait idéalement entre 10 % à 15 %, soit **environ 4 milliards d'euros supplémentaires par an sur les dix prochaines années**.¹⁶² La priorité doit être mise sur les secteurs de l'IA et du Cloud, comme l'un des leviers les plus rentables et stratégiques pour briser le monopole propriétaire des GAFAM. Cela serait bénéfique pour l'économie européenne, certaines études ayant révélé que l'Open Source contribue à hauteur de 65 à 95 milliards d'euros au PIB de l'UE.¹⁶³ Pourtant, les investissements directs actuels paraissent dérisoires par rapport aux enjeux car l'absence de solutions souveraines oblige l'UE à verser des milliards de dollars en licences à des acteurs non-européens (principalement les *Hyperscalers* américains).¹⁶⁴ Une autre proposition à court terme serait que l'UE et ses États membres **dédient 1% de leur budget informatique global à la maintenance et au développement de l'Open Source** qu'ils utilisent déjà ; cela représenterait des centaines de millions d'euros injectés directement dans la sécurité des bibliothèques logicielles. Enfin, pour détrôner les GAFAM, les solutions Open Source européennes doivent être aussi **ergonomiques** que celles de la Silicon Valley, ce qui

demanderait des investissements conséquents en **design et marketing**, et pas seulement en code, ce qui est loin d'être le cas aujourd'hui.

Il est important de noter que l'utilisation de l'Open Source n'est **pas une panacée**, car cela ne signifie pas automatiquement que l'on soit souverain dans le domaine du numérique. De manière paradoxale, les États-Unis dominent également le marché de l'Open Source, puisque les plus gros contributeurs au monde dans ce domaine sont Google, Microsoft et Meta. Si l'Europe utilise des outils Open Source maintenus à 90% par des développeurs américains, elle reste indirectement dépendante de leur feuille de route. De plus, en raison de son retard en termes de matériel (*hardware*), même si l'Europe possède des logiciels Open Source très performants, si ceux-ci tournent sur des processeurs Intel ou des serveurs AWS, la **souveraineté demeure partielle**.¹⁶⁵ Par conséquent, bien que l'Open Source soit nécessaire pour renforcer la souveraineté numérique de l'UE, il est **insuffisant en lui-même** et ne représente qu'une pièce du puzzle. Il doit être **combiné avec d'autres stratégies** mentionnées dans les sections précédentes, incluant le renforcement des infrastructures numériques (*data centers*, fabriques et « giga-fabriques » d'IA), ainsi qu'une intégration renforcée du marché unique européen, principalement via la stratégie « *Data Union* » et la mise en place d'une union des marchés de capitaux.

2 • Réguler l'IA : Limiter/interdire certains types d'IA, avec la définition de critères de « dangerosité »

Le débat sur la question de limiter ou d'interdire certains types d'IA, avec la définition de critères de « dangerosité », est actuellement au cœur des discussions en France, en Europe et dans le

¹⁶² Ces informations ont été collectées par le biais d'un entretien, la personne concernée ayant souhaité conserver l'anonymat. Voir la liste complète des personnes interviewées à la fin de la Bibliographie.

¹⁶³ European Commission (2021).

¹⁶⁴ Pour chaque euro investi par le secteur public dans l'Open Source, on estime un retour sur investissement multiplié par dix pour l'économie locale. Ainsi, une augmentation de seulement 10 % des contributions à l'Open Source au niveau de l'UE permettrait de créer plus de 600 start-ups et d'accroître le PIB européen d'environ 0,4 % à 0,6%. *Ibid.*

¹⁶⁵ L'initiative RISC-V (architecture de processeur ouverte) est ici un projet clé pour l'Europe.

monde et porte sur le **degré de réglementation nécessaire** pour permettre un **développement responsable de l'IA**. La législation européenne sur l'IA (*AI Act*) a déjà mis en place une telle approche.¹⁶⁶ Les dispositions principales de l'*AI Act* reposent sur une conception dite « **fondée sur les risques** », avec une réglementation juridique spécifiquement adaptée par rapport au degré de risque perçu découlant des différents types d'IA, avec quatre grandes catégories de risques :

- « **Risque inacceptable** » associé aux technologies d'IA interdites : une liste de logiciels d'IA est expressément interdite en raison du risque inacceptable qu'ils présentent pour les droits fondamentaux et la sécurité en général. Cela inclut les techniques de notation sociale basées sur l'IA ou la reconnaissance faciale, ainsi que les pratiques de manipulation subliminale.
- « **Risque élevé** » impliquant une réglementation stricte des pratiques d'IA : cette catégorie vise les logiciels d'IA qui affectent de manière substantielle les droits fondamentaux et/ou la sécurité. Ces pratiques ne sont pas interdites totalement, mais doivent respecter des procédures rigoureuses. Cela comprend des évaluations de conformité, avec une base de données à l'échelle de l'UE sous l'égide de la Commission qui enregistre ces logiciels d'IA « à haut risque » avant leur déploiement.¹⁶⁷
- « **Risque limité** » assorti d'exigences strictes en matière de transparence : cette catégorie regroupe les fameux Chatbots (ChatGPT, Gemini, etc.) et d'autres logiciels d'IA qui interagissent directement avec les personnes.¹⁶⁸ Ces pratiques d'IA sont soumises à des obligations de transparence clairement définies, incluant la mise au point de techniques

(comme le « filigrane numérique ») pour déterminer si un texte, un fichier audio ou une image est produit par l'IA.

- « **Risque minimal** » avec des lignes directrices volontaires sans obligation légale : des lignes directrices volontaires ou « codes de conduite » ont été élaborés afin d'encourager le respect de certaines des règles issues des autres catégories.

Néanmoins, l'*AI Act* contient un certain nombre de lacunes qu'il est essentiel de combler afin de renforcer la loi et les protections qu'elle comporte, par le biais d'une mise en œuvre plus rigoureuse à l'horizon 2028.¹⁶⁹ Parmi celles-ci figure le **manque de souplesse** pour s'adapter à **l'évolution rapide du secteur**, étant donné que toute modification importante de la loi nécessite un accord conjoint du Parlement européen et du Conseil de l'UE, un processus très lent qui peut prendre des années, tandis que l'IA évolue continuellement et à grande vitesse, le secteur étant méconnaissable d'une année sur l'autre. Ainsi, des mécanismes supplémentaires pourraient être introduits pour faciliter l'ajout rapide de nouveaux modèles d'IA dans les différentes catégories, surtout pour les « risques inacceptables » ou les « risques élevés », mais également pour la catégorie « risque limité », sans avoir à passer par une procédure législative. Cela pourrait être mis en œuvre par le **Comité européen de l'IA**¹⁷⁰ par exemple, qui serait ainsi autorisé à réaliser des **misés à jour régulières** sur les différentes catégories, selon les changements rapides dans le secteur, avec peut-être l'instauration de délais réglementaires pour des mises à jour tous les ans.¹⁷¹ De plus, les définitions utilisées pour les différentes catégories pourraient bénéficier d'un **élargissement des critères**, principalement la catégorie

¹⁶⁶ L'*AI Act* a été ratifié par le Parlement et le Conseil de l'UE en décembre 2023 et est entré en vigueur le 1^{er} août 2024, même si la plupart de ses dispositions ne s'appliquent que progressivement au fil du temps.

¹⁶⁷ D'autres exigences incluent la supervision humaine, la transparence, la gouvernance des données, la formation, la gestion des risques et les questions de cybersécurité.

¹⁶⁸ Cela comprend les systèmes qui créent ou gèrent des images, vidéos ou contenus audios.

¹⁶⁹ Barichella (2023a).

¹⁷⁰ Le Comité européen de l'IA, composé de représentants des États membres, de la Commission et d'autres institutions, a pour mission de veiller à ce que la loi soit correctement mise en œuvre au niveau de l'UE.

¹⁷¹ Ces informations ont été collectées par le biais d'un entretien, la personne concernée ayant souhaité conserver l'anonymat. Voir la liste complète des personnes interviewées à la fin de la Bibliographie.

« risque inacceptable » pour la biométrie et la reconnaissance faciale.¹⁷²

En outre, l'AI Act accorde une attention peut-être excessive aux risques de l'IA sur les individus, au détriment des **effets sociétaux plus larges**.¹⁷³

Il existe des situations où l'IA cause des problèmes sociétaux plus généraux, même si les effets directs sur les individus ne sont pas visibles ou mesurables.¹⁷⁴ Les impacts sociétaux de l'IA, tels que les « *fake news* » et la manipulation des processus démocratiques, devraient être mieux pris en compte dans la législation, ce qui implique d'**associer plus étroitement la société civile** dans la gouvernance de l'IA et la mise en œuvre de la loi à tous les niveaux. Cela pourrait prendre la forme d'une section représentant spécifiquement la société civile au sein du Comité européen de l'IA, avec des **évaluations d'impacts sociétaux** pour les nouveaux modèles d'IA, incluant des mécanismes pour mieux diffuser et propager ces impacts vers la société civile. Enfin, probablement la principale lacune de l'AI Act relève des mécanismes de mise en conformité, où l'auto-évaluation par les entreprises est devenue la norme, surtout depuis la loi **Omnibus** visant à simplifier les législations européennes sur le numérique. Les entreprises ont souvent intérêt à exagérer leur niveau de conformité, ce qui remet en cause l'aspect contraignant de la loi. Il existe certaines situations où des évaluations plus rigoureuses sont prévues, surtout pour les catégories de « risque élevé » ou « inacceptable », mais celles-ci sont minoritaires depuis la loi Omnibus, affaiblissant la loi dans son ensemble. C'est pourquoi il est essentiel d'augmenter la proportion d'**évaluations plus rigoureuses par tierce partie** (via des organismes indépendants

et certifiés) sous supervision de la Commission, tout en **diminuant les processus d'auto-évaluation** par les entreprises, en particulier pour les catégories de risques « inacceptables » et « élevés ».¹⁷⁵

3 • Sécurité pour les infrastructures numériques de l'IA, comparaison avec le secteur nucléaire

Le sujet de la **sécurité des infrastructures numériques** (*data centers*, fabrique d'IA) est devenu vital et hautement stratégique. Comme expliqué précédemment, ces infrastructures sont essentielles pour permettre à la France et l'UE d'être compétitives dans la course mondiale à l'IA, notamment pour développer des champions européens et renforcer la souveraineté numérique. À titre d'exemple, l'essor de l'IA a transformé les **data centers** car ils ne sont plus de simples entrepôts de données, mais sont devenus de véritables **usines de calcul intensif**. Cette mutation a introduit des risques de sécurité inédits, tant physiques que numériques (cybersécurité). Pour ces raisons et même s'ils fonctionnent de manière très différente, des comparaisons peuvent être établies entre les **data centers** et les **centrales nucléaires**, en lien avec les risques sécuritaires qu'ils représentent. En effet, le secteur nucléaire a mis au point une **approche de sécurité globale rigoureuse** afin de se prémunir contre les incidents tels que le sabotage, les incendies ou les fuites radioactives. Par exemple, les réseaux de communication et la plupart des équipements sont dupliqués de différentes manières.¹⁷⁶ Plusieurs

¹⁷² En outre, les critères mériteraient d'être plus détaillés, avec un cadre commun et obligatoire à la fois pour les fournisseurs et pour les utilisateurs de modèles d'IA.

¹⁷³ Barichella (2023a).

¹⁷⁴ Comme analysé ci-dessus, l'IA est utilisée pour générer des « *fake news* » sur Internet qui exacerbent la désinformation en ligne ; cela a un impact négatif sur les processus démocratiques, sans provoquer de préjudices directement aux individus.

¹⁷⁵ En effet, comme analysé dans la 1^{re} Annexe, un autre point de consensus qui est ressorti du sondage est que la réglementation actuelle est insuffisante, nécessitant le besoin d'accélérer et de renforcer la mise en œuvre de l'AI Act au niveau européen et national.

¹⁷⁶ Cela comprend l'utilisation d'un système Ethernet, des installations sur site où la redondance est assurée, ainsi que des équipements supplémentaires à distance.

câbles indépendants sont utilisés pour transmettre les informations relatives à l'état des sites nucléaires critiques, et ces informations ne sont considérées comme fiables qu'après avoir été vérifiées à plusieurs reprises. D'autres systèmes de sécurité incluent le « fonctionnement en îlot », grâce auquel une centrale nucléaire parvient, dans 80 % des cas, à s'alimenter en électricité de manière autonome si une attaque coupe l'alimentation, grâce à plusieurs générateurs de secours sur site.¹⁷⁷ Ces dispositifs ont été renforcés à la suite de l'accident de Fukushima en 2011, particulièrement en Europe, où l'accès aux sites nucléaires a été davantage restreint et de nombreux systèmes de contrôle industriel ont été dupliqués hors site avec des configurations différentes.

Ces mécanismes de sécurité physique offrent également une protection solide en matière de **cybersécurité**. Ils contribuent à réduire nettement le risque d'introduction de virus dans les systèmes critiques des centrales nucléaires et à limiter l'ampleur et la propagation des logiciels malveillants lors d'une cyberattaque.¹⁷⁸ Même si les *data centers* bénéficient déjà d'une sécurité importante, celle-ci doit être augmentée au cours des prochaines années. Il n'est évidemment pas possible de simplement copier et coller ce modèle sécuritaire des centrales nucléaires pour les *data centers* ou les fabriques d'IA, car ils ne fonctionnent pas du tout de la même manière. Cependant, le **secteur nucléaire** bénéficie de **plusieurs décennies d'expérience** dans ce domaine, et il est possible de s'en inspirer en **transposant** certaines de ces mesures pour les infrastructures numériques en Europe, avec bien entendu des adaptations. En particulier à l'horizon 2028, une duplication et redondance renforcée des systèmes, incluant une alimentation électrique de secours plus poussée avec un système de « **fonctionnement**

en îlot » intégral comme pour les centrales nucléaires afin d'assurer un approvisionnement en électricité totalement autonome, est pertinent pour les *data centers*.¹⁷⁹ Cela est non seulement lié aux risques sécuritaires, mais également aux difficultés d'approvisionnement avec les réseaux électriques locaux en raison de la grande consommation d'énergie des nouveaux modèles d'IA. De façon générale, il est essentiel de **limiter et restreindre au maximum de façon hermétique** l'accès aux *data centers* en renforçant les contrôles et les fouilles aux points d'ouverture.¹⁸⁰

4 • Bâtir une gouvernance mondiale du numérique avec la création de mécanismes multiniveaux

Depuis le lancement de *ChatGPT* version 3.5 en novembre 2022, de nombreuses voix se sont élevées pour réclamer la mise en place d'un mécanisme de **gouvernance mondiale** pour la réglementation des technologies de l'IA.¹⁸¹ Même certains grands patrons de la tech, comme Sam Altman et Elon Musk, ont mis en garde contre le risque possiblement existentiel posé par l'IA et le besoin d'instaurer un cadre international rigoureux, comme ce fut le cas pour le nucléaire à la fin de la deuxième guerre mondiale. En effet, de nombreux parallèles ont été établis entre l'IA et le nucléaire, notamment en ce qui concerne les risques et les opportunités pour l'humanité. Cependant, l'IA est plus complexe et globale dans son impact. Alors que le nucléaire peut être utilisé comme source d'énergie ou comme une arme, **l'IA est multiforme** avec le potentiel de devenir omniprésente en impactant progressivement tous les secteurs et tous les aspects de la société. Ainsi, même s'il est indéniable que

¹⁷⁷ Desarnaud (2017).

¹⁷⁸ Barichella (2023b).

¹⁷⁹ Ces informations ont été collectées par le biais d'un entretien, la personne concernée ayant souhaité conserver l'anonymat. Voir la liste complète des personnes interviewées à la fin de la Bibliographie.

¹⁸⁰ Certains analystes ont néanmoins mis en garde contre une réglementation trop extrême, qui pourrait constituer un frein à l'innovation. Le but est ainsi de trouver un juste équilibre entre compétitivité et sûreté des *data centers*.

¹⁸¹ Comme expliqué dans la 1^{re} Annexe, le besoin de mettre en place une forme de gouvernance mondiale sur l'IA est un autre sujet consensuel qui est ressorti du sondage.

l'Agence internationale de l'énergie atomique peut servir de source d'inspiration sur certains aspects de la gouvernance mondiale de l'IA, la technologie est trop complexe et multiforme pour être limitée à une dualité énergie-militaire. Pour ces raisons, le traité international de **l'Organisation pour l'interdiction des armes chimiques** (OIAC) offre probablement un cadre de comparaison plus pertinent. En effet, les produits chimiques, comme l'IA, sont de nature complexe, multiforme, avec un **impact global sur la plupart des secteurs** d'activité (santé, industrie, agriculture, construction), ainsi que dans le domaine militaire avec les armes chimiques.¹⁸²

L'OIAC pourrait devenir une source d'inspiration très utile pour la mise en place d'une gouvernance mondiale de l'IA, incluant le développement d'un **cadre de prévision et de surveillance** composé d'**experts internationaux**.¹⁸³ Cependant, il est indéniable que le **contexte international actuellement très instable** rendra toute collaboration sur l'IA difficile à l'horizon 2028. Cela est d'autant plus le cas en raison de la personnalité imprévisible du président américain actuel, qui mène une diplomatie nationaliste et insulaire (surnommée *America First*) avec le retrait de la plupart des organisations ou conventions internationales dont les États-Unis avaient été membre fondateur.¹⁸⁴ **Sans le leadership américain**, il sera compliqué d'avancer concrètement dans la création d'une nouvelle gouvernance mondiale pour l'IA, surtout qu'une grande partie des innovations technologiques dans le secteur proviennent justement des États-Unis. Cela ne veut pas dire qu'aucune mesure ne peut être prise à court terme. En effet, la **Chine**, en tant que deuxième grande puissance de l'IA, a indiqué à plusieurs reprises qu'il est nécessaire

de renforcer la collaboration internationale dans le domaine technologique. Les représentants chinois ont ainsi fait pression pour que l'ONU mette en place de nouveaux cadres de gouvernance mondiale sur l'IA, en collaboration notamment avec l'UE.

Ainsi, l'ONU a franchi des étapes cruciales ces deux dernières années avec le lancement du « **Pacte Numérique Mondial** » (*Global Digital Compact*) adopté en septembre 2024, le premier cadre mondial exhaustif pour la coopération numérique avec l'objectif d'établir des principes partagés pour un futur numérique ouvert, libre et sécurisé. Le pacte prévoit de « gouverner l'IA pour l'humanité » en s'assurant qu'elle respecte les droits de l'homme et aide à atteindre les objectifs de développement durable. En outre, l'ONU a également créé un **panel scientifique international sur l'IA** début 2026 (souvent comparé au GIEC pour le climat), afin de produire des évaluations scientifiques périodiques sur les capacités, les risques et les opportunités de l'IA. Enfin, l'année 2024 fut marquée par l'adoption de la **première résolution mondiale** dédiée exclusivement à l'IA¹⁸⁵ par l'**Assemblée générale de l'ONU**.¹⁸⁶ Cependant, comme la plupart des mécanismes onusiens, ces structures ne sont **pas contraignantes** et ne peuvent prodiguer que des recommandations, aucune sanction n'étant prévue en cas de non-respect des normes. En raison de l'instabilité chronique caractérisant les relations internationales, il sera en toute vraisemblance **impossible à court terme d'établir un cadre de gouvernance mondiale contraignant pour l'IA**, même en cas de changement de président aux États-Unis en 2028.¹⁸⁷ En effet, les grandes puissances sont généralement très réticentes

¹⁸² Barichella (2023a).

¹⁸³ Ces informations ont été collectées par le biais d'un entretien, la personne concernée ayant souhaité conserver l'anonymat. Voir la liste complète des personnes interviewées à la fin de la Bibliographie.

¹⁸⁴ Aux cours des dernières décennies, les États-Unis ont joué un rôle de leadership dans la mise en place de nombreux cadres de gouvernance mondiaux. Cela inclut la création de l'ONU à la fin la seconde guerre mondiale, ainsi que le traité international sur la non-prolifération des armes nucléaires pendant la guerre froide.

¹⁸⁵ Co-parrainée par plus de 120 pays, cette résolution appelle à combler le fossé numérique entre pays développés et en développement.

¹⁸⁶ Cela a permis la création d'un « Dialogue Mondial sur l'IA », un forum régulier au sein de l'ONU pour harmoniser les approches nationales et assurer l'interopérabilité des réglementations entre les pays.

¹⁸⁷ Même lorsque le contexte international était plus favorable, par exemple sous Obama lors des négociations pour l'Accord de Paris en 2015, aucun mécanisme contraignant n'avait été intégré dans le nouveau régime climatique mondial.

à l'idée de transférer une part de leur souveraineté aux institutions internationales, et très peu de mécanismes onusiens sont ainsi contraignants.¹⁸⁸

Pour ces raisons, la gouvernance mondiale de l'IA au cours des prochaines années a plus de chance d'aboutir par le biais de projets moins ambitieux, comme par exemple les **normes ISO** ou les **coalitions inter-régionale**. Les normes ISO (Organisation internationale de normalisation) constituent des points de repère mondiaux. Elles établissent des exigences, des spécifications ou des lignes directrices, pour garantir que des produits, processus ou services sont adaptés en matière de qualité et de sécurité. Elles sont **pratiquement universelles** car élaborées par des experts de 170 pays, mais sont de nature **volontaire** car toute entreprise choisit librement de s'y conformer, et procurent une crédibilité en servant de preuve de sérieux face aux clients et régulateurs. L'IA étant une technologie « boîte noire » par nature, les normes ISO peuvent aider à y injecter de la transparence et la confiance du public, prouvant que l'IA a été testée selon des standards internationaux rigoureux. En outre, les normes ISO facilitent l'interopérabilité en faisant en sorte que deux systèmes d'IA puissent échanger des données sans friction, et avec la conformité réglementaire.¹⁸⁹ Un **comité technique mondial** a été créé pour le développement de normes ISO spécifiquement dans le domaine de l'IA, incluant la gouvernance (ISO/IEC 42001), la gestion des risques (ISO/IEC 23894), la qualité des données (ISO/IEC 5259), ainsi que l'éthique et les biais (ISO/IEC TR 24027). À l'horizon 2028, ces normes ISO doivent être plus **systématique-**

ment intégrées au sein de l'AI Act ainsi que les législations contraignantes dans les domaines connexes au niveau national.

En ce qui concerne les **coalitions inter-régionales**, l'AI Act européen est une législation pionnière, l'UE étant la première grande entité politique à mettre en place un cadre réglementaire global pour les différents usages de l'IA, qui a servi de **source d'inspiration** pour de nombreux pays à travers le monde. Cela comprend des pays développés comme le Canada, l'Australie et le Japon, ainsi que plusieurs pays membres des BRICS comme le Brésil, l'Inde ou l'Afrique du Sud, qui ont indiqué s'être servis de l'AI Act pour la mise en place de leurs propres lois nationales.¹⁹⁰ L'UE a déjà utilisé par le passé sa position en tant que **plus grand marché unique au monde**, avec des opportunités de commerce lucratives, afin de faire pression lors des négociations commerciales pour l'inclusion de standards environnementaux ambitieux, ainsi que des règles pour la protection des données dans le cadre du RGPD.¹⁹¹ À l'horizon 2028, l'UE peut et doit exiger la même chose en ce qui concerne les **standards européens pour l'usage responsable de l'IA**, tels qu'ils sont mis place avec l'AI Act, d'autant plus que la loi européenne contient plusieurs clauses ayant une **dimension extraterritoriale**.¹⁹² L'Europe est donc idéalement positionnée pour mener la communauté internationale dans la création de **coalitions inter-régionales sur l'éthique de l'IA**, incluant à la fois des cadres bilatéraux et multilatéraux, avec ses principaux partenaires commerciaux dans le monde entier, notamment le Royaume-Uni, la Turquie, le Canada, le Japon, l'Afrique du Sud, le Mercosur et même l'Inde.¹⁹³

¹⁸⁸ Les seuls exemples qui existent sont le Conseil de sécurité de l'ONU, mais les cinq membres permanents parviennent rarement à se mettre d'accord ; en outre, l'Organisation mondiale du commerce (OMC) contient certains mécanismes contraignants, mais ceux-ci sont assez limités dans l'ensemble.

¹⁸⁹ A titre d'exemple, l'AI Act s'appuie fortement sur les normes ISO ; être certifié ISO, c'est avoir déjà fait une part importante du chemin vers la conformité légale en Europe.

¹⁹⁰ La Chine a également indiqué avoir pris en considération la législation européenne, même si le contexte est très différent en raison du système politique autoritaire chinois.

¹⁹¹ En effet, lors des négociations pour les accords commerciaux avec des pays comme le Canada, le Japon, l'Inde et même la Chine, l'UE a conditionné l'accès de son marché unique à l'inclusion de ces standards éthiques.

¹⁹² De la même manière que pour le RGPD, l'AI Act prévoit que toute entité qui met en place des biens ou des services utilisant les technologies de l'IA sur le marché européen, sera soumise à la loi européenne, même si cette entité n'est pas elle-même située sur le territoire de l'UE.

¹⁹³ Ces informations ont été collectées par le biais d'un entretien, la personne concernée ayant souhaité conserver l'anonymat. Voir la liste complète des personnes interviewées à la fin de la Bibliographie.

IV. Déclinaison des éléments précédents en trois scénarios prospectifs contrastés à moyen terme (horizon 2035)

Les **trois scénarios prospectifs à l'horizon 2035** dans cette dernière partie ont été construits sur la base des résultats quantitatifs issus du **sondage**, ainsi que sur les informations qualitatives collectées lors des **entretiens** ou dans la **littérature**. Les scénarios à moyen terme proposés ci-dessous sont exhaustifs, et incluent l'ensemble des différentes sections du sondage - voir la 1^{re} Annexe pour une analyse détaillée des résultats.

1 . Scénario optimiste à l'horizon 2035

I IMPACT SUR LE TRAVAIL, LES MÉTIERS ET LES COMPÉTENCES

Tout d'abord, ce scénario **optimiste** se base sur l'hypothèse d'une **accélération continue** en termes du **développement technologique** de l'IA, où des techniques efficaces sont trouvées pour contourner les problèmes identifiés dans la première partie de l'étude en lien avec le manque de nouvelles données pour entraîner les

modèles d'IA. Cela implique notamment le développement de données synthétiques hybrides élaborées qui imbriquent et enchevêtrent de manière complexe les données humaines avec celles de synthèse, ainsi que d'autres méthodes innovantes très sophistiquées,¹⁹⁴ permettant de continuellement renouveler les stocks de données et **d'éviter ainsi le piège de la stagnation des modèles d'IA**.¹⁹⁵

Dans ce scénario optimiste, l'accélération technologique continue permet à **l'humain d'être augmenté par l'IA plutôt que d'être remplacé**. Même si nous assistons à des pertes d'emplois dans certains secteurs, **plus d'emplois sont créés que ne sont détruits**, surtout dans les domaines émergents en lien avec les nouvelles technologies. Ainsi, grâce à une formation adéquate soutenue par les gouvernements, les travailleurs dans tous les secteurs de l'économie apprennent à se servir de l'IA pour augmenter de manière significative leurs capacités de production, accélérant la réalisation des différentes tâches au sein des entreprises, des organisations publiques, des universités, etc. Plutôt qu'un remplacement de l'humain par l'IA, nous assis-

¹⁹⁴ Au jour d'aujourd'hui, en plus des données de synthèse, d'autres techniques sont à l'étude ou en cours d'élaboration pour renouveler les stocks de données de l'IA, qui pourraient apporter des percées majeures au cours de la prochaine décennie. Cela inclut l'apprentissage « actif » (*active learning*) ou fédéré (*federated learning*), les pipelines de données en temps réel (IA Adaptive), ainsi que les recherches internet avancées à l'aide des IA agentiques (*Retrieval-Augmented Generation*).

¹⁹⁵ Les défis énergétiques sont également surmontés dans ce scénario, pour les raisons analysées ci-dessous.

tons à une **fusion entre l'homme et la machine** qui deviennent des binômes inséparables et complémentaires se renforçant mutuellement. L'IA réalisant rapidement toutes les tâches répétitives et récurrentes, cela permet aux humains de se concentrer sur les tâches où le contact humain et personnel, la créativité et l'ingéniosité stratégique sont nécessaires. Ce duo gagnant-gagnant, où l'IA et l'humain se concentrent chacun sur les domaines où ils peuvent apporter la meilleure contribution possible, transforme le travail, les métiers et les compétences en profondeur, mais de façon positive où **l'humain et la société toute entière en ressortent renforcés**. Dans ce scénario optimiste, l'accélération technologique continue de l'IA se propage dans tous les pays du monde, augmentant de façon substantielle la **productivité, la croissance économique et la richesse au niveau mondial**. Cela permet à l'humanité d'entrer dans une nouvelle ère de développement économique comparable aux « trente glorieuses » de l'après-guerre, avec le retour du plein emploi et un taux de forte croissance au niveau mondial supérieur à 5% par an. La richesse et la croissance économique générée par l'IA ruissellent du haut vers le bas au-delà des chefs d'entreprises, bénéficiant à l'ensemble de la société. Cela entraîne la **réduction des inégalités sociales** entre les classes aisées, moyennes et défavorisées, inversant ainsi une tendance qui avait commencé dans les années 1970 avec le creusement des inégalités dans de nombreux pays occidentaux. En outre, cette situation profite également aux pays en développement. En effet, grâce au boom économique généré par l'IA, ceux-ci regagnent des hauts niveaux de croissance leur permettent de continuer à rattraper leur retard, réduisant ainsi les inégalités entre les pays du Nord et ceux du Sud global. Pour ces raisons, la **prospérité mondiale** atteint un niveau sans précédent, et l'IA permet un « **miracle économique** » avec peu de parallèles dans l'histoire de l'humanité.

I IMPACT SUR L'ÉCOLOGIE

Dans ce scénario optimiste, malgré les craintes suscitées par la consommation excessive en électricité des nouveaux modèles d'IA, des mécanismes sont trouvés pour **limiter les impacts sur l'environnement**. Une utilisation efficace des **énergies renouvelables** permet d'alimenter les infrastructures numériques avec l'hydro-électrique, le solaire, l'éolien, la géothermie, ainsi que l'hydrogène. De nouvelles percées technologiques permettent de mieux capter ces énergies propres, notamment des nouveaux types de batteries aidant à stocker l'énergie plus efficacement en raison de l'intermittence des renouvelables. L'IA elle-même permet d'accélérer et de **stimuler la recherche scientifique** dans ce domaine, avec la découverte de nouvelles techniques pour limiter le changement climatique, y compris dans le secteur des énergies renouvelables mais pas uniquement. D'autres domaines de progression incluent les technologies pour la capture et le stockage du CO₂, où l'IA apporte des percées technologiques majeures permettant de réduire considérablement les coûts et d'augmenter l'efficacité des instruments pour capter et stocker le carbone, ainsi que ceux associés aux énergies renouvelables.¹⁹⁶ Dans ce scénario optimiste, les **technologies bas-carbone se propagent** dans le monde entier car elles sont compétitives sur le plan économique et efficaces sur le plan environnemental, accélérant la transition écologique. Cela aide les pays développés et en développement à mettre en œuvre plus rapidement leurs engagements dans le cadre de l'Accord de Paris, ainsi que de rehausser leur niveau d'ambition pour les « contributions déterminées au niveau national » (CDN). Pour ces raisons, la communauté internationale se rapproche à l'horizon 2035 des objectifs de long terme fixés par l'**Accord de Paris** de maintenir le réchauffement climatique en-deçà des 2°C d'ici la fin du siècle, car l'agrégation de toutes les CDN des pays du monde entier passe pour la première fois en dessous de ce seuil. Même si la mise en œuvre rigoureuse des CDN reste un défi de taille, la compétitivité des technologies

¹⁹⁶ Cela offre une période de transition très précieuse qui permet à la communauté internationale de continuer à utiliser les infrastructures fossiles, en attendant la transition complète vers les renouvelables.

bas-carbone grâce aux **percées technologiques liées à l'IA** entraîne progressivement l'économie mondiale dans une transition hors des énergies fossiles. Enfin, l'IA permet même d'atteindre les prémices de la **fusion nucléaire** plus tôt que prévu en aidant à résoudre des obstacles technologiques insurmontables pour l'humain seul. Cela ouvre pour la première fois la perspective d'un **avenir énergétique 100% décarboné** pour l'humanité avec une énergie propre potentiellement illimitée, **sauvant ainsi la planète** des effets du changement climatique. La fusion nucléaire décuplerait dans le même temps les perspectives de développement et de progrès pour la société dans virtuellement tous les domaines.

I IMPACT SUR LES PROCESSUS DÉMOCRATIQUES

Dans ce scénario optimiste à l'horizon 2035, les prévisions sur les impacts négatifs de l'IA en termes de désinformation et de manipulation des processus démocratiques ne se concrétisent pas. En effet, des nouveaux modèles d'IA plus performants permettent de **détecter et contrer** les « *fake news* » et les « *deepfakes* » sur Internet et les réseaux sociaux, **protégeant les populations** face aux tentatives de manipulation extérieures. Cela inclut par exemple le développement de nouveaux systèmes de **filigrane numérique** (« *watermarking* ») ultra-performants permettant d'identifier en temps réel lorsque des hackers tentent de propager la désinformation par le biais de l'IA. Lorsqu'ils sont identifiés, les services de l'État concernés, notamment le VIGINUM en France et l'équivalent dans d'autres pays, peuvent rapidement les bloquer et les mettre hors d'état de nuire sur la toile. Bien entendu, il y aura à l'avenir comme aujourd'hui un jeu de chat et de souris entre ceux chargés de la protection de l'information en ligne, et les hackers potentiels. Néanmoins, grâce à une collaboration étroite entre les gouvernements et les entreprises du secteur, notamment par le biais de conseils d'administration conjoints et de partenariats public-privés leur permettant de travailler de manière rapprochée, les services de l'État auront la plupart du temps une longueur d'avance sur le plan technologique. Cela passe

par une **règlementation rigoureuse** qui intègre des mécanismes de sûreté dès la conception des modèles d'IA, garantissant leur **auditabilité** et la **correction des biais** en continu sur l'ensemble du cycle de vie des modèles. Pour ces raisons, dans ce scénario optimiste, l'IA utilisée à bon escient et de manière responsable permet de mieux **protéger les démocraties** en Europe et à travers le monde, réduisant ainsi l'impact des campagnes de désinformation sur les processus électoraux. Cela aide à **contenir la montée du populisme** et des extrêmes politiques lors des élections à tous les échelons, favorisant ainsi les partis politiques modérés, sécurisant et renforçant les démocraties de l'intérieur. La distinction entre le vrai et le faux pour les informations accessibles sur Internet devient plus la facile et directe, réduisant ainsi le potentiel de manipulation politique. Grâce à cela, le niveau de polarisation dans l'électorat diminue nettement, avec le retour en force des partis politiques modérés et la réduction des fractures internes au sein des démocraties à travers le monde.

I IMPACT SUR LA SOCIÉTÉ EN GÉNÉRAL

Dans ce scénario optimiste, l'accélération du développement technologique de l'IA apporte des **progrès majeurs en médecine** et dans la **recherche scientifique** à l'horizon 2035. Les modèles d'IA avancés permettent de faire en quelques jours ou semaines ce que des humains faisaient en plusieurs années voire plusieurs décennies, notamment en ce qui concerne la compilation, le traitement et l'analyse d'une quantité massive de données. Grâce à cela, la recherche scientifique est stimulée de manière exponentielle, avec un effet en cascade où une découverte importante dans un secteur entraîne une autre dans un secteur connexe, avec un cycle d'auto-renforcement qui s'accélère et monte en puissance. La promesse de Sam Altman que **l'IA pourra guérir le cancer devient réalité**, avec de nombreux nouveaux médicaments et traitements découverts permettant de guérir des maladies longtemps incurables ou difficilement traitables. Pour ces raisons, l'IA apporte un **allongement de la durée et de la qualité de vie** en Europe et partout dans le monde. Ainsi, l'humanité vie mieux et plus longtemps en bonne

santé, améliorant les indices de bonheur et de bien-être. Cela est d'autant plus le cas puisque, dans ce scénario, l'IA est utilisée à bon escient par les gouvernements pour résoudre un certain nombre de problèmes sociétaux.

En effet, les progrès spectaculaires réalisés en médecine sont répliqués dans la plupart des domaines scientifiques, permettant à l'humanité d'entrer dans une phase de **progrès technologiques sans précédent**. Par exemple, l'IA permet de développer de nouveaux systèmes accessibles pour filtrer l'eau et pour rendre l'agriculture plus résiliente et abondante tout en diminuant les coûts¹⁹⁷, réduisant ainsi la faim et la soif dans les pays en développement. Dans les pays développés, des robots habiles programmés avec des IA avancées permettent de construire massivement de nouveaux logements à un prix réduit et plus rapidement par rapport à la main d'œuvre humaine, permettant de réduire considérablement le nombre de sans-abris ou d'habitats précaires. Dans ce scénario optimiste, les gouvernements dans la plupart des pays développés, ainsi que dans de nombreux pays en développement, parviennent à mettre en place des **cadres réglementaires nationaux rigoureux**, ainsi qu'une **gouvernance mondiale** sous l'égide de l'ONU, permettant un **développement responsable de l'IA**. Ainsi, l'AI Act européen encourage des lois comparables dans les pays du monde entier, avec une charte mondiale ratifiée par l'Assemblée générale et même par le Conseil de sécurité de l'ONU, où l'IA est régulée de manière éthique. Certains types d'IA à haut risque sont bannis complètement et d'autres comportant des risques élevés sont réglementés de façon stricte, avec des sanctions sévères au niveau international, régional et national en cas de non-respect. Pour ces raisons, la communauté internationale parvient à **garder le contrôle sur la technologie**, et l'IA demeure ainsi sous tutelle humaine et est utilisée pour servir le bien-être et le progrès de l'humanité.

2 • Scénario pessimiste à l'horizon 2035

I IMPACT SUR LE TRAVAIL, LES MÉTIERS ET LES COMPÉTENCES

Ce second scénario, **pessimiste**, est en **opposition totale** par rapport au premier scénario optimiste, à l'exception du fait qu'ils se basent tous les deux sur l'hypothèse d'une accélération continue du développement technologique de l'IA à l'horizon 2035. Dans ce second scénario, **les humains sont graduellement remplacés par la machine dans la plupart des secteurs**, car les entreprises se retrouvent dans un cercle vicieux de concurrence mondialisée. Cela les oblige à rechercher l'efficacité et la rapidité pour faire baisser les coûts vis-à-vis de leurs concurrents, l'IA étant devenue beaucoup plus performante en termes de vitesse et de résultats dans pratiquement tous les domaines que les **humains sont rendus progressivement obsolètes**. Cela inclut les métiers « à col blanc » les plus vulnérables mentionnés précédemment dans l'étude (administration et support de bureau, marketing, traducteurs ou designers web, etc.), où l'IA devient tellement plus rapide et performante, que ces métiers disparaissent presque totalement d'ici 2035. Cela bouleverse en profondeur le marché du travail avec des **pertes d'emplois massives** dans le secteur tertiaire, notamment dans des domaines clés comme les services marchands et non marchands qui employaient beaucoup de personnes jusque-là (secteur bancaire, finance, assurance, immobilier/hôtellerie, etc.). Seuls les cadres de direction, les actionnaires et certains métiers très spécialisés ou techniques survivent à ce véritable tsunami de l'emploi.¹⁹⁸ Même dans ces cas de figure, **l'IA bouleverse en profondeur la nature des métiers et des compétences**, avec une place de plus en plus importante pour la machine, et de moins en moins perceptible pour l'humain. En raison de ces gains d'efficacité et de compéti-

¹⁹⁷ Pour l'eau, cela inclut à la fois des nouvelles techniques pour purifier l'eau et pour la désalinisation. Concernant l'agriculture, cela comprend des innovations (machines, engrais) pour rendre les cultures plus résistantes notamment face aux sécheresses, ainsi que pour multiplier les récoltes en termes de la quantité produite.

¹⁹⁸ Ainsi, dans ce scénario pessimiste, ce sont principalement les grands patrons de la tech qui bénéficieront de la révolution de l'IA en continuant de s'enrichir aux dépens de la société dans son ensemble.

tivité, nous assistons même à l'émergence de scénarios où la direction est entièrement prise en charge par des nouveaux modèles d'IA dotés de capacités de raisonnement. Cela entraîne des situations où l'IA agentique dirige et prend des décisions sur d'autres IA au sein des entreprises et des administrations, **l'humain ayant perdu son rôle, sa place et sa pertinence sur le marché du travail.**¹⁹⁹

De plus, en raison de progrès technologiques continues, l'IA commence même à remplacer les humains dans les secteurs considérés comme étant protégés de l'automatisation, notamment les métiers manuels ou ceux impliquant « l'intelligence relationnelle » (*soft skills*). Dans ce scénario, les nouveaux modèles d'IA de type robotique parviennent à acquérir une certaine dextérité manuelle leur permettant d'interagir avec leur environnement en trois dimensions, pouvant réaliser des tâches jugées impossible jusque-là. Ils commencent même à rivaliser avec les humains en termes d'efficacité et de rapidité pour des métiers comme la plomberie, la maçonnerie, la menuiserie ou le jardinage. De façon similaire, les services à la personne et la santé, incluant l'accompagnement et le soutien psychologique, et même les professeurs dans l'enseignement sensés transmettre des valeurs et servir de guide moral, sont concurrencés par des IA très avancées capables d'imiter avec précision les émotions humaines. Même si les humains ne sont pas complètement remplacés, cela **bouleverse en profondeur les métiers** dans ces secteurs. Cette situation entraîne la montée d'un **chômage de masse** non seulement dans les pays développés, mais également ceux en développement, **augmentant considérablement les inégalités et la pauvreté**, et générant de **graves troubles sociaux** au niveau mondial que les gouvernements ont beaucoup de mal à contenir et à gérer (analysé ci-dessous).

I IMPACT SUR L'ÉCOLOGIE

En raison de l'accélération du développement technologique, les **nouveaux modèles d'IA deviennent de plus en plus énergivores** au cours de la prochaine décennie, atteignant une **situation critique à l'horizon 2035**. Cela entraîne des cas de figure extrêmes où des régions entières au sein de certains pays font face à des *blackout* ou **panne de courant d'ampleur** de façon régulière, car les réseaux électriques ne peuvent pas faire face à la demande en énergie massive des nouveaux modèles d'IA avancés. Des choix difficiles doivent être faits par les gouvernements concernant l'allocation en électricité, et en raison de la pression exercée par les lobbys, certains services publics à priori essentiels sont parfois sacrifiés au prétexte de la compétitivité. Les transports, la police, les administrations, voire même les hôpitaux et les écoles, font de plus en plus face à des pannes de courant importantes, les empêchant de fournir leurs services de manière constante à la population. Les citoyens eux-mêmes sont confrontés à des situations parfois difficiles, où certains quartiers moins favorisés dans les villes, ainsi que certaines zones rurales, n'ont **plus d'accès régulier à l'électricité**, induisant des risques sanitaires et sécuritaires notables. Au niveau mondial, la consommation de l'IA en électricité augmente de manière exponentielle pour atteindre un seuil critique à l'horizon 2035, devenant de loin le secteur le plus énergivore et ainsi celui responsable de la majorité des émissions de GES, loin devant les transports, le chauffage, l'agriculture ou les bâtiments (les plus polluants en 2026). Les progrès techniques pour les énergies renouvelables ne sont pas suffisants pour compenser une telle augmentation, ce qui entraîne un **accroissement significatif des émissions de GES** dans la plupart des pays du monde.

Ainsi, la « décennie cruciale » (2026-35) pour diminuer les émissions et mettre la communauté internationale sur une trajectoire positive pour atteindre la neutralité carbone en 2050 devient un échec total, avec le scénario inverse se profi-

¹⁹⁹ À ce sujet et comme indiqué dans la 1^{re} Annexe, il est notable comment le sondage fait apparaître un rejet unanime de l'IA comme décideur autonome complet à horizon 2035, avec l'adhésion au modèle de comité hybride avec supervision humaine.

lant impliquant une augmentation substantielle des émissions sur cette période. Pour ces raisons, **l'atteinte des objectifs de l'Accord de Paris** pour contenir le réchauffement climatique en-deçà des 2°C d'ici la fin du siècle **devient impossible, rendant l'Accord caduc**. À l'instar de Trump aux États-Unis, de plus en plus de pays choisissent même de s'en retirer par souci de compétitivité pour ne pas entraver le développement technologique de l'IA. Par conséquent, nous assistons dans ce scénario pessimiste à **l'accélération des effets du changements climatique**, qui arrivent **plus tôt** que les prévisions des scientifiques et entraînent une augmentation de la fréquence et de l'intensité des **phénomènes climatiques extrêmes**. Cela inclut les ouragans, tornades, typhons, tempêtes, sécheresses, ainsi que la montée des océans, causant des **dégâts catastrophiques** aux populations civiles dans les pays du monde entier. Pour faire face à la demande massive en électricité, de plus en plus de gouvernements font le choix de l'énergie nucléaire, parfois dans la précipitation et sans mettre en place des mesures sécuritaires appropriées. En plus d'augmenter considérablement le volume des déchets nucléaires avec leurs impacts sur l'environnement, cela entraîne une augmentation du nombre « **d'accidents** » dans les **centrales nucléaires**, à l'instar de Tchernobyl ou Fukushima. Cette situation est exacerbée par l'augmentation des phénomènes météorologiques extrêmes induits par l'accélération du changement climatique, rendant les centrales nucléaires plus vulnérables, et causant des catastrophes avec de nombreuses **victimes civiles**.

I IMPACT SUR LES PROCESSUS DÉMOCRATIQUES

Dans ce scénario pessimiste, les nouveaux modèles d'IA avancés contribuent à **démultiplier les campagnes de désinformation** et les manipulations des processus électoraux, **affaiblissant les démocraties** à l'horizon 2035. Malgré des avancés dans les technologies pour détecter la désinformation en ligne telles que les filigranes numériques (« *watermarking* »), les gouvernements **ne parviennent pas à mettre en place des réglementations suffi-**

samment strictes pour maîtriser les contenus sur Internet et les réseaux sociaux, comme l'auditabilité des systèmes et la correction des biais, notamment en raison de la pression exercée par les lobbys du secteur. Les gouvernements et les entreprises ne réussissent pas à créer des partenariats publics-privés robustes pour permettre l'intégration de mécanismes de sûreté dès la conception des modèles d'IA, entraînant **l'échec de la correction des biais en ligne** et la traçabilité des systèmes. Dans ce scénario pessimiste, ce sont donc les *hackers* qui obtiennent une longueur d'avance sur les agences nationales de protection des systèmes d'information et parviennent à les distancer, le jeu du chat et de la souris tournant clairement en leur faveur avant d'atteindre un point critique en 2035. Les pouvoirs publics **perdent ainsi totalement le contrôle sur Internet et les réseaux sociaux**, avec de graves conséquences en termes d'amplification de la désinformation en ligne. En effet, sans contraintes réglementaires pour limiter leurs actions, les *hackers* auront le champ libre pour amplifier à la fois le nombre et la magnitude des « *fake news* » et des « *deepfakes* » sur Internet en se servant de nouveaux modèles d'IA avancés, capables de lancer des campagnes de désinformation à très grande échelle et de manière autonome. Dans ce scénario pessimiste, les **démocraties sont minées de l'intérieur** par des manipulations massives de l'opinion public via la désinformation en ligne, où la **distinction entre le vrai et le faux devient presque impossible**. Cela contribue à exacerber le niveau de **polarisation** dans l'électorat et favorise la montée des partis antisystème, résultant en l'élection de figures populistes issues des extrêmes politiques à la tête des grandes démocraties en Europe, en Amérique du Nord et ailleurs. Ceux-ci mettent en place des programmes politiques semi-autoritaires et liberticides qui **affaiblissent la démocratie et les libertés**. Cette situation contribue à une détérioration des régimes démocratiques à travers le monde, avec le basculement vers des systèmes de type « **démocraties illibérales** » comme la Turquie d'Erdoğan, où les élections sont maintenues mais où les contre-pouvoirs ont été considérablement érodés.

I IMPACT SUR LA SOCIÉTÉ EN GÉNÉRAL

Pour ces raisons, ce scénario pessimiste contribue à exacerber les problèmes sociaux et les conflits à l'horizon 2035. En effet, l'arrivée d'un **chômage de masse** en raison de l'IA bouleverse les sociétés à travers le monde. Les gouvernements dans les pays développés et en développement tentent tant bien que mal de gérer une **situation hors de contrôle**. Les options incluent l'introduction de pensions publiques pour faire face à la démultiplication des nouveaux chômeurs (notamment des tentatives pour mettre en place un revenu universel de base - RUB), financées par des taxes sur l'IA et les robots. Même si ces mesures représentent de bonnes idées dans l'absolu (voir sections précédentes), elles deviennent difficilement réalisables dans le contexte d'un chômage de masse, où les gouvernements ne parviennent pas à taxer suffisamment les entreprises en raison de la pression exercée par les lobbys. Les fonds qui sont levés ne sont pas suffisants pour financer les pensions publiques et le RUB par rapport à l'augmentation exponentielle du nombre de demandeurs d'emplois. Dans ce scénario pessimiste, l'IA entraîne non seulement une **marche forcée vers une société post-travail** où une majorité de citoyens n'a plus d'activité, mais également vers une **précarisation généralisée de la société** car les gouvernements n'ont pas les ressources pour financer les transferts sociaux nécessaires.²⁰⁰ Par conséquent, cela entraîne une aggravation significative des inégalités et de la pauvreté, avec une augmentation correspondante de la **délinquance** et de la **criminalité**, où de plus en plus de personnes se tournent vers des activités extra-légales comme étant la seule manière de subvenir à leurs besoins vitaux. Dans ce scénario du pire, certains pays en développement, qui étaient déjà fragiles au point de départ, s'effondrent complètement et basculent dans **l'anarchie** voire la **guerre civile**. Dans les pays développés, les tensions sociales sont démultipliées pour atteindre un niveau critique en 2035. Cela contribue dans certains cas à **achever les régimes démocratiques**, déjà très affaiblis par

l'augmentation de la désinformation, et qui basculent vers **l'illibéralisme**, voire dans une forme de **semi-autoritarisme** comme la seule manière d'éviter l'anarchie.

Dans ce scénario pessimiste, l'accélération du changement climatique causée par les nouveaux modèles d'IA très énergivores (analysé précédemment), contribue à exacerber ces points de tensions à l'intérieur des pays, ainsi que les tensions entre les nations. Cela est particulièrement le cas pour les pays en développement en raison de la raréfaction des ressources naturelles (notamment l'eau), surtout dans les zones comme l'Afrique sub-saharienne, le Moyen-Orient ou le bassin de l'Indus, historiquement très exposés aux sécheresses. Cette situation génère un cocktail explosif de circonstances, avec des **risques de conflits armés** dans ces régions volatiles où les gouvernements ne parviennent pas à faire face à l'accumulation des problèmes existentiels. L'instabilité augmente sur le plan international et géopolitique, et voit les nations du monde se replier sur elles-mêmes, avec la montée du nationalisme. Cela entraîne **l'effondrement des cadres de collaboration international** dans la plupart des domaines (normes ISO, accords ONU, coalitions inter-régionales), y compris pour l'IA. Ainsi, la communauté internationale échoue dans la mise en place de cadres de gouvernance mondiaux rigoureux pour le développement responsable de l'IA. Pour ces raisons, **l'humanité perd progressivement le contrôle sur la technologie**, avec potentiellement de graves conséquences politiques, économiques et sociales, qui ne font qu'exacerber les problèmes mentionnés ci-dessus. Cela provoque des crises sans précédent qui **compromettent l'avenir de la civilisation humaine**.

²⁰⁰ Cela est d'autant plus le cas que les gouvernements sont déjà très endettés au point de départ en 2026.

3 • Scénario intermédiaire ou mixte à l'horizon 2035

I IMPACT SUR LE TRAVAIL, LES MÉTIERS ET LES COMPÉTENCES

Le troisième scénario proposé est **mixte** ou **intermédiaire**, car il reprend certains éléments issus à la fois du premier et du second scénario. Il se veut **réaliste**, car l'impact de l'IA à l'horizon 2035 sera vraisemblablement nuancé, combinant certains aspects positifs avec d'autres plutôt négatifs. En outre, ce scénario intermédiaire se base également sur l'hypothèse que même si les progrès technologiques de l'IA vont probablement continuer, **certaines limites** pourraient être atteintes, à la fois sur la plan technologique (manque de nouvelles données) et énergétiques (pas assez d'électricité pour entraîner les modèles d'IA avancés). Cela provoquerait un **ralentissement relatif de la vitesse du développement de l'IA**, surtout par rapport aux premières années (2022-26).²⁰¹ Ce ralentissement relatif aurait pour conséquence des impacts **moins extrêmes** et **plus nuancés** de l'IA sur la société, corroborant ainsi ce troisième scénario. Dans ce cas de figure, l'IA entraîne des **effets contrastés** sur le travail, les métiers et les compétences. D'un côté, la technologie permet **d'augmenter la productivité** dans de nombreux secteurs vitaux pour l'économie, créant également de nouveaux emplois dans les domaines en lien avec l'IA tels que l'informatique et la gestion des données. Les gouvernements investissent dans la formation aux compétences de l'IA, permettant aux travailleurs d'être augmentés dans certains secteurs en utilisant la technologie pour accroître leur vitesse de travail et leur productivité. Cela génère des bénéfices notables pour l'économie au niveau national et mondial, stimulant la **croissance économique** et augmentant le PIB. D'un autre côté, un **nombre non-négligeable d'emplois sont automatisés** dans d'autres secteurs et les pertes commencent à s'accroître à l'horizon 2035, avec la

balance entre création et destruction d'emplois devenant de plus en plus difficile à discerner.

Même lorsque les emplois ne disparaissent pas complètement, les métiers « à col blanc » sont souvent transformés en profondeur par la technologie, et beaucoup de cadres ont du mal à s'adapter suffisamment rapidement à ces bouleversements importants. Ainsi, plutôt que la fin du travail, **l'IA transformera probablement le marché du travail à l'horizon 2035** comme les révolutions industrielles aux siècles précédents, créant des emplois dans certains secteurs, en supprimant dans d'autres, et modifiant la nature des métiers et les compétences en profondeur dans la plupart des domaines. En outre, ce sont surtout les jeunes diplômés entrant sur le marché du travail qui auront le plus de mal s'intégrer dans la vie active, car les nouveaux modèles d'IA avancés permettront de réaliser la plupart des tâches répétitives préalablement confiées aux stagiaires. Cela contribuera à **augmenter le chômage chez les jeunes**, tout en créant de nouveaux emplois pour certains profils plus expérimentés et techniques. Malgré les politiques des États pour accompagner la reconversion professionnelle dans les secteurs où l'IA permettra une expansion, un pourcentage non-négligeable de travailleurs plutôt seniors ne **parviendront pas à s'adapter à cette nouvelle révolution technologique**, et seront ainsi laissés pour compte sur le marché du travail, perdant définitivement leur emploi ou étant relégués à des tâches mal rémunérées. Pour ces raisons, les effets de l'IA pourraient être **paradoxaux**, entraînant dans le même temps **l'accroissement de la productivité** et de la croissance économique mondiale en termes de PIB, tout en **augmentant les inégalités sociales** au sein des pays.²⁰² Se faisant, l'IA ne fait que renforcer des **tendances qui sont déjà visibles dans le monde en 2026** et depuis les dernières décennies pour un certain nombre de pays, à commencer par les États-Unis. En effet, même si le PIB américain a beaucoup augmenté depuis les années 1980, les inégalités sociales se sont considérable-

²⁰¹ En effet, ces premières années avaient été marquées par l'enthousiasme suscité par la nouveauté technologique.

²⁰² Comme indiqué dans la 1^{re} Annexe, un point de consensus dans le sondage est lié aux inquiétudes profondes sur le risque de renforcement des inégalités sociales en raison de l'IA.

ment creusées dans le même temps,²⁰³ avec une répartition très inégale des fruits de cette croissance.²⁰⁴ Selon ce scénario intermédiaire, la révolution de l'IA pourrait contribuer à exporter ce modèle économique américain de type **néolibéral** dans de nombreux autres pays, y compris en Europe où les inégalités sociales n'avaient jusqu'alors pas augmenté autant qu'aux États-Unis.²⁰⁵ Ce scénario mixte présente donc à la fois des **avantages** et des **inconvénients**.

I IMPACT SUR L'ÉCOLOGIE

Dans ce scénario intermédiaire, l'IA permet des **avancées technologiques** notables dans le domaine des **énergies renouvelables** et des **technologies bas-carbone**, incluant les techniques de capture et de stockage du carbone. Cela permet de renforcer la compétitivité de ces technologies par rapport aux énergies fossiles en augmentant leur efficacité et en réduisant leur coût. Ainsi, les forces du marché incitent les entreprises dans le monde entier à **accélérer leur décarbonation**, facilitant la mise en œuvre des CDN par les gouvernements car les impacts sur l'économie sont réduits, voir positifs dans certains cas. Cela amène des bénéfices tangibles en termes de diminution des GES, aidant la communauté internationale à se rapprocher des objectifs de l'**Accord de Paris** à l'horizon 2035 sur certains points et dans certains secteurs. Néanmoins, dans ce scénario mixte, le revers de la médaille est sans surprise l'augmentation significative des émissions de GES dans d'autres secteurs, en raison de la **consommation énergétique de plus en plus importante des nouveaux modèles d'IA**. En effet, même si les avancées apportées par l'IA pour le développement des technologies bas-carbone permettent d'atténuer quelque-peu ce phénomène, elles ne permettent pas de compenser complètement l'augmentation significative des émissions dans un nombre de secteurs clés. Ainsi, **l'impact global de l'IA en termes d'éco-**

logie est contrasté. En raison de la pression exercée par les lobbys, la balance a plutôt tendance à pencher du côté de l'augmentation des émissions, même si le résultat global demeure nuancé. Cela est d'autant plus le cas que beaucoup de gouvernements choisissent le nucléaire comme solution de facilité pour le développement de l'IA, entraînant une augmentation du nombre de déchets nucléaires avec des impacts négatifs sur l'environnement, sans compter les risques sécuritaires.

Pour ces raisons, l'IA pourrait avoir des conséquences **paradoxales** sur la mise en œuvre de l'Accord de Paris à l'horizon 2035. D'un côté, les avancées dans le domaine des technologies bas-carbone génèrent un réel enthousiasme, encourageant les gouvernements à **rehausser le niveau d'ambition de leurs CDN** soumises lors du processus de la COP. Sur le papier, l'agrégation de toutes les CDN permet à la communauté internationale, pour la première fois, de respecter le seuil de limitation du réchauffement climatique **en dessous des 2°C** d'ici la fin du siècle. Ce jalon, atteint lors de la COP40 en 2035, est célébré comme un point de bascule majeur dans la mise en œuvre de l'Accord de Paris, suscitant beaucoup d'espoir. Cependant, l'augmentation significative des émissions de GES dans un certain nombre de secteurs critiques, en raison des nouveaux modèles d'IA de plus en plus énergivores, fait que la mise en œuvre de ces nouvelles CDN plus ambitieuses reste une chimère et devient comparable à du « **greenwashing** ». De manière paradoxale, l'IA entraîne donc l'augmentation du niveau d'ambition des CND, mais sans que cela ne soit suivi par une mise en œuvre concrète, et l'écart entre les deux devient de plus en plus notable à l'horizon 2035. Sur ce point, **l'IA ne fait qu'exacerber des tendances qui sont déjà visibles en 2026**, où les nouvelles CND soumises par les gouvernements à la COP30 en novembre 2025 reflètent certes une augmentation du niveau d'ambition global sur le papier, mais sans être suivi pas des

²⁰³ Cela inclut les « laissés pour comptes de la mondialisation », auxquels viendront se rajouter les « laissés pour compte de la révolution de l'IA » à l'horizon 2035.

²⁰⁴ Ces inégalités sociales aux États-Unis comme ailleurs sont mesurable via l'[indice de Gini](#).

²⁰⁵ L'UE possède historiquement un modèle économique et social plus équilibré par rapport au système néolibéral aux États-Unis, ce qui permet de réduire les inégalités sociales. Voir : Hay et Wincott (2012), Hay et Payne (2015).

politiques concrètes au niveau national. Dans ce scénario intermédiaire, nous assistons donc à une **accélération des effets du changement climatique**, même si cela se déroule **modérément** par rapport au deuxième scénario pessimiste.

I IMPACT SUR LES PROCESSUS DÉMOCRATIQUES

Dans ce scénario mixte, l'IA a un **effet contrasté** sur les régimes démocratiques à l'horizon 2035. Sous l'angle négatif, l'IA contribue à **amplifier les campagnes de désinformation**, incluant l'augmentation des « *fake news* » et « *deepfakes* » sur Internet, ce qui accentue les tentatives de manipulation des processus électoraux. Dans le même temps, des **règlementations efficaces et rigoureuses** sont mises en place pour contrer et mieux maîtriser la désinformation sur Internet et les réseaux sociaux dans certains pays comme en France et en Europe, mais pas dans d'autres. Aux États-Unis par exemple, la culture du **laissez-faire économique** d'inspiration néolibérale, combinée à la pression exercée par les lobbys, empêchent la mise en œuvre de réglementations appropriées dans ce domaine. À l'instar des États-Unis, un certain nombre d'autres démocraties comme le Royaume-Uni et l'Australie, sous l'influence de la culture du laissez-faire économique et par souci de compétitivité, préfèrent ne pas mettre en place de réglementations trop rigoureuses qui pourraient entraver l'innovation. D'autres démocraties comme le Japon, le Canada ou la Corée du Sud ont tendance à **suivre le modèle européen**, tandis que celles dans les BRICS comme le Brésil ou l'Afrique du Sud, ainsi que le Mexique, tentent tant bien que mal de suivre ce modèle. Néanmoins, certains pays sont minés par la corruption et n'ont souvent **pas les ressources ou les infrastructures nécessaires** pour maîtriser les contenus sur Internet ou les réseaux sociaux. Dans les pays européens, au Japon, au Canada et en Corée du Sud qui parviennent à mettre en place une **législation suffisamment rigoureuse**, le jeu du chat et de la souris tourne à l'avantage des pouvoirs publics. Ceux-ci uti-

lisent les techniques d'IA avancées comme les filigranes numériques, ainsi que des réglementations strictes pour assurer l'**auditabilité** des systèmes et la **correction des biais** en continu, afin de lutter efficacement contre la désinformation en ligne. Cela permet de diminuer la manipulation des systèmes électoraux, faisant baisser le degré de polarisation et les fractures au sein de ces démocraties, tout en favorisant les partis politiques modérés. Au contraire, dans les pays qui n'ont pas mis en œuvre des lois suffisamment robustes dans ce domaine, soit par manque de moyens ou de volonté politique, la désinformation en ligne est amplifiée et le jeu du chat et de la souris tourne à l'avantage des hackers. Cela entraîne de graves conséquences en termes de **polarisation** avec la montée du populisme et des extrêmes politiques, **affaiblissant la démocratie de l'intérieur**.

Dans les États autoritaires, les systèmes de censure permettent un contrôle drastique des contenus en ligne et des réseaux sociaux. Les technologies d'IA avancées sont utilisées par les dictatures pour **accroître la surveillance sur leurs populations**, notamment par l'utilisation de la reconnaissance faciale et de la notation sociale, qui ont pour conséquence d'exacerber le niveau de **contrôle autoritaire**. Ainsi, il est clair que dans ce scénario intermédiaire, l'IA a des **effets contrastés** sur les processus démocratiques dans les pays à travers le monde, même si la balance penche plus du côté de l'**affaiblissement des libertés** que de leur renforcement. En effet, mis à part l'UE²⁰⁶ et certains de leurs partenaires démocratiques, peu de pays dans le monde seront en toute vraisemblance capables, avec la volonté politique nécessaire, de mettre en place des règles suffisamment rigoureuses pour lutter efficacement contre la désinformation en ligne. Par conséquent, dans une majorité des cas, l'IA risque au contraire de continuer à **affaiblir certaines démocraties de l'intérieur** en amplifiant la désinformation et la manipulation des processus électoraux. Pour les dictatures, la technologie leur prodigue même de nouveaux outils très performants pour accroître leur contrôle sur les populations. Une fois de plus,

²⁰⁶ Même en Europe, le processus actuel de simplification des législations sur le numérique (Digital Omnibus – voir ci-dessus) risque de compromettre la capacité de l'UE à lutter efficacement contre la désinformation.

selon ce scénario, l'IA ne fera que de **renforcer à l'horizon 2035 beaucoup de tendances déjà visibles aujourd'hui en 2026** que ce soit en Europe, aux États-Unis, en Chine et ailleurs, où les démocraties se retrouvent fragilisées, les dictatures renforcées, et les libertés en recul dans beaucoup de régions du monde.

I IMPACT SUR LA SOCIÉTÉ EN GÉNÉRAL

Dans ce scénario mixte, l'IA permet des **progrès importants en science et en médecine** à l'horizon 2035, contribuant à améliorer la durée et la qualité de vie des citoyens. Néanmoins, les nouveaux traitements disponibles grâce à l'IA coûtent cher et sont souvent inabordables, **renforçant ainsi les inégalités**.²⁰⁷ L'IA n'apporte **pas les miracles escomptés** comme la guérison totale du cancer comme promis par Sam Altman, même si elle permet le développement de nouveaux médicaments plus efficaces. De plus, les percées technologiques de l'IA permettent de trouver de nouvelles **solutions à certains problèmes sociétaux**, comme pour l'agriculture et l'accès à l'eau potable dans les pays en développement. Néanmoins, nous sommes loin des solutions miracles décrites dans le premier scénario optimiste, et bien des problèmes persistent. L'IA contribue également à **exacerber un certain nombre d'autres enjeux sociétaux**, sans sombrer néanmoins dans les catastrophes du second scénario pessimiste. Le **décrochage scolaire** s'accroît dans les pays développés, surtout au sein des milieux défavorisés, avec une utilisation répandue de l'IA comme substitut à l'enseignant en raison des faibles coûts associés, l'éducation avec des instituteurs ou professeurs humains devenant un luxe onéreux accessible uniquement à l'élite.²⁰⁸ Cela accentue les difficultés pour les jeunes à entrer sur le marché du travail, précisément au moment où l'IA transforme profondément les métiers à l'horizon 2035 avec une diminution du nombre de postes juniors, contribuant à une hausse du taux de **chômage chez les**

jeunes. De plus et comme expliqué précédemment, malgré les politiques pour accompagner les reconversions professionnelles dans les secteurs où l'IA crée de nouveaux emplois, une part non-négligeable de la population risque de ne **pas parvenir à s'adapter** et à trouver sa place dans ce nouveau marché du travail, où l'IA joue un rôle majeur. Ces différents facteurs contribuent à **l'augmentation de la précarité et des inégalités sociales** dans plusieurs couches de la population, avec une aggravation correspondante de la délinquance et de la criminalité. En exécutant des mesures telles que les taxes sur l'IA, certains gouvernements, notamment en Europe, parviennent difficilement à financer les pensions publiques pour ces « **laissés pour compte** » de la révolution de l'IA ; cela accentue les **dettes publiques** et fragilise davantage les finances des États. Dans d'autres nations plus pauvres qui ne possèdent pas les mêmes systèmes de protection sociales, ou à l'instar des États-Unis avec leur orientation néolibérale, les impacts sociétaux sont encore plus prononcés, affaiblissant les systèmes politiques et **renforçant les tensions et les degrés de polarisation**.

Enfin, la mise en place de cadres de gouvernance efficaces au niveau national dans plusieurs pays permet de garder un certain contrôle sur l'IA. En effet, l'UE est en position de force pour collaborer avec d'autres pays ayant mis en œuvre des mesures similaires, notamment le Canada ou le Japon. Cela entraîne la création de **cadres bilatéraux et multilatéraux** utilisant les **normes ISO**, afin de mieux réglementer l'IA et favoriser son usage responsable. Néanmoins, la gouvernance au niveau mondial via l'ONU continue d'être très difficile, en raison notamment de la **rivalité sino-américaine sur l'IA**, qui fait que les deux grandes puissances ne sont pas engagées de manière sincère dans ces processus multilatéraux par peur que cela ne compromette leur avantage compétitif. **Aucun cadre réglementaire contraignant au niveau mondial** sur l'IA n'est donc établi dans ce scénario, avec seulement des résolutions non-contraignantes votées

²⁰⁷ En outre, les bénéfices seraient concentrés au sein des pays développés qui ont pu investir suffisamment dans les infrastructures médicales renforcées par l'IA, et moins dans ceux en développement.

²⁰⁸ Les élèves n'ont bien entendu pas le même type de motivation pour apprendre avec une IA, en comparaison à la stimulation intellectuelle et émotionnelle prodiguée par l'interaction humaine.

par l'Assemblée générale de l'ONU, avec une portée surtout symbolique. Grâce à la **décélération** relative du développement technologique de l'IA (voir ci-dessus), l'humanité parvient tant bien que mal à **conserver un certain contrôle sur la technologie**, même si l'absence de cadres de gouvernance mondiaux rigoureux, et également nationaux dans de nombreux pays, fragilise cette maîtrise et entraîne l'apparition de **failles sécuritaires inquiétantes** pour l'avenir. Ce scénario intermédiaire à l'horizon 2035 illustre une fois encore un **renforcement de nombreuses tendances déjà visibles aujourd'hui en 2026** à propos de l'IA et qui risquent de s'accroître au cours de la prochaine décennie.

Conclusion

En conclusion, la révolution des technologies de l'IA va **profondément bouleverser la société à l'horizon 2035**. Cela inclut des impacts sur l'économie comme sur le travail, les métiers et les compétences, sur l'écologie en termes d'empreinte environnementale, sur les processus démocratiques en raison de la désinformation exacerbée par l'IA, ainsi que sur la société plus généralement notamment la recherche scientifique et médicale, l'éducation et les inégalités sociales. En raison de ces transformations profondes, cette étude met en avant l'importance fondamentale de **conserver l'humain dans la boucle par la réglementation, afin de définir des limites pour garder le contrôle** sur la technologie. L'étude a présenté un état des lieux avec des scénarios prospectifs à court terme (horizon 2026-28) sur la cohabitation humain-IA, suivi par une analyse des risques systémiques qui en découlent et les mécanismes de régulation applicables, en proposant des recommandations opérationnelles pour les entreprises et les décideurs publics. La dernière partie a décliné ces éléments en trois scénarios prospectifs contrastés (optimiste, pessimiste et intermédiaire/mixte) à moyen terme à l'horizon 2035. L'étude a proposé des solutions concrètes pour reprendre le contrôle sur l'IA, réparties en sept axes d'action principaux, incluant la mobilisation de **€100 milliards supplémentaires d'ici 2030** pour la protection socio-économique et **€300 milliards supplémentaires d'ici 2035** pour la souveraineté numérique :

1. **Protéger les secteurs économiques impactés par l'IA** : Obligation pour les employeurs de maintenir l'employabilité de leurs salariés, création de comités d'éthique et de supervision numérique, clauses de non-substitution dans les conventions col-

lectives, dividende numérique et fonds de reconversion sectorielle, réduction du temps de travail, et, à moyen terme, un « Revenu Universel de Base ». Le tout financé par les gains de productivité de l'IA, notamment une « taxe robot » et des fonds souverains technologiques.

2. **Soutenir la formation à l'IA et l'augmentation humaine** : Crédit d'impôt « augmentation humaine », financement de « congés de conversion IA » rémunérés, passeport européen de compétences IA, micro-accréditations sectorielles obligatoires et droit à la transparence algorithmique au travail. Création d'un « Index de l'augmentation humaine » conditionnant l'accès aux marchés publics, avec l'intégration de l'IA dans une nouvelle « 5e liberté » du marché unique européen dédiée à l'éducation, à la recherche et à l'innovation.

3. **Renforcer la souveraineté numérique européenne** : Fonds souverain numérique (€30 Mds/an) pour renforcer les infrastructures de l'UE, préférence européenne dans la commande publique, relocalisation de la production de puces (*Chips Act* accéléré), mutualisation des données stratégiques (*Data Union*), soutien à l'Open Source (€4 milliards/an) et union des marchés de capitaux.

4. **Contenir la désinformation exacerbée par l'IA** : Passeport algorithmique obligatoire pour les modèles à haut risque, système d'alerte inter-plateformes en temps réel, « Bacs à sable de données » sécurisées gérées par le Bureau européen de l'IA, et filigrane numérique renforcé.

5. **Diminuer les impacts environnementaux de l'IA** : Droit à l'audit environnemental, taxe carbone numérique proportionnelle, quotas d'eau par unité de calcul (Téraflop), standardisation des modèles SLM moins énergivores, et création d'un Index de réparabilité et durabilité des infrastructures d'IA.
6. **Renforcer la réglementation sur l'IA** : Mise à jour annuelle sans procédure législative des catégories de risque de l'*AI Act*, élargissement des « risques inacceptables » (biométrie, reconnaissance faciale), intégration des impacts sociétaux (manipulation démocratique), renforcement des évaluations de conformité tierce partie, ainsi qu'une sécurisation renforcée (physique et cyber) des *data centers* sur le modèle des centrales nucléaires.
7. **Gouvernance mondiale pragmatique de l'IA avec mécanismes multiniveaux** : Faute de traité mondial contraignant à court terme, l'UE doit s'appuyer sur les normes ISO en les intégrant systématiquement, ainsi que sur sa position de plus grand marché unique pour développer des cadres bilatéraux et multilatéraux afin de diffuser ses standards éthiques sur l'IA à l'international.

En outre, l'objectif de cette Conclusion est **d'ouvrir la réflexion** plus largement autour de l'impact de l'IA sur la société à l'**horizon 2050 ou 2100**, en abordant la question de comment préparer la société à l'arrivée potentielle de l'**IA générale** (*Artificial General Intelligence* - AGI) et à la « singularité technologique » qui en découle. En effet, dans tous les scénarios examinés, même les plus extrêmes, cette étude s'est basée sur l'hypothèse que **l'AGI et la singularité technologique ne seront pas atteintes d'ici 2035**. Même si certains prédisent le contraire, le consensus scientifique aujourd'hui est soit que ces technologies ultimes sont impossibles à développer même avec des percées techniques majeures, soit que l'AGI et la singularité ne seront pas atteintes avant encore au moins

plusieurs décennies.²⁰⁹ Même s'il n'existe pas d'accord sur les définitions précises de ces concepts, l'AGI est souvent décrite comme possédant une **intelligence similaire à celle de l'Homme** en étant capable d'accomplir toutes les tâches intellectuelles que l'être humain possède, incluant l'apprentissage, le raisonnement et l'adaptation à de nouvelles situations. Pour ces raisons, l'AGI serait aussi en théorie dotée d'une **forme de conscience**. Celle-ci est parfois présentée comme le précurseur à la « **super-intelligence** » (*Artificial Super Intelligence* - ASI), qui surpasse l'intelligence humaine dans tous les domaines et pourrait potentiellement accomplir des choses dépassant l'entendement humain.²¹⁰ Ainsi, le concept de la « singularité technologique » découle de l'arrivée de l'AGI et ensuite de l'ASI, marquant un **point de rupture historique** où la croissance technologique devient incontrôlable car **l'IA s'améliore d'elle-même de façon autonome**, entraînant ainsi des changements imprévisibles et irréversibles pour la civilisation humaine, conduisant à un **basculement civilisationnel**.

Avec une IA capable de surpasser l'humain dans toutes les tâches cognitives et physiques (via la robotique avancée), le concept de « travail », tel qu'il existe aujourd'hui comme pilier de l'identité sociale, pourrait disparaître. Cela imposerait soit une **redistribution massive** (via un revenu universel de base par exemple), soit le risque de **fractures sociales profondes** et de révolutions ; en effet, si la productivité est assurée par les machines, la redistribution des richesses ne pourra plus passer par le salaire. Nous devons peut-être passer d'une société centrée sur l'humain et sa productivité (l'homme-outil), à une **société centrée sur le sens, l'art et les relations interpersonnelles**, ainsi qu'à une **coexistence complexe entre l'humain et la machine**. Dans un scénario optimiste, l'AGI ou l'ASI permettraient de résoudre des défis existentiels tels que le changement climatique grâce à une optimisation radicale des ressources incluant la fusion nucléaire maîtrisée par l'IA, ainsi que l'éradication de la plupart des maladies. L'IA

²⁰⁹ Sheng (2025), Fjelland (2020).

²¹⁰ <https://www.ibm.com/think/topics/artificial-general-intelligence>

pourrait déchiffrer les mécanismes du vieillissement et créer des thérapies géniques personnalisées en temps réel, provoquant un bondissement de l'espérance de vie, posant des questions de surpopulation et d'accès équitable à cet allongement de la durée de vie. Le plus grand défi serait de s'assurer que les objectifs d'une super-intelligence restent **alignés avec les valeurs humaines**. En effet, une IA très intelligente n'a pas besoin d'être « méchante » pour être dangereuse ; il suffit qu'elle soit indifférente aux besoins de l'humanité pour provoquer des catastrophes. Dans un scénario plus optimiste, nous nous dirigeons peut-être vers la **fusion entre l'Homme et la machine d'ici la fin du siècle**, avec une interface cerveau-machine avancée (de type *Neuralink*) pour permettre aux humains de ne pas être dépassés et devenir obsolètes. Cela impliquerait une communication directe avec l'IA, redéfinissant ce que signifie être « humain », amenant ainsi une **marche forcée vers la prochaine grande étape de l'évolution**.

Ces scénarios à l'horizon 2050 voire 2100 relèvent bien entendu de la science-fiction, et rien ne garantit qu'ils deviendront réalité. Cependant, le fait qu'ils soit même concevables souligne une nouvelle fois l'importance fondamentale de garder l'humain dans la boucle de la révolution technologique de l'IA, en **régulant et en définissant des limites pour garder le contrôle**, comme explicité dans cette étude. L'objectif est bien entendu de faire en sorte que la technologie puisse servir à **augmenter le bien-être de l'humanité**, plutôt que d'être le précurseur de sa fin ou le déclencheur de cataclysmes sans précédent.²¹¹ Comme l'a affirmé Stephen Hawking : « Réussir à créer l'IA serait le plus grand événement de l'histoire de l'humanité. Malheureusement, ce pourrait aussi être le dernier, à moins que nous n'apprenions comment éviter les risques ».²¹²

²¹¹ En effet, comme indiqué dans la 1^{re} Annexe, l'une des convergences les plus frappantes du sondage reste celle autour de la peur d'une perte de contrôle de l'IA à l'horizon 2035. Quelle que soit la population interrogée, la nécessité de maintenir un contrôle humain effectif sur l'IA fait l'objet d'un large consensus. Le message central du sondage est clair : l'avenir de l'IA se jouera autant sur le plan technologique que sur celui de la gouvernance et de la capacité des sociétés à définir et imposer des limites au nom de l'intérêt général.

²¹² Hawking *et al.* (2014).

Bibliographie

- Barichella A. (2026). *Threats, Prospects and Policy Approaches for the Twin Digital and Energy Transitions*, dans Barichella A. (Éditeur-en-Chef) et Yada J. (2026). *The Palgrave Handbook of Cybersecurity, Technologies and Energy Transitions*. Palgrave Macmillan.
- Barichella A. (2025). "Addressing fragmentation to create a common EU framework for a true Digital Single Market: a case study of the cybersecurity sector", in *Why Europe needs a true Digital Single Market*. Vodafone Vision Paper.
- Barichella A. (2023a). *Regulating artificial intelligence at the EU level: obstacles and prospects*. Policy Paper n°294 – Jacques Delors Institute.
- Barichella A. (2023b). "Cybersecurity and Data Protection in the Power Sector: Challenges, Perspectives and Policy Approaches", in Considine J., Cote S., Cook D. and Wood G. (Eds.), *A Research Agenda for Energy Politics*. Edward Elgar Publishing – Research Agenda series.
- Barichella A. (2022). *Cyberattacks in Russia's hybrid war against Ukraine and its ramifications for Europe*. Policy Paper n°281 – Jacques Delors Institute.
- Badshah N. (2026), *Teenager died after asking ChatGPT for 'most successful' way to take his life, inquest told*. The Guardian.
- Bellan R. (2025), *Sam Altman says ChatGPT has hit 800M weekly active users*, TechCrunch.
- Blind K. et al. (2021). *The impact of Open Source Software and Hardware on technological independence, competitiveness and innovation in the EU economy*. European Commission – Final Study Report.
- Blum J., *Nuclear fusion, the 'holy grail' of power, was always 30 years away—now it's a matter of when, not if, fusion comes online to power AI*. Fortune, 2 octobre 2025.
- Bick A. et al. (2026), *Why Does AI Adoption Differ So Much across Countries?*, St. Louis Federal Reserve Bank On the Economy.
- Challapally A. et al. (2025), *The GenAI Divide: State of AI in Business 2025*, MIT NANDA.
- Dastin J. et Nellis S. (2023). *Focus: For tech giants, AI like Bing and Bard poses billion-dollar search problem*. Reuters, 22 février 2023.
- Desarnaud G. (2017). *Cyber Attacks and Energy Infrastructures: Anticipating Risks*. Études de l'Ifri, Ifri.
- Draghi M (2024). *The Draghi Report on The future of European Competitiveness. Part A: A competitiveness strategy for Europe*.
- Draghi M (2024). *The Draghi Report on The future of European Competitiveness. Part B: In-depth analysis and recommendations*.
- Emerson G. et al. (2026), *AI Will Reshape More Jobs Than It Replaces*, Boston Consulting Group.
- Fjelland R. (2020). "Why general artificial intelligence will not be realized". *Humanities and Social Sciences Communications*, 7, 10.
- Frank N. (2026), *Thinking small: How small language models could lessen the AI energy burden*, Virginia Tech.
- George A. L. et Bennett A. (2005), *Case Studies and Theory Development in the Social Sciences*. The MIT Press.
- Hatzius J. et al. (2023), *The Potentially Large Effects of Artificial Intelligence on Economic Growth*, Goldman Sachs - Global Economics Analyst.
- Hawking S. et al., *Stephen Hawking: 'Transcendence looks at the implications of artificial intelligence - but are we taking AI seriously enough?*, The Independent, 1 May 2014.
- Hay C. et Payne A. (2015), *Civic Capitalism*, Polity Press.
- Hay C. et Wincott D. (2012). *The Political Eco-*

- onomy of European Welfare Capitalism (21st Century Europe). Palgrave Macmillan.
- Hu K. (2023), *ChatGPT sets record for fastest-growing user base*, Reuters.
 - International Energy Agency (2024), *Energy and AI*. World Energy Outlook Special Report.
 - International Energy Agency (2025), *Energy and AI*. World Energy Outlook Special Report.
 - International Labour Organization (2026), *Employment and Social Trends – World of Work Series*.
 - Jahangir R., *What's Behind Europe's Push to "Simplify" Tech Regulation?*, Tech Policy Press, 24 avril 2025.
 - Kenny R. (2025), *Report on Water Use in AI and Data Centers*, UK Government Publishing Service.
 - Lazaro Cabrera L. *Europe's Deregulatory Turn Puts the AI Act at Risk*. Tech Policy Press, 3 juin 2025.
 - Letta E. (2025). *Much More Than A Market: Speed, Security, Solidarity. Empowering the Single Market to deliver sustainable future and prosperity for all EU citizens*.
 - Liu Y. et Wang H. (2024). *Who on Earth Is Using Generative AI? Global Trends and Shifts in 2025*. Policy Research Working Paper, World Bank.
 - May B. et Badawy Y., *Evidence of an AI-driven shakeup of job markets is patchy*, Global Research Briefing - Oxford Economics.
 - Murray C. (2025), *Why AI 'Hallucinations' Are Worse Than Ever*, Forbes.
 - Nathan A., Grimberg J. et Rhodes A. (2024), *Gen AI: Too Much Spend, Too Little Benefit?*, Global Macro Research: Issue 129.
 - Porębski A. et Figura J. (2025). "There is no such thing as conscious artificial intelligence". *Humanities and Social Sciences Communications*, 12, 1647.
 - Ren S. (2023), *How much water does AI consume? The public deserves to know*, OECD AI Observatory.
 - Réseau de transport d'électricité (RTE – 2024), *Bilan électrique 2024 : Principaux Résultats*.
 - Ryseff J., De Bruhl B. et Newberry S. J. (2024), *The Root Causes of Failure for Artificial Intelligence Projects and How They Can Succeed*, Research Report – Rand Corporation.
 - Schenker J., *How AI will help get fusion from lab to grid by the 2030s*. World Economic Forum, 12 décembre 2025.
 - Secrétariat général de la défense et de la sécurité nationale – République française, *Stratégie nationale de cybersécurité française 2026-30*.
 - Sheng H. (2025). "Why Can't Artificial General Intelligence Be Achieved?" *Journal of Historical Archeology & Anthropological Sciences*, 10(2):43~47.
 - Sajadieh S. et al. (2026), *Artificial Intelligence Index Annual Report 2026*. Institute for Human-Centered AI (HAI), Stanford University.
 - Shumailov I. et al. (2024). "AI models collapse when trained on recursively generated data". *Nature* 631, 755–759.
 - Wimmers. A. J. et al. (2025). "Power generation from nuclear fusion not expected in the foreseeable future: Applied research developing dynamically", *Deutsches Institut für Wirtschaftsforschung (DIW)*, Vol. 15, Issue 14/15, 90-95.
 - Wood D. et al. (2025), *Sovereign AI: Own your AI future, From managing risk to accelerating growth*, Accenture.
 - Worland J. *The IMF's Kristalina Georgieva on the AI 'Tsunami' Hitting Jobs*. Time, 15 janvier 2026.
 - World Economic Forum (2025), *Future of Jobs Report – Insight Report*.

I SITES WEB:

- Aissa A. B., *Bug OpenClaw : quand l'agent IA efface les mails d'une directrice de Meta*, Numerama, 24 février 2026 : <https://www.numerama.com/cyberguerre/2186121-bug-openclaw-quand-lagent-ia-efface-les-mails-dune-directrice-de-meta.html>
- BBC news, *ChatGPT boss says government regulation of AI is 'crucial'*: <https://www.bbc.com/news/live/world-us-canada-65610337>
- BusinessStats, *Artificial Intelligence (AI) Worldwide – Statistics & Facts 2026*: <https://businessstats.com/artificial-intelligence-ai-worldwide-statistics/?hl=fr-FR#>
- Eadicicco L. et Duffy C., *Meta to cut 10% of staff as it pours billions into AI*, CNN Business: <https://edition.cnn.com/2026/04/23/tech/meta-layoffs-10-percent-staff-ai>

- Eurostat, *Le taux de chômage dans la zone euro et l'UE*, octobre 2025: <https://ec.europa.eu/eurostat/fr/web/products-euro-indicators/w/3-02122025-bp#estat-in-page-nav-heading-0>
 - Eurostat, *20% of EU enterprises use AI technologies*, décembre 2025: <https://ec.europa.eu/eurostat/web/products-eurostat-news/w/ddn-20251211-2>
 - FranceNum, *Le portail de la transformation numérique des entreprises* : <https://www.francenum.gouv.fr/formations/quest-ce-que-le-cloud-computing-ou-informatique-en-nuage-7-minutes-pour-comprendre#:~:text=Le%20recours%20au%20cloud%20computing,local%20ou%20un%20ordinateur%20personnel>
 - Groupe Crédit Agricole, *Contexte politico-économique – États-Unis* : <https://international.groupecreditagricole.com/fr/accompagnement-a-l-international/etats-unis/contexte-economique#:~:text=L'économie%20américaine%20est%20essentielle,les%20plus%20liquides%20du%20monde>.
 - Hays K., *Thousands lose their jobs in deep cuts at tech giant Oracle*. BBC news, 1er Avril 2026: <https://www.bbc.com/news/articles/cm296jzzl9yo>
 - IBM, *What is artificial general intelligence (AGI)?*: <https://www.ibm.com/think/topics/artificial-general-intelligence>
 - Industrie Numérique, *Le secteur tertiaire en Europe : dynamique et enjeux actuels*: <https://www.industrie-numerique.com/le-secteur-tertiaire-en-europe-dynamique-et-enjeux-actuels/>
 - INSEE Statistiques, *Combien y a-t-il de chômeurs en France ?* : <https://www.insee.fr/fr/statistiques/4805248#graphique-chomage-g1-fr>
 - INSEE Statistiques, *Emploi, chômage, revenus du travail* (Édition 2025) : <https://www.insee.fr/fr/statistiques/8376826?sommaire=8376908>
 - Islam F. and Clun R., *Google boss says trillion-dollar AI investment boom has 'elements of irrationality'*. BBC News World, 18 Novembre 2025: <https://www.bbc.com/news/articles/cwy7vrd8k4eo>
 - Kuznia R., Gordon A. et Lavandera E., *'You're not rushing. You're just ready:' Parents say ChatGPT encouraged son to kill himself*, CNN investigates: <https://edition.cnn.com/2025/11/06/us/openai-chatgpt-suicide-lawsuit-invs-vis>
 - Le Monde, *Open AI, Microsoft face lawsuit over ChatGPT's alleged role in a murder-suicide*, 12 December 2025: https://www.lemonde.fr/en/pixels/article/2025/12/12/open-ai-microsoft-face-lawsuit-over-chatgpt-s-alleged-role-in-a-murder-suicide_6748404_13.html
 - Mount Sinai, *New Study Links Altered Cellular States to Brain Structure*, Press Release, 22 janvier 2026: <https://www.mountsinai.org/about/newsroom/2026/new-study-links-altered-cellular-states-to-brain-structure>
 - Trading Economics, *China Unemployment Rate*: <https://tradingeconomics.com/china/unemployment-rate>
 - US Bureau of Labor Statistics, *Civilian unemployment rate*: <https://www.bls.gov/charts/employment-situation/civilian-unemployment-rate.htm>
 - World Population Review, *Gini Coefficient by Country 2026*: <https://worldpopulationreview.com/country-rankings/gini-coefficient-by-country>
- LISTE DES PERSONNES INTERVIEWÉES DANS LE CADRE DE CETTE ÉTUDE :**
1. **Claudia RODA (France)** : Professeur en Informatique à l'Université américaine de Paris, détentrice de la Chaire UNESCO sur l'intelligence artificielle et droits humains, le 16 février 2026.
 2. **Ivan KULCHYTSKY (Ukraine)** : Président de l'ONG *Agency of European Innovations* (Ukraine) le 20 février 2026.
 3. **Nazar PODOLCHAK (Ukraine)** : Responsable du parc scientifique de l'Université nationale polytechnique de Lviv (Ukraine), le 25 février 2026.
 4. **Régis CHATELLIER (France)** : Chargé des études et des explorations prospectives pour la Commission nationale de l'informatique et des libertés (CNIL - France), le 2 mars 2026.

5. **Johnny DESCHAMPS (France)** : Professeur à l'Unité Chimie et Procédés (UCP), Institut Polytechnique de Paris (IPP), le 4 mars 2026.
6. **Pagona TSORMPATZOU DI (Royaume-Uni)** : Directrice, vie privée/protection des données & gouvernance de l'IA à Mastercard, le 11 mars 2026.
7. **Susan NESBITT (États-Unis)** : Directrice Générale - The Cantellus Group, *Human Data Red Team Research* - OpenAI, 3 et 10 avril 2026.
8. **Antonio KUNG (France)** : Co-fondateur et membre du conseil d'administration - Trialog (France), le 7 avril 2026.
9. **Hynd REMITA (France)** : Directrice de Recherche à l'Institut de Chimie Physique, Université Paris-Saclay/CNRS (France), le 8 avril 2026.
10. **Nathan HEATH (États-Unis)** : Officier de sécurité et sûreté de l'IA pour Anthropic et OpenAI : 23 avril 2026.
11. **Alexandru GEORGESCU (Roumanie)** : Chercheur en cybersécurité et protection des infrastructures critiques au *National Institute for Research & Development in Informatics* (ICI) à Bucharest: 24 avril 2026.
12. **Alvaro ARTIGAS (France)** : Chercheur associé au Centre d'études européennes et de politique comparée (CEE), Sciences Po Paris : 8 mai 2026.

Annexes

1 . ANNEXE 1: Analyse comparative des résultats issus du sondage

I 1. INTRODUCTION ET MÉTHODOLOGIE

Un sondage a été mené auprès d'un large panel de participants issus de divers secteurs en France, en Europe et à l'international. Ce sondage se base sur un questionnaire d'une trentaine de questions avec plusieurs types de réponses possibles, notamment des degrés d'accord allant de 0 à 5²¹³, permettant le traitement de données quantitatives à grande échelle. Le sondage a été mené sur trois mois (février à mai 2026), apportant une large gamme de données avec 460 réponses au total. Les questions couvrent l'ensemble des sujets et thématiques abordés dans cette étude, et ont été divisées en quatre parties :

1. Questions préliminaires sur le profil des personnes interrogées.
2. Questions sur les impacts de l'IA sur le travail, les métiers et les compétences.
3. Questions sur le rôle de la gouvernance politique et la réglementation de l'IA.
4. Questions sur les impacts plus généraux de l'IA sur la société.

L'axe central du sondage est la question du maintien de l'humain dans la boucle décisionnelle, entendu comme la capacité à réguler, contrôler et définir des limites sur l'IA tout en préservant la primauté humaine. Pour des raisons linguistiques et opérationnelles, trois versions parallèles du même questionnaire ont été administrées à des publics distincts²¹⁴ :

Version du sondage	Réseau associé	Nombre de répondants	Concentration sectorielle
Sondage Institut Jacques Delors – réseau francophone	Professionnels francophones au sein du réseau de l'Institut Jacques Delors	327 répondants	Très large spectre sectoriel (entreprises, institutions internationales, administrations publiques, université, think tank, ONG, etc.).
Sondage Institut Jacques Delors – réseau anglophone	Professionnels anglophones au sein du réseau de l'Institut Jacques Delors, plus large diffusion internationale	96 répondants	Similaire à la version francophone mais avec un fort poids université/recherche, ONG et organisations internationales

²¹³ Avec 0 marquant un désaccord total et 5 indiquant un accord parfait.

²¹⁴ Cette Annexe n'a pas pour but de présenter l'ensemble des résultats du sondage, mais seulement d'en faire une analyse globale.

Version du sondage	Réseau associé	Nombre de répondants	Concentration sectorielle
Sondage réseau CRiP	Membres du Club des Responsables d'infrastructure, de technologies et de production informatique (CRiP)	37 répondants	Chefs d'entreprises, directeurs des systèmes d'information (DSI) ou directeurs techniques (CTO) dans les secteurs IT, banque, industrie, énergie, etc.

La différence de taille des échantillons est notable : le sondage IJD francophone est de loin celui qui a pu recueillir le plus de réponses, le plus représentatif avec 327 répondants, contre 96 pour la version anglophone et 37 pour la version du CRiP. Malgré ces disparités, les tendances observées sont remarquablement convergentes sur la plupart des grandes thématiques. Compte tenu de la nature graphique des résultats (principalement des histogrammes), l'analyse est conduite de manière comparative sur les tendances dominantes, les convergences et les divergences entre les trois groupes par rapport aux différentes questions.

I 2. L'IA ET LE TRAVAIL : TRANSFORMATION PLUTÔT QUE SUPPRESSION

— L'IA comme outil indispensable (Q4) :

Les trois panels convergent sur un point : l'IA est déjà perçue comme un outil important sur le marché du travail aujourd'hui, avec des scores concentrés entre 3 et 5 sur l'échelle de 0 à 5. Le panel CRiP, composé de professionnels de l'IT et des infrastructures numériques, affiche la conviction la plus forte – les scores 4 et 5 y semblent plus prononcés – ce qui est cohérent avec leur proximité quotidienne avec les outils technologiques.

Convergence

Les trois sondages s'accordent largement sur une vision nuancée de l'impact de l'IA sur le marché du travail. Ils s'accordent à reconnaître l'IA comme un outil déjà significatif, même si la perception de son caractère « indispensable » varie selon le secteur d'appartenance.

— Création d'emplois vs. suppression (Q5) :

La majorité des répondants rejette le scénario catastrophiste d'une destruction massive d'emplois à court terme ; les options « pas vraiment d'accord » ou « cela dépend » dominent dans les trois versions. La question de la compensation par de nouveaux emplois est celle qui suscite le plus de scepticisme dans les trois panels. Les réponses « pas vraiment d'accord » et « pas du tout d'accord » dominent, avec une nuance significative : la réponse « cela dépend » recueille également une part importante des votes, signe d'une prudence analytique plutôt que d'un rejet catégorique. Le panel anglophone IJD, plus académique, semble légèrement plus nuancé dans ses réponses.

— Transformation des métiers et formation (Q6) :

C'est sur cette question que les trois panels expriment la convergence la plus forte et la plus positive. Les scores se concentrent massivement vers 4 et 5, indiquant un accord très large sur l'idée que le défi principal n'est pas la suppression massive d'emplois mais bien la transformation de leur nature, et que la réponse adéquate réside dans la formation aux compétences IA. Cette conviction est particulièrement marquée dans le panel CRiP, où les directions IT sont en première ligne de cette transformation.

Point clé

La formation aux compétences IA émerge comme la réponse prioritaire dans les trois panels – c'est le consensus le plus net de l'ensemble du sondage, et il est directement lié à la question du maintien de l'humain dans la boucle.

— Les utilisateurs d'IA remplaceront-ils ceux qui ne s'en servent pas ? (Q7) :

L'une des questions les plus consensuelles de l'ensemble du sondage, la thèse selon laquelle les professionnels maîtrisant l'IA remplaceront progressivement ceux qui ne l'utilisent pas recueille un accord très élevé dans les trois panels (scores concentrés entre 3 et 5). Les membres du CRiP, confrontés quotidiennement à ces enjeux dans leurs organisations, expriment une conviction encore plus marquée, suivi par le panel IJD francophone, tandis que celui anglophone semble légèrement plus polarisé. Ce résultat traduit un sentiment d'urgence partagé quant à la nécessité de former les travailleurs et la nécessité d'une politique active de montée en compétences.

— Fin du travail à l'horizon 2035 ? (Q8) :

La vision d'une « fin du travail » à horizon 2035 est reçue avec beaucoup plus de prudence. Les distributions sont plus étalées, avec une concentration autour des valeurs médianes (2-3), suggérant ni rejet ni adhésion. Le panel CRiP, plus technique, se montre légèrement moins pessimiste sur ce scénario extrême. Cela reflète une incertitude profonde sur les effets à moyen terme de l'IA, motivant d'autant plus les questions de gouvernance et de contrôle humain.

I 3. GOUVERNANCE DE L'IA : L'HUMAIN DANS LA BOUCLE

— Comités décisionnels hybrides aujourd'hui (Q9) :

La question de la gestion partielle d'une entité par l'IA via un comité hybride – avec maintien d'une supervision humaine – suscite une réponse modérément positive. Les scores se situent principalement entre 2 et 4 dans les trois panels. Celui du CRiP tend à être légèrement plus ouvert à ce scénario, peut-être en raison de leur expérience des systèmes automatisés de décision en IT, suggérant que les praticiens du numérique perçoivent déjà cette évolution comme réaliste et en cours. Au contraire, le panel IJD francophone manifeste un accord

plus modéré (scores centrés sur 2/3), tandis que le panel anglophone est plus favorable, possiblement lié à une culture de la gouvernance technologique plus développée.

Nuance importante

La majorité des répondants accepte l'idée d'une IA dans la boucle décisionnelle à condition que le contrôle et la vérification humaine soient maintenus – c'est précisément la définition de l'« humain dans la boucle ». Une tension structurelle émerge entre la reconnaissance des potentialités de l'IA et la crainte de perdre la maîtrise humaine sur ses décisions.

— IA dirigeantes à l'horizon 2035 (Q10) :

Le scénario de l'IA comme décideur autonome et complet à horizon 2035 est clairement rejeté par les trois panels, avec des scores concentrés en bas de l'échelle (0-2). Ce résultat est l'un des plus convergents et des plus tranchés de l'enquête. La résistance est légèrement plus forte dans le panel francophone IJD et dans celui du CRiP – deux groupes proches des instances réglementaires et des réalités opérationnelles.

— Changement de donne avec l'arrivée de l'IA « agentique » (Q11) :

L'arrivée des IA agentiques – capables de prendre des décisions autonomes et de s'optimiser en temps réel – est perçue comme un changement de paradigme significatif par une majorité dans les trois panels (« tout à fait d'accord » et « plutôt d'accord » dominant). Cela souligne que les répondants anticipent que ce type d'IA représente effectivement une rupture qualitative par rapport aux systèmes actuels. Le panel CRiP, confronté aux implications d'infrastructure de tels systèmes, est particulièrement sensible à cet enjeu. Cette reconnaissance amplifie d'autant la nécessité perçue d'un cadre de contrôle robuste ; c'est précisément sur ce point que la question du contrôle humain prend toute son acuité.

– 3.4 Biais, auditabilité et transparence (Q17) :

Les trois panels s'accordent très largement pour reconnaître que les biais algorithmiques, l'auditabilité et la transparence des systèmes constituent des obstacles majeurs au développement d'une IA responsable. Le panel anglophone IJD – plus académique – et le panel CRiP – plus technique – sont particulièrement sensibles à ces enjeux. Ce constat fonde directement la nécessité de mécanismes de supervision humains robustes.

I 4. RÉGLEMENTATION ET SOUVERAINETÉ NUMÉRIQUE

– Les législations française et européenne sont-elles suffisantes ? (Q18) :

La question de l'adéquation entre la réglementation française et celle européenne (notamment l'AI Act) face aux enjeux de biais et d'auditabilité recueille des réponses majoritairement sceptiques dans les trois panels. Les scores se concentrent entre 1 et 3, indiquant un sentiment que les législations actuelles, bien que nécessaires, restent insuffisantes. Le panel anglophone, plus familier des standards internationaux, est peut-être légèrement plus critique.

– Protection des travailleurs et reconversion (Q19) :

La perception est encore plus négative sur la capacité de la France et de l'UE à protéger les travailleurs des secteurs vulnérables. Les trois panels affichent des distributions fortement orientées vers les scores bas (0-2), signifiant un déficit perçu de politique publique active en matière de reconversion professionnelle liée à l'IA.

– Excès de réglementation ? (Q21) :

Sur les questions de régulation, les trois sondages révèlent une position nuancée. Contrairement à ce que l'on pourrait attendre d'un panel de professionnels, la thèse d'un excès réglementaire freinant l'innovation recueille des

avis partagés et modérés, avec des scores centrés autour de 2-3. Le panel CRiP est légèrement plus sensible à cet argument, ce qui reflète les préoccupations des DSI face aux contraintes de conformité. Les répondants IJD (francophones et anglophones), plus proches des institutions, sont plus sceptiques face à cette critique. Cette tension entre régulation nécessaire et risque de blocage de l'innovation est un fil rouge de l'ensemble du débat sur le contrôle humain de l'IA.

– Dépendance technologique de l'UE (Q22) :

La dépendance vis-à-vis des technologies IA américaines et chinoises est jugée problématique avec des scores élevés dans les trois groupes (3-5). Le panel CRiP et celui de l'IJD francophone sont particulièrement préoccupés par l'absence de champions européens. Cette convergence est remarquable : elle transcende les différences sectorielles et nationales pour établir un consensus sur la nécessité d'une souveraineté technologique européenne comme condition du contrôle humain collectif.

Consensus fort

Les trois panels s'accordent sur la nécessité de champions européens du numérique et jugent insuffisantes les politiques actuelles de protection des travailleurs face à l'IA – deux conditions indissociables d'une gouvernance humaine effective.

– Interdiction des IA à haut risque (Q23) :

En accord avec l'AI Act européen, les trois panels approuvent largement l'interdiction des systèmes IA qui manipulent les décisions ou classifient les comportements sociaux. Les scores sont concentrés vers 4 et 5, avec une convergence particulièrement forte dans le panel CRiP, confronté aux enjeux de sécurité des systèmes d'information. Cette position illustre un consensus sur la nécessité de lignes rouges pour protéger les droits fondamentaux, maintenir la dignité et l'autonomie humaines face à l'IA.

– **Data centers régulés comme des centrales nucléaires (Q24) :**

Cette proposition plus radicale divise davantage les panels, avec des distributions plus étalées (1-4). Elle suscite un accord légèrement plus prononcé chez les répondants IJD que chez les membres du CRiP, peut-être parce que ces derniers perçoivent mieux les contraintes opérationnelles d'une telle réglementation. La métaphore nucléaire – stricte régulation avec protocoles de sécurité et seuils de dangerosité – fait son chemin, mais sans emporter une adhésion massive.

I 5. RISQUES SOCIÉTAUX ET IMPACTS ENVIRONNEMENTAUX

– **Impacts environnementaux de l'IA (Q20) :**

L'inclusion systématique de l'impact environnemental de l'IA dans les politiques de transition énergétique aux niveaux national et européen fait l'objet d'un accord très large. Les réponses « tout à fait d'accord » et « plutôt d'accord » dominent largement, en particulier dans les panels IJD. Celui du CRiP, plus tourné vers les opérations IT, est légèrement plus nuancé mais reste majoritairement favorable. Ce consensus est cohérent avec la prise de conscience croissante de la consommation massive en énergie et en eau des *data centers* ; ce résultat souligne que la durabilité de l'IA est désormais une préoccupation transversale, et non plus marginale.

– **Désinformation et processus démocratiques (Q27) :**

Le risque que l'IA exacerbe les campagnes de désinformation et nuise aux processus démocratiques est perçu comme élevé dans les trois groupes (scores entre 3 et 5). Le panel IJD francophone et le panel anglophone, plus proches des institutions politiques et académiques, sont les plus préoccupés. Ce risque est directement lié à la question du contrôle humain : une IA non régulée dans l'espace informationnel constitue une menace pour la délibération démocratique elle-même.

– **Avancées scientifiques et santé (Q28) :**

À rebours des préoccupations dominantes, les trois panels expriment un optimisme marqué sur la capacité de l'IA à générer des avancées importantes dans la recherche scientifique et la santé. Les réponses « tout à fait d'accord » et « plutôt d'accord » dominent dans les trois panels, avec une convergence notable. Cette vision positive constitue l'une des rares dimensions d'enthousiasme net dans l'ensemble du sondage : c'est l'une des questions où l'accord est le plus fort et le plus homogène.

– **Inégalités sociales (Q30) :**

Le risque que l'IA renforce les inégalités sociales – avec des gains de productivité bénéficiant surtout à une minorité d'entrepreneurs tech – est perçu comme probable ou très probable dans les trois groupes, avec des scores concentrés entre 3 et 5. Néanmoins, alors que les membres du CRiP expriment des positions plus nuancées sur ce sujet, les répondants IJD, notamment francophones, montrent une préoccupation plus marquée pour la justice distributive des bénéfices de l'IA.

I 6. MAINTIEN DU CONTRÔLE HUMAIN À L'HORIZON 2035

– **Transformation sociétale par la collaboration humain-IA (Q26) :**

L'idée que la collaboration humain-IA va bouleverser la société fait l'objet d'un accord quasi-unanime (scores concentrés entre 4 et 5). C'est l'une des distributions les plus univoques de l'ensemble du sondage. Cette conviction partagée fonde l'urgence des questions de gouvernance et de régulation abordées par le reste du questionnaire.

– **Architectures d'interface et « kill switch » (Q31) :**

La question de la capacité des architectures d'interface, de l'ergonomie et des mécanismes de sécurité (dont les « *kill switches* ») à favoriser la maîtrise humaine sur les IA avancées reçoit

un accueil positif dans les trois groupes, avec des scores centrés entre 3 et 4, indiquant que ces outils sont perçus comme utiles. Le panel CRiP est naturellement le plus favorable à ces mécanismes techniques, qui relèvent de leur domaine d'expertise. Toutefois, ils ne sont pas considérés comme une panacée : l'absence de scores massivement concentrés sur 5 suggère que les répondants estiment que des solutions techniques seules ne suffisent pas et doivent s'inscrire dans un cadre réglementaire plus large, corroborant la notion de « l'humain dans la boucle ».

Implication pratique

Les répondants font confiance aux solutions techniques (interfaces, « kill switches ») pour maintenir le contrôle humain à court terme, mais expriment des doutes croissants pour le moyen terme, notamment avec l'émergence des IA agentiques.

— Gouvernance mondiale du numérique (Q25) :

La faisabilité d'une gouvernance mondiale incluant mécanismes multiniveaux (sanctions, normes ISO, accords ONU) reçoit un accueil mitigé mais optimiste dans les trois groupes. Les réponses « plutôt d'accord » dominent, mais « cela dépend » recueille également une part importante, reflétant un réalisme sur les difficultés géopolitiques. Le panel anglophone, plus institutionnel et international, est légèrement plus confiant dans la faisabilité de tels mécanismes.

— Risque de perte de contrôle à l'horizon 2035 (Q32) :

Cette question constitue l'un des marqueurs les plus révélateurs de l'ensemble du questionnaire. Les trois panels expriment une préoccupation significative et transversale quant au risque de perte de contrôle humain sur l'IA à horizon 2035, avec de potentielles conséquences sociales, économiques et politiques graves. Les distributions montrent des scores entre 3 et 5, sans toutefois atteindre le consensus des questions sur les bénéfices scientifiques ou la

dépendance technologique. Le panel CRiP est légèrement moins alarmiste, peut-être en raison de leur confiance dans les solutions techniques. Le panel IJD anglophone est le plus préoccupé, tandis que les répondants IJD francophones, plus nombreux et plus diversifiés sectoriellement, montrent une distribution légèrement plus étalée, mais la tendance générale reste à la préoccupation.

7. CONCLUSION : CONVERGENCES, DIVERGENCES ET RECOMMANDATIONS

— Convergences majeures entre les trois panels

- La transformation des métiers plutôt que leur suppression est le consensus le plus fort de l'ensemble du sondage — la formation aux compétences IA est la réponse prioritaire.
- Le rejet unanime de l'IA comme décideur autonome complet à horizon 2035, et l'adhésion au modèle de comité hybride avec supervision humaine.
- L'IA agentique est perçue comme un changement de paradigme, renforçant l'urgence de mécanismes de contrôle.
- Les biais, l'auditabilité et la transparence sont les principaux obstacles à une IA responsable.
- La dépendance technologique de l'UE vis-à-vis des USA et de la Chine est un problème qui appelle à la création de champions numériques européens.
- L'optimisme sur les applications médicales et scientifiques coexiste avec des inquiétudes profondes sur les inégalités et la désinformation.

— 7.2 Divergences et spécificités par panel

- Le panel CRiP (DSI/CTO) se distingue par une plus grande ouverture aux comités décisionnels hybrides, une sensibilité particulière aux questions de sécurité des infrastructures (Q24) et une relative confiance dans les solutions techniques. Son pragmatisme opérationnel le rend légèrement moins alarmiste sur la perte de contrôle à long terme.
- Le panel IJD anglophone, plus académique et international, affiche une préoccupation

plus marquée pour les risques de perte de contrôle à l'horizon 2035 et pour les droits fondamentaux face à la désinformation. Il est également plus favorable à une gouvernance mondiale multilatérale.

- Le panel IJD francophone, le plus large et le plus diversifié, reflète une tension entre l'enthousiasme pour les bénéfices de l'IA et l'inquiétude pour les inégalités sociales et la

protection des travailleurs.

— Recommandations pour maintenir l'humain dans la boucle

Les résultats croisés des trois panels dessinent une feuille de route cohérente articulée autour de six axes :

Priorités identifiées	Recommandations opérationnelles	Temporalité
Formation et montée en compétences	Priorité absolue identifiée dans les trois panels. Développer des programmes de formation IA à tous les niveaux, du salarié au cadre dirigeant.	Court terme (1-3 ans)
Architectures de contrôle technique	Investir dans les « <i>kill switches</i> », interfaces de supervision et mécanismes d'auditabilité.	Court terme (1-3 ans)
Mise en œuvre complète et consolidation de l'AI Act	Les trois panels jugent la réglementation actuelle insuffisante. Accélérer et consolider la mise en œuvre de l'AI Act au niveau européen et national.	Court à moyen terme (3-5 ans)
Renforcement de la protection des travailleurs	Les trois panels jugent également qu'il est nécessaire de renforcer les droits et les protections des travailleurs dans les secteurs vulnérables à l'automatisation.	Court à moyen terme (3-5 ans)
Souveraineté numérique européenne	Consensus fort sur la nécessité de faire émerger des champions européens dans le secteur numérique. Besoin d'une politique industrielle coordonnée à l'échelle de l'UE.	Moyen terme (5-10 ans)
Gouvernance mondiale multilatérale	Mécanismes ISO, ONU et coalitions inter-régionales. Priorité pour les systèmes IA à fort impact transnational (désinformation, santé, finance).	Moyen terme (5-10 ans)

— Conclusion : une convergence sur l'impératif du contrôle humain

En définitive, les trois panels convergent vers une vision cohérente : l'IA est une révolution rapide et profonde qui doit être accompagnée plutôt que subie. Maintenir l'humain dans la boucle ne signifie pas freiner l'innovation, mais construire les conditions – techniques, réglementaires, éducatives et géopolitiques – d'une IA qui reste un outil au service des choix humains plutôt qu'un substitut à la délibération démocratique. L'analyse comparative de ces trois sondages fait ressortir une conclusion centrale : qu'importe les répondants interrogés, la nécessité de maintenir un contrôle humain effectif sur l'IA fait l'objet d'un large consensus. Ce consensus ne traduit pas un rejet de l'IA, dont les bénéfices

potentiels sont largement reconnus, mais une exigence de gouvernance.

Les divergences entre groupes sont davantage de degré que de nature. Les membres du CRiP, immergés dans la transformation numérique de leurs organisations, tendent à percevoir l'IA comme déjà incontournable et adoptent une posture plus pragmatique sur la gouvernance hybride. Les répondants IJD, plus proches des sphères académiques et institutionnelles, insistent davantage sur les risques systémiques et les impératifs réglementaires. La convergence la plus frappante reste celle autour de la peur d'une perte de contrôle à l'horizon 2035 : c'est précisément ce risque qui justifie, aux yeux des répondants, l'ensemble du cadre réglementaire – de l'AI Act aux normes ISO en passant par

les mécanismes d'urgence techniques. Le message central de ces sondages est clair : l'avenir de l'IA se jouera autant sur le plan technologique que sur celui de la gouvernance et de la capacité des sociétés à définir et imposer des limites au nom de l'intérêt général.

Malgré l'exhaustivité des personnes interrogées dans le cadre de ce sondage, ces résultats n'ont pas de vocation universelle. En effet, il y a certaines limites en termes de généralisation, puisque la grande majorité de l'échantillon provient du réseau de l'Institut Jacques Delors incluant des institutions internationales, administrations publiques, universités, think tanks, ONGs. Même si l'échantillon des membres du CRiP permet d'élargir le spectre d'opinion vers le secteur privé et le monde de l'entreprise, celui-ci est moins bien représenté par rapport aux autres secteurs en termes du nombre de réponses. Une autre limite est la focalisation géographique sur l'Europe dans les trois versions du sondage.

2 • ANNEXE 2 :

Clause de non-responsabilité (« Disclaimer ») : Élaborer une proposition législative complète est un travail exhaustif qui va au-delà du cadre de cette étude. Pour ces raisons, l'objectif de cette 2^e Annexe est d'offrir une application concrète des solutions développées dans l'étude, afin de donner un ancrage plus réaliste aux recommandations politiques. Les deux propositions ci-dessous visent donc à prodiguer des éléments de réflexion qui pourraient peut-être former une base liminaire pour de « vrais » projets de lois à l'avenir, mais ne doivent en aucun cas être considérées ou traitées comme des propositions de lois authentiques en tant que telles. C'est pour cette raison que les montants financiers ne sont pas précisés (indiqués avec « XXX »).

• PROPOSITIONS LÉGISLATIVES

Intelligence artificielle, protection des travailleurs et formation

I DOCUMENT I

Proposition de directive de l'Union européenne

I DOCUMENT II

Proposition de loi française de transposition et mesures complémentaires

I DOCUMENT I

PROPOSITION DE DIRECTIVE DU PARLEMENT EUROPÉEN ET DU CONSEIL relative à la protection des travailleurs face à l'automatisation par l'intelligence artificielle, à la formation professionnelle et au financement de la transition numérique

Le Parlement Européen et le Conseil de l'Union Européenne, considérant les articles 153 (droits

des travailleurs), 166 (formation professionnelle), 114 (marché intérieur) et 173 (politique industrielle) du Traité sur le fonctionnement de l'Union européenne (TFUE), ont adopté la présente directive.

La directive respecte le principe de subsidiarité : elle fixe des exigences minimales communes tout en laissant aux États membres la latitude de maintenir ou d'introduire des dispositions plus favorables.

– CHAPITRE I – OBJET, CHAMP D'APPLICATION ET DÉFINITIONS

• Article 1 – Objet

La présente directive établit un cadre minimum commun visant à : (i) protéger les travailleurs dont les emplois sont exposés à une automatisation substantielle par des systèmes d'intelligence artificielle (IA) ; (ii) soutenir leur droit à la formation et à l'employabilité tout au long de leur vie professionnelle ; (iii) assurer la transparence des algorithmes déployés sur le lieu de travail ; (iv) financer la transition numérique par des mécanismes fiscaux harmonisés.

• Article 2 – Champ d'application

La directive s'applique à tout employeur établi dans l'Union européenne qui déploie des systèmes d'IA susceptibles d'automatiser, de surveiller ou d'évaluer des tâches exercées par des travailleurs salariés ou assimilés. Elle s'applique sans préjudice des dispositions plus favorables en vigueur dans les États membres.

• Article 3 – Définitions

- « Système d'IA à impact sur l'emploi » : tout système au sens du règlement (UE) 2024/1689 (règlement européen sur l'IA) dont le déploiement entraîne ou est susceptible d'entraîner la suppression, la transformation ou la surveillance automatisée d'au moins 20 % des tâches d'un poste de travail.
- « Obligation de maintien de l'employabilité » (*Upskilling Obligation*) : obligation légale pesant sur l'employeur de garantir l'accès à un parcours de formation certifiant pour tout

travailleur dont le poste est concerné à plus de 50 % par l'automatisation.

- « Augmentation humaine » : démarche par laquelle l'IA assiste le travailleur sans se substituer à ses décisions critiques ni à sa relation avec les tiers.
- « Fonds européen de transition numérique » : mécanisme financier européen dédié au financement des congés de conversion et des formations de longue durée.

– CHAPITRE II – DROITS DES TRAVAILLEURS ET OBLIGATIONS DES EMPLOYEURS

• Article 4 – Obligation de maintien de l'employabilité

Toute entreprise employant 50 salariés ou plus est tenue, avant tout déploiement d'un système d'IA à impact sur l'emploi :

- d'identifier les postes concernés par une automatisation supérieure à 50 % des tâches ;
- de proposer un parcours individualisé de montée en compétences vers des fonctions à valeur ajoutée supérieure (conception, supervision, relation client complexe) assistées par l'IA ;
- de consacrer un pourcentage de la masse salariale, fixé par les États membres dans la fourchette de 0,5 % à 1,5 %, au financement de ces formations.

Cette obligation est particulièrement applicable aux travailleurs jeunes diplômés en début de carrière, ainsi qu'aux travailleurs de 55 ans ou plus.

• Article 5 – Droit à la transparence algorithmique

Avant tout déploiement ou mise à jour substantielle d'un système d'IA au sens de l'article 3, l'employeur publie une « notice d'impact sur les compétences » précisant :

- les compétences humaines renforcées par le système d'IA ;
- celles susceptibles de s'atrophier ou d'être substituées ;
- les formations compensatoires proposées.

Les représentants des travailleurs disposent d'un droit de consultation préalable d'un délai minimum de 60 jours. Les États membres peuvent prévoir un droit d'opposition temporaire des comités d'entreprise en cas d'évaluation insuffisante des risques.

- **Article 6** – Comités de supervision numérique

Dans les entreprises de 50 salariés ou plus, un comité de supervision numérique composé paritairement de représentants de la direction et des salariés réalise annuellement un audit de l'utilisation de l'IA, vérifie le respect de l'obligation de maintien de l'employabilité et s'assure que l'IA n'est pas utilisée à des fins de substitution totale des décisions à caractère humain.

- **Article 7** – Index européen de l'augmentation humaine

La Commission européenne établit un index public notant les entreprises de plus de 250 salariés sur leur capacité à maintenir l'emploi tout en intégrant l'IA. Les entreprises affichant la meilleure notation bénéficient d'un accès prioritaire aux marchés publics européens et aux financements des fonds structurels.

– CHAPITRE III – FORMATION PROFESSIONNELLE ET PASSEPORT EUROPÉEN DE COMPÉTENCES IA

- **Article 8** – Passeport européen de compétences IA

La Commission crée un système de certification interopérable dans toute l'Union européenne, le « Passeport européen de compétences IA », fondé sur le cadre DigComp (*Digital Competence Framework*) mis à jour. Ce passeport atteste la capacité d'un travailleur à collaborer avec des systèmes d'IA et facilite la mobilité professionnelle vers des postes augmentés. Il est reconnu de plein droit dans tous les États membres.

- **Article 9** – Fonds européen de transition numérique

Il est créé, sur le modèle du Fonds pour une transition juste, un Fonds européen de transition numérique doté d'une enveloppe pluriannuelle d'au moins XXX milliards d'euros à l'horizon 2030. Celui-ci finance :

- des « congés de conversion IA » d'une durée de 6 à 12 mois destinés aux travailleurs dont les emplois sont exposés à une automatisation à plus de 50% des tâches, rémunérés à 90 % du salaire brut de référence pour le 1^{er} mois, et ensuite 70% pour les mois suivants ;
- des micro-accréditations sectorielles obligatoires dans les domaines stratégiques : santé, finance, droit, cybersécurité, infrastructures numériques ;
- des allocations de transition pour les travailleurs en attente de reconversion.

La moitié du financement (50%) proviendra du budget de l'UE, incluant le Cadre financier pluriannuel et des mécanismes d'emprunt commun (sur le modèle de *NextGenerationEU*). L'autre moitié proviendra d'un mélange de cofinancement par les États membres, de capitaux privés (investisseurs/capital-risque), ainsi que de nouvelles « ressources propres technologiques » (taxes IA/données), dont les mécanismes fiscaux sont explicités au chapitre IV. Chaque État membre sera libre de fixer la répartition exacte des composantes de cette seconde moitié du financement.

- **Article 10** – Cinquième liberté du marché intérieur de l'UE – Éducation, recherche et innovation

La Commission européenne est invitée à proposer, dans un délai de 18 mois suivant l'entrée en vigueur de la présente directive, une initiative législative que le Parlement et le Conseil de l'Union Européenne ont vocation à adopter consacrant une cinquième liberté fondamentale du marché intérieur, dédiée spécifiquement à la libre circulation des connaissances, des formations certifiantes et des données de recherche, incluant celles liées à l'IA.

– CHAPITRE IV – FISCALITÉ ET FINANCEMENT

• Article 11 – Cadre fiscal harmonisé

Les États membres adoptent, dans le respect de leur souveraineté fiscale et des règles relatives aux aides d'État, au moins l'une des mesures suivantes pour financer la transition numérique :

- une contribution sur les gains de productivité liés à l'automatisation, sous forme de surtaxe sur l'impôt des sociétés appliquée aux entreprises dont le ratio chiffre d'affaires/effectif dépasse un seuil sectoriel défini par décret ;
- la suppression ou la réduction progressive des avantages fiscaux liés à l'amortissement accéléré de logiciels d'IA de substitution au travail humain ;
- une taxe proportionnelle sur l'usage des données – plus une IA utilise de données pour s'entraîner et fonctionner, plus l'entreprise paie une redevance – afin de redistribuer la valeur créée par l'IA vers le bien public ;
- une « taxe sur l'usage » ou l'équivalence salariale, traitant l'IA comme un employé virtuel. L'entreprise paie une taxe équivalente aux cotisations sociales qu'elle aurait versées si un humain occupait le poste, avec l'objectif de maintenir le financement de la protection sociale (retraites, santé).

La Commission présentera, au plus tard le 1er janvier 2028, une proposition d'assiette commune pour l'Union en vue d'une harmonisation renforcée.

• Article 12 – Crédit d'impôt augmentation humaine

Les États membres peuvent instituer un crédit d'impôt spécifique en faveur des entreprises qui investissent dans des outils d'IA conçus pour l'assistance du travailleur (« cobotique », copilotes de codage, outils d'aide au diagnostic) plutôt que dans des systèmes d'automatisation complète. Ce crédit est conditionné à la justification d'une formation certifiante pour l'ensemble des employés concernés.

– CHAPITRE V – MÉCANISMES DE CONTRÔLE ET SANCTIONS

• Article 13 – Autorités nationales compétentes

Chaque État membre désigne une autorité nationale compétente chargée de contrôler l'application de la présente directive. Cette autorité coordonne son action avec la Commission européenne, ainsi qu'avec l'Autorité européenne du travail (AET) et le Bureau européen de l'IA (*European AI Office*) dans le cadre du règlement sur l'IA.

• Article 14 – Sanctions

Les États membres prévoient des sanctions effectives, proportionnées et dissuasives. Les sanctions pécuniaires pour violation de l'obligation de maintien de l'employabilité (article 4) ou du droit à la transparence algorithmique (article 5) ne peuvent être inférieures à :

- 0,5 % du chiffre d'affaires annuel mondial pour une première violation ;
- 1 % pour une récidive dans un délai de 3 ans.

En cas de manquement grave, les États membres peuvent prévoir la suspension du droit d'utiliser le système d'IA concerné jusqu'à mise en conformité.

• Article 15 – Recours et protection des lanceurs d'alerte

Les travailleurs disposent d'un droit de recours individuel ou collectif devant les juridictions compétentes en cas de violation de leurs droits découlant de la présente directive. Les représentants des travailleurs qui signalent une violation de la présente directive bénéficient des protections prévues par la directive (UE) 2019/1937 sur la protection des personnes qui signalent des violations du droit de l'Union européenne.

— CHAPITRE VI – CALENDRIER D'APPLICATION

- **Article 16** – Transposition et mise en œuvre

Les États membres transposent la présente directive selon le calendrier suivant :

- au plus tard le X 2027 : Transposition dans les droits nationaux des articles 4 (employabilité), 5 (transparence) et 13 (autorités de contrôle).
- au plus tard le X 2028 : Mise en place des comités de supervision numérique (article 6), lancement du Passeport européen de compétences IA (article 8) et adoption des premières mesures fiscales nationales (article 11).
- au plus tard le X 2030 : Pleine opérationnalité du Fonds européen de transition numérique (article 9), publication de l'index européen de l'augmentation humaine (article 7) et harmonisation fiscale renforcée (article 11 alinéa 3).

La Commission procède à une évaluation de la directive au plus tard le 31 décembre 2030 et, si nécessaire, propose des révisions au regard des évolutions technologiques.

- **Article 19** – Entrée en Vigueur

La présente directive entre en vigueur le vingtième jour suivant celui de sa publication au *Journal officiel de l'Union européenne*.

- **Article 20** – Destinataires

Les États membres sont destinataires de la présente directive.

Fait à Strasbourg, le X.
Par le Parlement européen
Par le Conseil

I DOCUMENT II

PROPOSITION DE LOI FRANÇAISE relative à l'adaptation du droit du travail français à l'intelligence artificielle, à la protection des travailleurs, à la formation professionnelle et à la fiscalité du numérique

L'Assemblée Nationale et le Sénat ont adopté, vu la décision X du Conseil Constitutionnel, le Président de la République promulgue la loi dont la teneur suit :

La présente loi transpose la directive (UE) relative à la protection des travailleurs face à l'automatisation par l'intelligence artificielle, à la formation professionnelle et au financement de la transition numérique. Elle met en œuvre, dans le cadre des compétences nationales, des mesures complémentaires relevant du droit du travail, de la fiscalité, de la négociation collective et de la formation professionnelle. Elle s'inscrit dans le respect du bloc de constitutionnalité et du principe de participation des travailleurs garanti par le huitième alinéa du Préambule de la Constitution de 1946.

— TITRE I – PROTECTION DES TRAVAILLEURS FACE À L'AUTOMATISATION

- **Article 1** – Garantie d'employabilité IA – Obligation de maintien

Après l'article L. 6321-1 du Code du travail, il est inséré un article L. 6321-2 ainsi rédigé :

« Dans les entreprises d'au moins 50 salariés, lorsqu'un système d'intelligence artificielle est susceptible d'automatiser plus de 50 % des tâches d'un poste, l'employeur est tenu d'élaborer, en concertation avec les représentants du personnel, un parcours individualisé de maintien de l'employabilité. Ce parcours comporte à minima un accès à une formation certifiante vers des fonctions de supervision, de conception ou de relation client complexe assistée par l'IA. L'employeur consacre à cet effet au moins 1 % de la masse salariale dans les entreprises concernées. »

Cette obligation s'applique de manière prioritaire aux jeunes travailleurs diplômés en début de carrière, ainsi qu'aux travailleurs âgés de 55 ou plus, pour lesquels l'employeur présente un programme spécifique au comité social et économique (CSE).

- **Article 2 – Comités de supervision numérique**

Dans les entreprises d'au moins 50 salariés, le comité social et économique (CSE) exerce, en sus de ses attributions actuelles, les compétences d'un comité de supervision numérique. À ce titre, il :

- est consulté au moins 60 jours avant tout déploiement ou mise à jour substantielle d'un système d'IA ayant un impact sur les conditions de travail ou sur les emplois ;
- reçoit annuellement un rapport d'audit sur l'utilisation de l'IA dans l'entreprise comprenant l'évaluation de son impact sur l'emploi et l'autonomie des travailleurs ;
- peut exercer un droit d'opposition temporaire de 30 jours, renouvelable une fois, si l'évaluation des risques est jugée insuffisante ; en cas de désaccord persistant, la Direction régionale de l'économie, de l'emploi, du travail et des solidarités (DREETS) est saisie.

- **Article 3 – Clause de non-substitution technologique**

Il est inséré dans la partie relative aux négociations obligatoires en entreprise du Code du travail (section L. 2242) une obligation de négocier, tous les trois ans, une clause de non-substitution technologique dans les branches professionnelles et dans les entreprises d'au moins 50 salariés. Cette clause garantit que les systèmes d'IA ne peuvent être utilisés que pour traiter des volumes supplémentaires ou des tâches à faible valeur ajoutée, les décisions critiques et les interactions humaines complexes demeurant de la compétence exclusive de travailleurs formés.

- **Article 4 – Droit à la transparence algorithmique au travail**

Avant tout déploiement d'un système d'IA au travail, l'employeur remet à chaque salarié concerné et aux représentants du personnel une « notice d'impact sur les compétences » précisant les compétences humaines renforcées, celles exposées à l'atrophie, ainsi que les formations proposées en compensation. Cette notice est versée au dossier du salarié et opposable en cas de contentieux prud'homal.

- **Article 5 – Réduction du temps de travail et partage des gains de productivité**

Les accords de branche et d'entreprise peuvent prévoir, dans le cadre des négociations sur le temps de travail et la rémunération, une réduction du temps de travail sans baisse de salaire lorsque des gains de productivité mesurables résultent du déploiement de systèmes d'IA. Le temps libéré peut être affecté prioritairement à des actions de formation certifiante. La méthodologie de calcul des gains de productivité est définie par décret en Conseil d'État.

– TITRE II – FORMATION PROFESSIONNELLE ET RECONVERSION

- **Article 6 – Fonds national de transition numérique**

Il est créé un Fonds national de transition numérique, géré par France compétences en lien avec les opérateurs de compétences (OPCO), abondé par les contributions des employeurs et par les recettes fiscales prévues au titre III. Ce fonds finance :

- des « congés de conversion IA » d'une durée de 6 à 12 mois destinés aux travailleurs dont les emplois sont menacés, rémunérés à 90 % du salaire brut de référence pour le 1er mois, et ensuite 70% pour les mois suivants, sur la base d'un accord individuel conclu avec l'employeur et validé par la DREETS ;
- des allocations de transition pour les demandeurs d'emploi dont le poste a été supprimé en raison de l'automatisation ;

- des formations longues dans les secteurs en expansion identifiés : santé et biotechnologie, cybersécurité, infrastructures numériques, services à la personne et métiers manuels qualifiés.
- **Article 7** – Micro-accréditations sectorielles obligatoires

Dans les branches professionnelles des secteurs bancaire, assurantiel, juridique et de santé, des micro-accréditations spécifiques à l'IA sont rendues obligatoires par accord de branche homologué par le ministère du Travail. La liste des branches concernées est actualisée annuellement par arrêté. À titre d'exemple, un architecte utilisant des outils d'IA générative dans le cadre de ses missions doit valider un module d'éthique et de pilotage de la conception assistée pour conserver son droit de signature professionnelle.

- **Article 8** – Passeport national de compétences IA

En articulation avec le Passeport européen de compétences IA instauré par la directive, il est créé un Passeport national de compétences IA interopérable avec le compte personnel de formation (CPF) et le passeport de prévention. Le Passeport national atteste des formations suivies et des certifications obtenues en matière de collaboration humain-IA. Sa reconnaissance dans les procédures de recrutement au sein des marchés publics nationaux est obligatoire pour les appels d'offres supérieurs à XXX euros.

- **Article 9** – Formations ciblées vers les secteurs à expansion

Le service public de l'emploi (France Travail), en lien avec les conseils régionaux, développe une offre de formation spécifiquement orientée vers les secteurs dans lesquels l'IA crée de nouvelles opportunités d'emploi. Sont identifiés à titre prioritaire :

- la santé et la biotechnologie : bio-informaticiens, spécialistes de la médecine génomique, experts en diagnostic radiologique augmenté ;

- la cybersécurité et l'éthique algorithmique : éthiciens de l'IA, responsables de conformité algorithmique, analystes en cyberdéfense prédictive ;
- les infrastructures numériques et la gestion des données : ingénieurs de centres de données, architectes de solutions informatiques en périphéries de réseaux, ingénieurs en données ;
- les services à la personne, l'éducation et les métiers manuels qualifiés, dont la valeur ajoutée est renforcée par l'IA en raison du caractère irremplaçable des compétences interpersonnelles et de la gestuelle complexe impliquées par ces métiers.

– TITRE III – FISCALITÉ DU NUMÉRIQUE ET FINANCEMENT

- **Article 10** – Contribution sociale sur les gains d'automatisation

Il est institué, dans le Code général des impôts, une contribution additionnelle à l'impôt sur les sociétés applicable aux entreprises dont le ratio chiffre d'affaires consolidé rapporté aux effectifs salariés dépasse, par secteur, un seuil fixé par décret. Cette contribution est assise sur la fraction du bénéfice imputable à l'automatisation, évaluée selon une méthode standardisée fixée par arrêté ministériel. Son produit est affecté au Fonds national de transition numérique (article 6).

- **Article 11** – Suppression des avantages fiscaux à l'automatisation substitutive

L'amortissement accéléré prévu à l'article 39 A du Code général des impôts est exclu pour les logiciels et systèmes d'IA dont la finalité principale est la suppression nette de postes de travail, telle que documentée dans la notice d'impact sur les compétences (article 4). Les entreprises bénéficiant d'un crédit d'impôt formation (article 12) ne peuvent cumuler cet avantage avec un amortissement accéléré sur les mêmes investissements en IA.

- **Article 12** – Crédit d'impôt augmentation humaine

Les entreprises qui investissent dans des systèmes d'IA conçus pour assister le travailleur sans le substituer (« cobotique », copilotes de codage, outils d'aide au diagnostic médical) bénéficient d'un crédit d'impôt de 30 % des dépenses éligibles, plafonné à XXX euros par an. Ce crédit est conditionné à la justification d'une formation certifiante pour l'ensemble des salariés concernés et à la production d'un rapport annuel remis au comité social et économique (CSE).

- **Article 13** – Redevance sur l'usage des données

Les entreprises dont le chiffre d'affaires lié à l'entraînement ou à l'exploitation commerciale de modèles d'IA dépasse XXX millions d'euros annuels en France acquittent une redevance proportionnelle à la valeur des données collectées sur le territoire national. Le produit de cette redevance est partiellement affecté au Fonds national de transition numérique et partiellement au budget de l'État pour le financement de la stratégie nationale de cybersécurité.

– **TITRE IV – INDEX DE L'AUGMENTATION HUMAINE ET TRANSPARENCE**

- **Article 14** – Index national de l'augmentation humaine

Sur le modèle de l'index de l'égalité professionnelle femmes-hommes, il est créé un index national public de l'augmentation humaine. Les entreprises d'au moins 250 salariés se voient attribuer une note sur une échelle de 0 à 100 par l'autorité nationale de régulation de l'IA créée dans le cadre de la mise en œuvre du règlement (UE) 2024/1689 sur l'IA. Cette note est calculée selon un référentiel fixé par décret intégrant le maintien des effectifs, l'effort de formation, la qualité du dialogue social relatif à l'IA et l'application de la clause de non-substitution et doit être publiée annuellement par les entreprises concernées. Les entreprises notées au-dessus de 75/100 bénéficient d'un accès prioritaire aux marchés publics nationaux et d'un abattement

de 5 % sur leur contribution sociale d'automatisation.

– **TITRE V – MÉCANISMES DE CONTRÔLE ET SANCTIONS**

- **Article 15** – Autorité de contrôle

Les DREETS sont compétentes pour contrôler l'application de la présente loi. Elles agissent en coordination avec l'inspection du travail, la CNIL (pour les questions de données personnelles) et l'autorité nationale de régulation de l'IA créée dans le cadre de la mise en œuvre du règlement européen sur l'IA. Un rapport annuel relatif au contrôle de l'application de la loi est transmis à l'Assemblée Nationale et au Sénat.

- **Article 16** – Sanctions administratives

Le non-respect de l'obligation de maintien de l'employabilité (article 1), de consultation du CSE (article 2) ou de remise de la notice d'impact (article 4) est sanctionné par :

- une mise en demeure de mise en conformité dans un délai de 3 mois ;
- en l'absence de régularisation, une amende administrative de 0,5 % de la masse salariale pour une première violation et de 1 % en cas de récidive dans un délai de 3 ans ;
- la suspension du droit de déployer le système d'IA concerné jusqu'à régularisation, prononcée par la DREETS.

- **Article 17** – Sanctions en droit du travail et protection des représentants

Tout licenciement ou mesure défavorable visant un salarié ou un représentant du personnel en raison de l'exercice de ses droits prévus par la présente loi est nul de plein droit. Les représentants du personnel bénéficient de la protection élargie prévue par la législation sur les lanceurs d'alerte lorsqu'ils signalent une violation.

– TITRE VI – CALENDRIER D'APPLICATION

• Article 18 – Entrée en vigueur

Les dispositions de la présente loi entrent en vigueur selon le calendrier suivant :

- dès promulgation (2026) : articles 2 (comités de supervision), 4 (transparence algorithmique) et 12 (crédit d'impôt augmentation humaine).
- au plus tard le X 2027 : articles 1 (obligation de maintien de l'employabilité), 3 (clause de non-substitution), 6 (Fonds national), 8 (Passport national IA) et 15 (sanctions).
- au plus tard le X 2028 : articles 7 (micro-ac créditations sectorielles), 9 (formations sectorielles), 10 (contribution sociale d'automatisation), 11 (révision des avantages fiscaux), 13 (redevance données) et 14 (index augmentation humaine).
- au plus tard le X 2029 au plus tard : pleine application de l'ensemble des dispositions ; remise d'un rapport d'évaluation au Parlement avec, le cas échéant, propositions de révision.

• Article 19 – Dispositions réglementaires

Les décrets en Conseil d'État nécessaires à l'application de la présente loi seront publiés dans les 12 mois suivant sa promulgation pour les dispositions entrant en vigueur en 2026-2027, et dans les 24 mois pour les dispositions entrant en vigueur en 2028-2029.

La présente loi sera exécutée comme loi de l'État

Fait à Paris le X,

Directeur de la publication: Sylvie Matelly • La reproduction en totalité ou par extraits de cette contribution est autorisée à la double condition de ne pas en dénaturer le sens et d'en mentionner la source • Les opinions exprimées n'engagent que la responsabilité de leur(s) auteur(s) • L'Institut Jacques Delors ne saurait être rendu responsable de l'utilisation par un tiers de cette contribution • Version originale • © Notre Europe - Institut Jacques Delors • Image de couverture © Getty Images

Institut Jacques Delors
Penser l'Europe • Thinking Europe • Europa Denken
17 rue d'Antin, 75002 Paris, France
www.delorsinstitute.eu • info@delorsinstitute.eu