HAL open science

IA : Notre Ambition Pour La France

Philippe Aghion, Anne Bouverot, Arno Amabile, Cyprien Canivenc, Gilles Babinet, Joëlle Barral, Alexandra Bensamoun, Nozha Boujemaa, Bertrand Charlès, Luc Julia, et al.

To cite this version:

Philippe Aghion, Anne Bouverot, Arno Amabile, Cyprien Canivenc, Gilles Babinet, et al.. IA : Notre Ambition Pour La France. Gouvernement Français. 2024. ⟨hal-04825691⟩

HAL Id: hal-04825691

https://hal.science/hal-04825691v1

Submitted on 16 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Autorisation HAL

HAL Authorization

GOUVERNEMENT

Liberté
Égalité
Fraternité

COMMISSION DE L'INTELLIGENCE ARTIFICIELLE

IA : NOTRE AMBITION POUR LA FRANCE

MARS 2024

SOMMAIRE

SYNTHÈSE GÉNÉRALE, NOTRE VISION ET NOS RECOMMANDATIONS CLÉS 4
MÉTHODE. 14
INTRODUCTION 17
L'intelligence artificielle n'est pas une nouveauté : elle date des années 1950 17
L'IA recouvre un ensemble d'outils numériques, déjà omniprésents dans notre société 18
L'IA générative constitue un tournant majeur de cette histoire de l'innovation 18
Comment fonctionne l'intelligence artificielle générative ? 20
Comment se décompose la chaîne de valeur économique de l'IA générative ? 21
En France et en Europe, nous accusons un net retard… 23
…mais nous disposons d'atouts : rien n'est donc définitivement joué 23
Quelques débats au cœur du plan d'action que nous proposons 24
Le débat incontournable : à quoi ressemblera le monde de demain ? 26

1. DÉDIABOLISER L'IA, SANS POUR AUTANT L'IDÉALISER. . . . 28
1.1 L'IA ME CONCERNE-T-ELLE ? 29
1.2 FAUT-IL AVOIR PEUR DE L'IA ?. 31
1.3 L'IA NOUS RENDRA-T-ELLE PLUS PROSPÈRES ? 34
1.4 L'IA : CRÉATRICE OU DESTRUCTRICE D'EMPLOIS ? 41
1.5 L'IA VA-T-ELLE DÉGRADER OU AMÉLIORER LA QUALITÉ DE VIE AU TRAVAIL ? . 48
1.6 L'IA MET-ELLE EN DANGER LA CRÉATION ARTISTIQUE ? 52
1.7 L'IA PEUT-ELLE NUIRE À LA QUALITÉ DE L'INFORMATION ? 54
1.8 FAUT-IL DIFFUSER L'IA DES AUTRES OU CRÉER LA NÔTRE ? 56
1.9 L'IA DEVRAIT-ELLE RESTER ENTRE LES MAINS DE QUELQUES ACTEURS ? 58
1.10 L'IA MET-ELLE EN DANGER LA PLANÈTE ?. 61
1.11 Y A-T-IL UNE BULLE DANS L'IA GÉNÉRATIVE ? 63
1.12 DOIT-ON SE PRÉPARER À UNE IA PLUS INTELLIGENTE QUE NOUS ? 65

2. HUMANISME, SOUVERAINETÉ, RESPONSABILITÉ : INNOVONS, DÉPLOYONS ET MAÎTRISONS L'IA 67

2.1 HUMANISME : PLAÇONS L'IA À NOTRE SERVICE 68

2.1.1. Faire du dialogue social et de la co-construction la pierre angulaire du recours à l'IA 68

2.1.2 Former : sans délai, massivement et en continu 70

2.1.3 Équiper les agents publics : une opportunité pour transformer l'administration 76

2.1.4 Mieux soigner grâce à l'IA : plus de temps au soin 81

2.1.5 Mieux éduquer grâce à l'IA : l'accompagnement individualisé des élèves 83

2.2 SOUVERAINETÉ : INVESTIR POUR NOTRE AUTONOMIE STRATÉGIQUE 86

2.2.1 Financer durablement l'innovation : l'indispensable changement d'échelle . . .88

2.2.2 Disposer de capacités de calcul souveraines : une condition *sine qua non* de l'autonomie stratégique 91

2.2.3 Accéder à des données de qualité 99

2.2.4 Attirer les talents pour construire les technologies et les usages de demain . . 106

2.2.5 Déployer massivement l'intelligence artificielle dans notre économie 110

2.3. RESPONSABILITÉ : MAÎTRISER, AUDITER, PROTÉGER 113

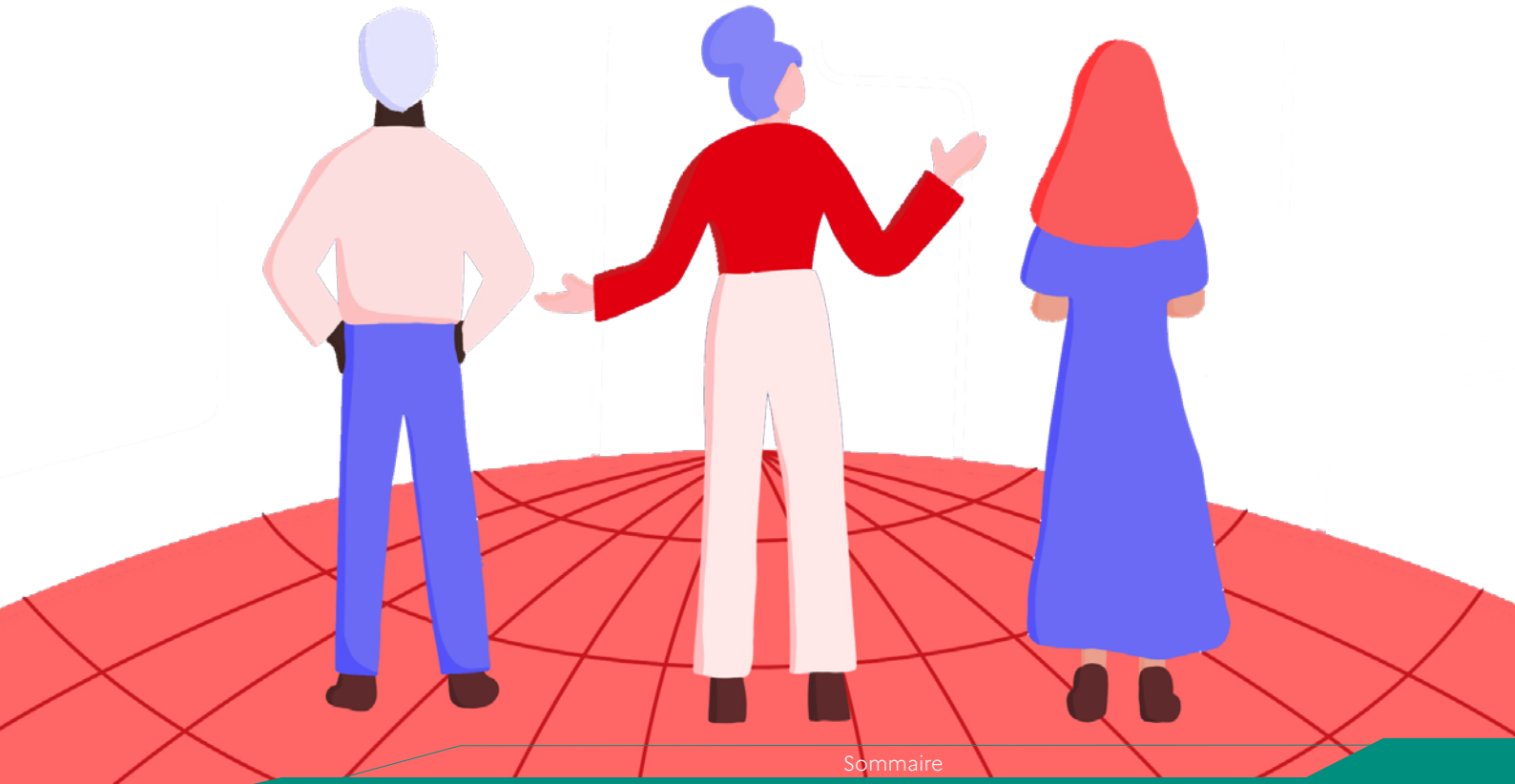
2.3.1 Bâtir une gouvernance internationale qui fait aujourd'hui défaut 113

2.3.2 Disposer en France d'une capacité d'évaluation des systèmes d'IA 117

2.3.3 Éviter les positions concurrentielles dominantes 121

NOS RECOMMANDATIONS 124

LETTRE DE MISSION 128



SYNTHÈSE GÉNÉRALE, NOTRE VISION ET NOS RECOMMANDATIONS CLÉS

L'intelligence artificielle est une révolution technologique incontournable. L'émergence soudaine et la diffusion de l'IA générative marquent une étape importante de cette révolution. Nous ne pouvons que constater la simplicité d'utilisation de certains outils, la rapidité de la génération du contenu, le réalisme des textes, images et sons générés, et plus généralement les aptitudes des récents modèles d'IA.

Cette révolution technologique affecte tous les domaines d'activité. Elle a des effets sur l'économie, l'emploi, les services publics, l'environnement, l'information, le secteur culturel… Tous les pans de notre société sont concernés et le seront davantage à l'avenir, tant son potentiel est considérable.

L'IA ne doit susciter ni excès de pessimisme, ni excès d'optimisme : nous n'anticipons ni chômage de masse, ni accélération automatique de la croissance. Dans les prochaines années, l'IA ne remplacera pas l'humain, de même qu'elle ne sera pas la solution à tous les défis de notre temps. Nous ne devons ni surestimer l'impact à très court terme, ni le sous-estimer à long terme.

L'Europe et la France ont des atouts pour être des acteurs de cette révolution, en premier lieu du fait de l'excellence de nos talents. Cette richesse et le dynamisme exceptionnel de l'écosystème français dans l'IA ne doivent néanmoins pas masquer une réalité préoccupante. Depuis plusieurs décennies, la tendance est celle d'un déclassement technologique et économique de notre continent, qui hypothèque sa prospérité et son indépendance.

Alors que les États-Unis et la Chine ont fait de la maîtrise de l'IA l'un des piliers de leur stratégie de puissance, nous devons relever le défi de l'IA, faute de quoi nous n'aurons pas la maîtrise de notre avenir. Il faut réformer nos institutions et nos politiques publiques, pour que l'IA puisse être pleinement un facteur de progrès.

Nous proposons six grandes lignes d'action :

- **lancer immédiatement un plan de sensibilisation et de formation de la nation** : animation de débats publics en continu sur les impacts économiques et sociétaux de l'IA au plus près des lieux du quotidien, structuration de l'offre de formation d'enseignement supérieur, massification de la formation continue aux outils d'IA, intégration de l'IA comme objet et outil du dialogue social ;

- **réorienter structurellement l'épargne vers l'innovation et créer, à court terme, un fonds « France & IA » de 10 Md€**, pour financer l'émergence de l'écosystème d'IA et la transformation du tissu économique français ;
- **faire de la France un pôle majeur de la puissance de calcul** : approvisionnement collectif sécurisé d'ampleur nationale et européenne, appel à projets d'implantation de centres de calcul avec garantie publique d'utilisation et simplification des procédures, crédit d'impôt IA pour l'entraînement de modèles ;
- **faciliter l'accès aux données** : en matière de données à caractère personnel, modernisation du mandat de la Commission nationale de l'informatique et des libertés (Cnil) et de son collège, suppression de certaines procédures d'autorisation préalable d'accès aux données de santé et réduction des délais de réponse ; en matière culturelle, mise en place de l'infrastructure technique favorisant l'entraînement des modèles d'IA dans le respect des droits de propriété intellectuelle ;
- **assumer le principe d'une « exception IA » dans la recherche publique** : libération des chercheurs des contraintes administratives, revalorisation de leur rémunération, doublement des moyens de la recherche publique spécialisée en IA ;
- **promouvoir une gouvernance mondiale de l'IA** : création d'une Organisation mondiale de l'IA pour évaluer et encadrer les systèmes d'IA, d'un Fonds international pour l'IA au service de l'intérêt général et d'un mécanisme de solidarité « 1 % IA » pour les pays en voie de développement.

Une mobilisation collective, massive, sans délai et au long cours est impérative. C'est dans cette perspective que notre Commission s'est attachée à élaborer un plan d'action aussi ambitieux que réaliste, au service des personnes, de nos besoins, de nos valeurs et de nos principes. Le plan représente un investissement public annuel de 5 Md€ pendant cinq ans. Il comprend des investissements technologiques mais aussi des investissements pour catalyser à la fois la diffusion de l'IA dans l'économie, son déploiement au service des citoyens et une appropriation et formation de toute la société.

Cet investissement est significatif mais il est nécessaire pour faire de la France un pays à la pointe dans l'intelligence artificielle et pour que notre société en tire tous les bénéfices. Cette ambition est atteignable, étant donné les atouts de la France et de l'Europe. Elle est également réaliste et accessible pour notre pays : le « plan IA » que nous proposons représenterait 0,3 % des dépenses publiques totales. Le coût de l'inaction serait, à l'inverse, très élevé. Nous renoncerions à des gains économiques et sociaux importants, et risquerions un déclassement historique. Il s'agit donc de choisir les dépenses qui permettront à la France d'assurer la maîtrise de son avenir.

A. AFFIRMONS LE PRINCIPE DE RESPONSABILITÉ : L'INNOVATION AU SERVICE D'UN PROJET DE SOCIÉTÉ

Partout, les sociétés sont mises au défi par la diffusion des technologies numériques. Les réseaux sociaux ébranlent les systèmes politiques. La concentration technologique polarise la répartition des richesses. Les algorithmes contribuent aux inégalités de travail et d'emploi. La massification des usages s'accompagne d'un impact environnemental croissant. Le pouvoir de certaines entreprises limite la capacité d'action des États souverains.

L'intelligence artificielle prolonge et approfondit ce mouvement. À rebours des scénarios d'épouvante avancés par certains, les actuels systèmes d'IA ne conduiront pas à la fin de l'humanité. Ils sont cependant loin d'être infaillibles et s'accompagnent d'effets indésirables : reproduction de stéréotypes, divulgation d'informations confidentielles, violation des droits de propriété intellectuelle, etc. Ils ouvrent de nouvelles possibilités d'actes malveillants, en particulier en matière de cyberattaques ou de désinformation. Ils sont source de risques systémiques, notamment liés au potentiel de concentration technologique entre les mains d'un petit nombre de pays, d'entreprises ou de personnes.

Face à ces défis d'ampleur, nous ne pouvons pas reproduire les erreurs du passé. Au cours de ces deux dernières décennies, la France et l'Europe ont réagi trop tard et trop peu, avec un faible engagement dans l'innovation technologique et une réglementation tardive. Aujourd'hui, il nous revient de tirer parti de l'IA en l'installant à sa juste place : celle d'un moyen technologique au service d'une ambition d'humanité, d'égalité, de solidarité, de justice, de prospérité, de liberté.

Ces enjeux concernant le monde entier, plusieurs modèles d'organisation internationale ont été envisagés ces dernières années. La France a fait partie des pays pionniers avec la co-fondation, en 2020, du programme mondial pour l'intelligence artificielle (PMIA). Les forums de discussion internationale sur l'IA sont foisonnants, il y en a une cinquantaine au moins à ce jour. Pour aller plus loin et ancrer ces initiatives dans l'action opérationnelle, des parallèles ont été faits avec les enjeux mondiaux du climat ou de l'énergie. Notre Commission constate que ces parallèles sont insuffisants : l'IA ne peut se conformer à un modèle antérieur. Nous considérons également que la communauté internationale doit profiter de la fenêtre d'opportunité qui se présente en 2024 pour faire converger le foisonnement d'initiatives.

Afin d'assurer la maîtrise de l'IA, nous recommandons de fonder une gouvernance mondiale avec une coalition de pays partageant les mêmes objectifs. Notre Commission envisage trois avancées majeures. Premièrement, une coalition de pays érigerait l'**Organisation mondiale de l'IA** (*World AI Organization*). Cette organisation internationale partagerait des constats scientifiques sur le fonctionnement et les effets de l'IA et définirait des normes contraignantes sur les systèmes d'IA et les modalités de leur audit. Elle serait gouvernée démocratiquement, en réunissant les États, la société civile (chercheurs, citoyens, syndicats) et les entreprises.

Deuxièmement, la France pourrait soutenir la création d'un **Fonds international pour l'IA au service de l'intérêt général** (*International Fund for Public Interest AI*), doté d'un budget annuel de 500 M€, avec pour objectif de financer des initiatives d'intérêt général : des services d'IA libres et gratuits (*open*), des projets de recherche indépendante, des innovations (dans l'environnement, la science, la santé…). Troisièmement, la France pourrait promouvoir le **mécanisme de solidarité « 1 % IA »**, par lequel tous les acteurs internationaux investissant dans la puissance de calcul s'engageraient à en allouer 1 % aux pays en voie de développement.

Cette gouvernance mondiale devra se décliner au niveau national. La France a l'occasion de se placer parmi les pionniers de l'évaluation des systèmes d'IA, notamment en structurant son réseau d'évaluation et de surveillance. Il convient également de mener, en continu et avec ambition, des travaux prospectifs et d'anticipation des évolutions de l'IA, afin d'anticiper ses effets sur la société et préparer les transformations nécessaires.

B. VISONS UN OBJECTIF D'HUMANISME DANS LE DÉPLOIEMENT DE L'IA

La révolution technologique de l'intelligence artificielle devra accorder davantage de pouvoir aux citoyens et aux travailleurs. Ne pas s'y employer ouvre le risque d'un refus massif de l'IA. Par le passé, d'autres innovations technologiques ont affecté la cohésion sociale. Surtout, l'innovation n'a de sens que si elle est au service du libre épanouissement de notre humanité. Autrement dit, le déploiement de l'IA doit viser un objectif d'humanisme. Pour y parvenir, notre Commission a identifié trois piliers principaux : la formation, le dialogue social et le service public.

Nous recommandons de lancer immédiatement un plan de sensibilisation et de formation de la nation. Pour y parvenir, nous devons d'abord créer les conditions d'une **appropriation collective de l'IA** et de ses enjeux. Cela suppose d'animer en continu des débats publics dans notre société, de susciter la création de lieux d'expérimentation et d'appropriation de la technologie (les « cafés IA »), de mettre à disposition un outil numérique d'information ou encore de lancer un concours de cas d'usages positifs de l'IA.

Nous devons également investir dans la **formation de tous et à tout âge** : des jeunes dans le temps scolaire et périscolaire, des étudiants spécialisés ou non, des salariés, des indépendants et des agents publics, des retraités. Cela implique de préparer les métiers de demain, notamment en structurant une offre de formation d'enseignement supérieur hybride, comme « IA + biologie » et « droit + IA », ou en créant des chaires sur l'IA dans les écoles de création. Cela implique aussi de permettre l'usage de l'IA dans les métiers d'aujourd'hui, par exemple en prévoyant un parcours de sensibilisation à l'IA pour l'ensemble des agents publics.

Le renouveau du dialogue social devrait constituer la pierre angulaire du recours à l'IA. À l'échelle nationale comme à l'échelle de l'entreprise, il est nécessaire de construire les usages de l'IA selon une démarche partenariale. Dans le même temps, des investissements devront être consacrés à l'analyse des impacts de l'IA sur la quantité et la qualité de l'emploi. L'IA elle-même peut être mise au service du dialogue social, avec la création et le déploiement d'outils spécialisés.

Enfin, les systèmes d'IA devront être mis au profit de la qualité du service public. L'intelligence artificielle peut améliorer le service public, en contribuant à personnaliser l'éducation, à accorder plus de temps aux patients, à mieux accompagner et anticiper les transitions professionnelles, à réduire la bureaucratie. Nous obtiendrons ces gains à condition de faire la mue de nos institutions. De l'évolution des infrastructures numériques à la conduite de projets d'IA, la mobilisation des administrations publiques sur les enjeux tenant à l'IA doit être accélérée, amplifiée, généralisée et déclinée par service public.

C. UN POTENTIEL ÉCONOMIQUE ET SOCIAL MAJEUR

Si nous nous mobilisons pour la déployer et la maîtriser, l'IA devrait augmenter la prospérité collective et peut contribuer à l'amélioration de la qualité du travail et à la réduction des inégalités. Selon notre analyse, la croissance économique annuelle de la France pourrait doubler grâce à l'automatisation de certaines tâches. Au bout de dix ans, la hausse de PIB serait comprise entre 250 et 420 Md€, soit du même ordre de grandeur que l'activité actuelle de l'industrie dans son ensemble. Cette hausse serait cependant temporaire : une fois l'IA adoptée par l'ensemble du tissu économique, il n'y aurait plus de gains de productivité à attendre.

Notre prospérité devrait être renforcée par une deuxième caractéristique majeure de l'IA : elle semble accélérer l'innovation. Cet effet demeure incertain, mais il est étayé par de nombreux usages récents de l'IA : trouver de nouvelles protéines, identifier de nouveaux matériaux, etc. Si cet effet se confirme, il s'agirait là d'une caractéristique remarquable de l'IA : elle pourrait induire une augmentation permanente du taux de croissance de l'économie. Autrement dit, en plus d'un effet *temporaire* lié à l'automatisation, l'IA pourrait produire un effet *à plus long terme* lié à l'émergence de nouvelles innovations, de nouveaux produits, de nouvelles formes d'organisation, etc.

Dans les prochaines années, les systèmes d'IA conduiront à la transformation de nombreux emplois. La plupart des travailleurs bénéficieront de l'automatisation de tâches parfois ingrates. La grande majorité des métiers évolueront, avec des tâches en plus, et des tâches en moins. Oui, il faut se préparer à ce que l'automatisation permise par l'IA supprime certains emplois et accélère l'obsolescence de certaines compétences. Au niveau sectoriel ou à titre individuel, cette évolution représentera un défi de formation et de reconversion. Cependant, au niveau national et malgré les incertitudes, notre Commission estime que les effets de l'IA seront globalement favorables à l'emploi : l'IA pourra générer des emplois dans de nouveaux métiers, en partie inconnus à ce jour, ainsi que dans des métiers existants.

La seule existence de la technologie ne garantit en aucun cas ces gains économiques et sociaux. L'histoire récente le prouve. Alors que les technologies numériques ont contribué à la croissance américaine, la France n'en a que peu bénéficié. L'activité française a crû nettement moins qu'outre-Atlantique et aucun véritable acteur mondial du numérique n'a émergé : entre 2001 et 2022, la richesse par habitant a progressé de 29 % aux États-Unis, contre seulement 14 % en France.

Il est donc essentiel de mettre en place un ensemble de politiques publiques adaptées pour maximiser les gains : politique d'innovation, politique industrielle, politique concurrentielle… L'accompagnement des reconversions professionnelles et la formation continue seront également déterminants. Il est d'autant plus important de relever le défi de l'accompagnement des évolutions de carrières individuelles que la rapidité de la diffusion de l'IA rendra les transitions difficiles.

Seuls les États qui se donneront les moyens de maîtriser l'IA en obtiendront les principaux bénéfices. Or, à ce jour, la France et l'Europe sont en retard.

D. L'AMPLEUR DU DÉFI ÉCONOMIQUE : LA PRÉOCCUPANTE FAIBLESSE DE LA FRANCE ET DE L'EUROPE

L'économie du numérique est deux à trois fois plus faible en Europe qu'aux États-Unis, et l'IA suit pour l'instant une trajectoire comparable. Sur les 100 entreprises de technologies à la plus grande capitalisation fin 2023, 10 sont européennes. Le problème n'est pas seulement que l'Europe ne produit pas de géant du numérique, mais qu'elle ne produit pas non plus d'entreprise au 2e ou 3e rang : ni Adobe, ni Uber, ni AirBnB, ni Shopify, ni Stripe ne sont européennes, alors que leurs plus grands marchés ou leurs fondateurs sont européens. Dans le secteur des services numériques (logiciels, traitement de données…), l'activité est ainsi 2,5 fois plus élevée aux États-Unis que dans l'Union européenne et le Royaume-Uni pris ensemble. Nous retrouvons un ordre de grandeur comparable dans l'IA. Le nombre d'entreprises spécialisées qui ont été financées sur la période 2013-2022 est effectivement 2,5 fois plus élevé aux États-Unis qu'en Europe.

Si cette supériorité se poursuit ou se renforce, la France et l'Europe encourent le risque d'un rapide déclassement économique. Ce risque est de deux ordres : être largement dépourvu d'entreprises *spécialisées* dans l'IA et voir les entreprises *existantes* perdre en compétitivité. À l'image de la précédente vague d'innovations numériques, nous pourrions donc non seulement manquer l'économie de l'IA, ce qui conduirait à une captation croissante par d'autres de notre valeur économique, mais aussi voir l'affaiblissement des autres secteurs d'activité.

Aucune entreprise n'est à l'abri. Depuis dix ans, l'intégration de l'IA dans les entreprises est dans l'ensemble plus lente et moins profonde en France qu'aux États-Unis, au Royaume-Uni ou dans les pays scandinaves. En l'absence d'adoption rapide et structurelle, chacune des entreprises françaises sera confrontée à l'érosion de ses parts de marché, de ses marges et de sa valeur, mais aussi au risque de son éviction par un nouvel acteur (*disruption*). Cette perspective se renforce à mesure que les systèmes d'IA sont plus accessibles et plus performants.

Au-delà, le retard en matière d'intelligence artificielle porte atteinte à notre souveraineté. Une faible maîtrise de la technologie implique effectivement un lien de dépendance à sens unique vis-à-vis d'autres pays. Dans le domaine privatisé et si évolutif qu'est l'IA, la puissance publique apparaît largement dépassée, limitant ainsi notre capacité collective à faire des choix alignés avec nos valeurs et nos intérêts.

Cette faiblesse de l'innovation française s'explique par de multiples freins. Le frein de la méconnaissance des enjeux technologiques sous-jacents à l'IA et de ses effets potentiels sur la société. Le frein de notre aversion collective au risque, qui nous amène à éviter les dans des technologies et des modèles d'affaires non prouvés. Le frein de la bureaucratie qui entrave en particulier la recherche publique.

Le retard n'est pas inéluctable et il n'est pas trop tard pour agir. Notre continent dispose d'atouts qu'il ne faut pas négliger. Des entreprises européennes sont positionnées sur l'ensemble de la chaîne de valeur de l'IA. L'enseignement supérieur français forme des ingénieurs et des chercheurs d'excellence en IA. Par ailleurs, les innovations technologiques au cœur de l'IA générative sont récentes et la chaîne de valeur économique est très loin d'être arrivée à maturité. L'économie de l'IA est encore en devenir et la plupart des modèles d'affaires restent à inventer. Depuis le rapport de Cédric Villani (2018), l'État s'est également mis en mouvement dans le cadre des investissements d'avenir et de France 2030, en développant des formations spécialisées dans l'IA, accentuant l'effort de recherche publique et soutenant l'innovation privée.

E. INNOVONS POUR ASSURER LA MAÎTRISE DE NOTRE AVENIR

Nous n'aurons pas la maîtrise de notre avenir par le seul déploiement de l'IA des autres. Les retards français et européen, s'ils ne sont pas comblés, renforceront notre dépendance vis-à-vis d'autres pays, affecteront notre cohésion sociale et affaibliront notre économie. De même, il est illusoire de croire que nous pouvons emprunter un chemin autarcique. Il nous revient donc de tirer parti, dès à présent, du potentiel des systèmes d'IA, d'où qu'ils viennent, tout en créant les conditions d'une offre européenne d'IA. Trois axes d'interventions conditionnent l'émergence d'un écosystème d'IA.

Premièrement, les financements actuels de l'écosystème de l'IA sont insuffisants pour faire émerger des acteurs de rang mondial : nous recommandons de réorienter une partie de l'épargne vers l'innovation. Les montants investis dans l'IA aux États-Unis sont aujourd'hui 20 fois supérieurs à ceux investis en France. À richesse comparable, nous investissons environ trois ou quatre fois moins que les Américains et l'écart risque d'augmenter. À moyen terme, un accroissement structurel de l'allocation de l'épargne vers l'innovation est indispensable. Des actions volontaristes doivent être rapidement prises en ce sens, par exemple en matière de fiscalité de l'assurance vie, afin de disposer d'ici quelques années d'une capacité de financement significativement accrue.

À court terme, nous proposons la création d'un fonds d'investissement « France & IA ». **Le fonds mobilisera 10 Md€ de** capital-investissement d'entreprise et de soutien public, selon plusieurs modalités d'intervention, pour faire émerger l'écosystème d'IA et accélérer la transformation du tissu économique par l'IA. Aux côtés des moyens financiers, le fonds **s'accompagnera d'une mise en commun de données d'activité pour conduire certains projets numériques. Une telle ampleur des moyens et le tandem de financement et de données sont inédits en France. Face au risque de déclassement économique, l'audace participera de l'émergence de solutions innovantes performantes et de l'accélération de la modernisation des entreprises françaises.**

Deuxièmement, nous ne tirerons pas les bénéfices de l'IA générative sans accéder à des données fiables de qualité : nous recommandons donc de repenser la gouvernance de la donnée. D'abord, il est essentiel de **faciliter l'accès aux données** à caractère personnel pour permettre leur utilisation dans des innovations thérapeutiques, notamment en supprimant certaines procédures d'autorisation préalable d'accès aux données de santé et en réduisant les délais de réponse de la Commission nationale de l'informatique et des libertés (Cnil). Cela suppose de réformer le mandat de la Cnil, pour y ajouter un objectif d'innovation, de réviser la composition de son collège et de renforcer ses moyens.

Ensuite, nous devons **veiller à l'application du principe de transparence** des données d'entraînement des grands modèles d'IA. Prévu par le règlement européen relatif à l'IA, ce principe doit permettre le respect du droit de la propriété littéraire et artistique. Il doit pouvoir être mis en œuvre de la façon la plus simple possible, pour les développeurs de modèles d'IA comme pour les ayants droit, en particulier par l'élaboration de standards pour la publication des informations sur les modèles d'IA et la mise en œuvre du droit de retrait.

Enfin, dans une logique plus prospective, il nous revient de **concevoir une nouvelle gestion, collective, des données**. Alors que la protection des données est aujourd'hui centrée sur l'individu avec le règlement européen de protection des données personnelles, l'accès aux données des modèles d'IA et les bénéfices qui en découlent sont principalement collectifs.

Cette dichotomie a jusqu'ici favorisé les géants du numérique, seuls à disposer de centaines de millions d'utilisateurs qui génèrent chaque jour des flux de données pour entraîner leurs modèles. Il convient d'explorer de nouveaux modèles de gouvernance commune des données, sans bien sûr affaiblir la protection des individus.

Troisièmement, la puissance de calcul est l'autre ingrédient incontournable de l'IA générative : nous recommandons donc de faire de la France un pôle majeur du domaine. C'est une condition *sine qua non* de l'autonomie stratégique et les capacités publiques de supercalculateurs, qu'il convient de soutenir, ne suffiront pas. Du côté de l'offre, il faut sécuriser sans délai l'approvisionnement de l'écosystème français par le biais d'une commande européenne de puissance de calcul privée. Parallèlement, un appel à projets d'implantation de centres de calcul pourrait être lancé sur le territoire européen, assorti à la fois d'une garantie publique d'utilisation de la puissance de calcul et d'un accompagnement à l'implantation et au raccordement électrique. Du côté de la demande, un crédit d'impôt IA soutiendrait les projets de recherche et de développement dans la location de la puissance de calcul, sous la condition d'utiliser un centre de calcul établi sur le territoire. Enfin, la politique industrielle pourra être orientée vers l'émergence d'une filière électronique adaptée à l'IA.

Il ne s'agit pas de courir derrière les avancées technologiques, mais de créer nos avantages comparatifs. Le ciblage et la concentration des moyens seront donc primordiaux pour fonder notre supériorité sur certains segments de la chaîne de valeur, et ainsi être en mesure de parler d'égal à égal avec nos concurrents et nos partenaires. Le ciblage doit aller de pair avec l'émergence progressive d'écosystèmes d'innovation dans le domaine de l'IA en France et en Europe. La voie de leur différenciation pourra notamment porter sur la dimension environnementale, en visant de nouvelles générations d'IA, de l'architecture matérielle au choix des modèles, qui consommeront moins d'énergie.

F. MISONS SUR L'OUVERTURE DES SYSTÈMES D'INTELLIGENCE ARTIFICIELLE

Le développement de l'IA fait apparaître un risque majeur de domination du marché par quelques acteurs. Une seule entreprise — américaine — détient actuellement 80 % des parts de marché mondial de conception des processeurs graphiques. Trois entreprises — américaines — se partagent 80 % de l'augmentation des dépenses françaises en infrastructures et applications de services de *cloud*. Ces trois mêmes entreprises associent à leurs services de *cloud* des outils — américains — d'IA générative.

On ne peut pas se satisfaire que seules quelques entreprises, *a fortiori* extra-européennes, soient les seules à maîtriser — sous toutes les acceptions du terme — la révolution technologique de l'IA. Il convient donc de veiller à l'émergence d'une diversité d'acteurs économiques, notamment français et européens, pour des raisons de souveraineté et économiques. Limiter les positions concurrentielles dominantes favorise la croissance et une juste répartition des gains économiques.

La politique européenne de concurrence devra donc être pleinement mobilisée pour prévenir l'émergence de positions dominantes. À court terme, il importe de recourir à la palette d'actions prévue par le règlement européen sur les marchés numériques. Ce règlement pourrait par ailleurs être complété pour tenir compte des spécificités de la chaîne de valeur de l'IA. À moyen terme, il convient d'envisager un changement de doctrine de la politique de concurrence, en passant d'un système *statique* (quelles parts de marché détient *aujourd'hui* cette entreprise ?) à une vision *dynamique* (quelles parts de marché pourrait *demain* détenir cette entreprise et quelles entreprises pourraient *demain* entrer sur ce marché ?).

Au-delà, notre Commission recommande de soutenir un écosystème ouvert de développeurs d'IA, présentant des bénéfices de transparence, de pluralisme et de concurrence. Nous considérons qu'un tel écosystème constitue un puissant levier d'innovation et peut participer de la sécurité des systèmes d'IA et du développement d'usages bienveillants, y compris de contre-mesures vis-à-vis des usages malveillants. Il contribue aussi à la confiance des citoyens et à la réduction de certains impacts négatifs de l'IA sur les individus. Il faut donc apporter à l'écosystème de la sécurité juridique et des données de qualité, mais aussi développer les capacités d'inspection et d'évaluation des modèles.

Enfin, pour concevoir l'avenir de l'IA, il est indispensable de libérer les chercheurs des contraintes administratives : nous recommandons d'assumer le principe d'une « exception IA » dans la recherche publique. Sous la forme d'une expérimentation, ce principe vise un objectif de « zéro entrave pour les chercheurs », notamment par un engagement sur les délais de réponse aux sollicitations et la mise en place d'un indicateur de simplicité administrative. L'exception IA doit également permettre de revaloriser les rémunérations des chercheurs et des enseignants-chercheurs et de faciliter les temps partiels avec des entreprises ou d'autres acteurs socio-économiques de l'IA. Ces actions pourront s'accompagner d'un doublement des moyens de la recherche publique spécialisée en IA, amplifiant ainsi les investissements menés dans le cadre des instituts interdisciplinaires et des prochains *IA cluster*.

Des Lumières à nos jours, l'ouverture est au cœur de notre continent européen et de nos valeurs. Inscrivons-nous dans cette tradition fondatrice.

G. QUEL AVENIR À L'ÈRE DE L'INTELLIGENCE ARTIFICIELLE ?

Avec l'IA générative, une étape importante dans l'histoire de l'innovation a été franchie. Cette étape est loin d'être la dernière. Dans les mois et années à venir, nous devrions connaître de nouvelles avancées rapides et de grande ampleur. Les modèles seront progressivement capables d'être factuels, de mener des raisonnements, de comprendre le monde physique autour de nous. L'IA accompagnera les personnes en continu et dans toutes leurs tâches, peut-être sous la forme d'assistants personnalisés. Des produits et des gestes seront inventés pour nous permettre de tirer tout le parti de ces nouvelles IA. La robotique fera également des progrès majeurs.

Les transformations sociétales que susciteront ces innovations dépendront de notre ambition et de notre engagement. L'IA peut être mise au profit de la réduction des inégalités sociales, de la prospérité collective et de la qualité du travail. Ces bénéfices ne seront pas spontanément obtenus. En l'absence de projet politique et d'engagement collectif, l'IA peut, à l'inverse, affaiblir notre démocratie, altérer notre souveraineté et concentrer les richesses. Veillons à ne pas sous-estimer les potentiels effets de l'IA à moyen terme, et à ne pas les surestimer à court terme.

Donnons-nous donc, collectivement et sans attendre, les moyens de tirer parti de l'IA. Ses effets seront d'autant plus bénéfiques que la France et l'Europe maîtriseront la technologie et sa chaîne de valeur. Cette maîtrise est incontournable. Notre Commission recommande donc de combler les retards français et européen et de lancer une nouvelle stratégie relative à l'IA d'ici la fin du premier semestre 2024. Les mesures devront faire l'objet d'une évaluation continue et d'une révision annuelle, voire semestrielle si le rythme des innovations demeure très soutenu.

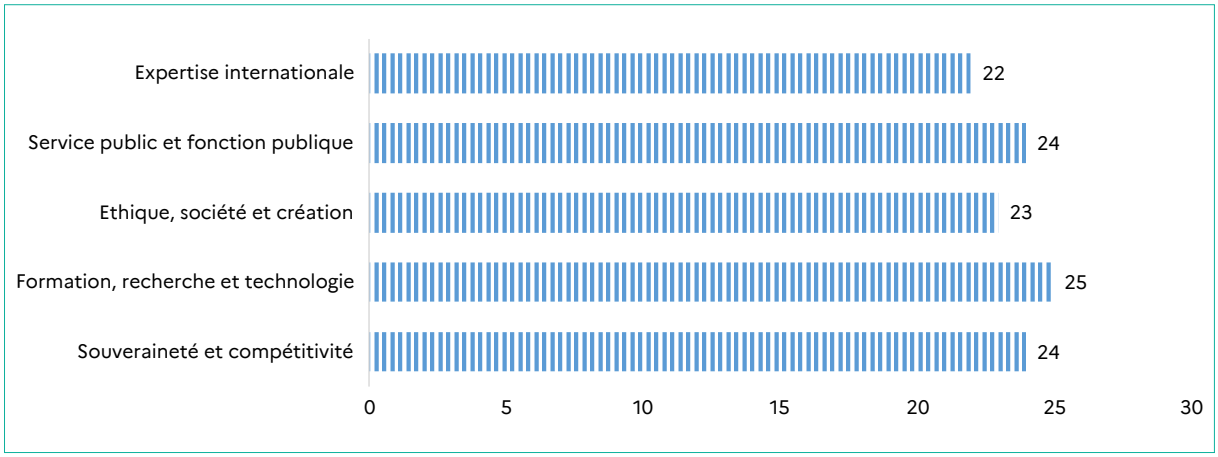
C'est une course de fond qu'il convient d'engager. La stratégie recommandée par notre Commission n'en constitue que les premiers kilomètres. Aller au-delà nécessitera de la constance dans l'intervention publique, mais aussi des actions visant à renforcer la plasticité de nos organisations, publiques et privées. Une démarche collégiale d'anticipation apparaît également incontournable, afin de préparer notre pays aux effets de la révolution technologique.

MÉTHODE

COMPOSITION

Présidents	Philippe Aghion	Anne Bouverot	
Rapporteurs généraux	Arno Amabile	Cyprien Canivenc	
Membres	Glles Babinet Joëlle Barral Alexandra Bensamoun Nozha Boujemaa	Bernard Charlès Luc Julia Yann Le Cun Arthur Mensch	Cédric O Isabelle Ryl Franca Salis-Madinier Martin Tisné Gaël Varoquaux
Rapporteurs	Marc Auberger Simon Bunel Philippe Chantepie Eloy Dorado	Emilie-Pauline Gallié Paul Jolie Arnaud Mazier Vincent Montreuil	Erwan Paitel Timothée Paris Christophe Ravier Ulrich Tan Louis Charles Viossat

CHAMPS D'EXPERTISE



MISSION

En septembre 2023, le Gouvernement a installé la Commission de l'intelligence artificielle pour « contribuer à faire de la France un pays à la pointe de la révolution de l'IA ». La Commission a ainsi été chargée de présenter des propositions opérationnelles, réalistes et ambitieuses soutenues par une vision long terme, globale et objectivée. C'est à ce mandat précis que nous entendons répondre avec ce rapport.

PRINCIPES ET MÉTHODES

Expertise

Les membres et les rapporteurs de la Commission ont été nommés, à titre personnel, en raison de leurs compétences en matière d'intelligence artificielle. Cette expertise est mise à la disposition du Gouvernement pour concourir à l'efficacité de l'action publique.

Collégialité

Les séances plénières de la Commission, qui ont réuni les quinze membres et les quinze rapporteurs selon un rythme hebdomadaire, incarnent la collégialité. La pluralité des expertises et la libre discussion contradictoire participent de l'objectivité des travaux, ainsi que de la mesure et de l'équilibre du plan d'action recommandé. Les membres de la Commission ont donné leur accord général au rapport et aux recommandations qui représentent un consensus majoritaire.

Concertation

Les débats et les conclusions de la Commission se sont nourris des auditions d'experts et de parties prenantes, aux profils, aux expériences et aux horizons géographiques les plus divers. Ils se sont enrichis de la consultation de citoyens, qui contribue à l'opportunité, au réalisme et au pragmatisme des recommandations.

Indépendance

La Commission a défini librement le programme de ses travaux et les modalités de son organisation, dans le cadre de la mission qui lui a été confiée le 19 septembre 2023. Elle a mené ses réflexions et a élaboré ses conclusions indépendamment des pouvoirs exécutif et législatif.

CHIFFRES CLÉS

600 auditions d'experts et de parties prenantes de l'IA

200 prises de contact spontanées avec la Commission

7 000 personnes consultées sur l'application Agora

23 séances de travail plénières de la Commission

25 recommandations à l'attention du Gouvernement

INTRODUCTION

L'intelligence artificielle (IA) est omniprésente dans le débat public. Ses applications défraient la chronique : synthèse de texte, génération de musique, traduction et interprétation, montage de photos, etc. Il en va de même pour certains outils à base d'IA dont l'usage s'est rapidement diffusé dans le monde. Certains y voient une source d'inquiétudes, d'autres d'espoirs. Des observateurs constatent ou prédisent une révolution sociétale, d'autres n'identifient pas de rupture dans leur quotidien personnel ou professionnel. En réalité, de quoi parle-t-on ?

L'INTELLIGENCE ARTIFICIELLE N'EST PAS UNE NOUVEAUTÉ : ELLE DATE DES ANNÉES 1950

L'IA est en réalité sur le devant de la scène depuis des décennies. Le système américain Mycin de diagnostic de maladies du sang et de prescription (années 1970), la construction du premier véhicule à conduite autonome Navlab (1986), la victoire de la machine *Deep Blue* sur le champion du monde d'échec Gary Kasparov (1997), l'assistant virtuel Siri intégré aux portables iPhone (2011), ou encore la défaite du champion mondial Ke Jie au jeu de go face à la machine *AlphaGo* (2017) ont tous été décrits, à un moment, comme de l'intelligence artificielle.

L'histoire de l'IA a effectivement plus de 70 ans. Dès 1950, le mathématicien et cryptologue britannique Alan Turing s'est intéressé à la capacité d'une machine à imiter une conversation. Durant plusieurs décennies, cette aptitude n'était pas suffisante pour tromper un humain, qui savait distinguer une conversation simulée d'une conversation réelle. Cette époque est désormais révolue, nous y reviendrons.

C'est quelques années plus tard, en 1956, qu'apparaît pour la première fois le terme d'intelligence artificielle. La recherche dans le domaine prend progressivement de l'ampleur et fait apparaître plusieurs approches technologiques. L'IA s'est d'abord développée sous la forme de règles déductives du type « si… alors ». Cette approche dite symbolique, fondée sur le raisonnement et des instructions, était largement majoritaire jusque dans les années 1990.

Sans que cette approche symbolique soit abandonnée, une approche statistique de l'IA a pris de l'ampleur à compter des années 1990 : l'apprentissage automatique, également appelé apprentissage machine. Contrairement à l'approche symbolique, l'humain ne détermine pas un ensemble de règles « si… alors ». Il veille à ce que l'ordinateur « apprenne » à identifier des relations statistiques entre les données. Il n'y a donc pas d'instruction explicite d'un humain : la machine est entraînée à reconnaître des liens à partir d'un ensemble de données dites d'entraînement. La machine applique ensuite ces liens à des données nouvelles pour effectuer une tâche.

Le succès de cette seconde approche repose sur deux ingrédients indispensables : les données et une puissance de calcul, soutenues par l'émergence du *cloud*. L'accessibilité de ces deux ingrédients a fortement crû ces 30 dernières années, sous le triple effet de la numérisation de notre société (produisant donc plus de données), de l'amélioration des matériaux semi-conducteurs (augmentant la puissance de calcul) et du progrès technique. Cette évolution a permis à l'apprentissage machine d'enregistrer des progrès aussi rapides qu'importants. Les techniques permettant aux machines « d'apprendre » automatiquement les règles à partir des données se sont diversifiées et affinées.

L'IA RECOUVRE UN ENSEMBLE D'OUTILS NUMÉRIQUES, DÉJÀ OMNIPRÉSENTS DANS NOTRE SOCIÉTÉ

Il n'existe pas de définition unique et universelle de l'IA, en particulier parce que ce terme recouvre de nombreuses technologies : il y a 30 ans, qu'une machine puisse distinguer à coup sûr un chat d'un chien semblait loin d'être atteinte. Il s'agit par ailleurs d'une notion théorique, car ce sont les *systèmes* d'intelligence artificielle que nous utilisons au quotidien. Que sont donc ces systèmes d'IA ?

Les systèmes d'IA sont en mesure d'établir des prévisions, de formuler des recommandations, ou de prendre des décisions. Ils répondent à un ensemble d'objectifs donné et ont une influence sur leur environnement.

Les systèmes d'IA ont de très nombreuses applications dans notre quotidien, dans notre économie et dans nos services publics. Citons par exemple la reconnaissance vocale des téléphones portables, la robotique industrielle, les véhicules à conduite automatisée, la détection de pathologies en imagerie médicale, les assistants commerciaux virtuels, la reconnaissance faciale des ordinateurs, la publicité ciblée sur internet ou encore l'identification d'anomalies financières pour lutter contre la fraude fiscale.

L'IA GÉNÉRATIVE CONSTITUE UN TOURNANT MAJEUR DE CETTE HISTOIRE DE L'INNOVATION

Ces nombreuses applications professionnelles ou personnelles sont amplifiées par les systèmes d'IA dite générative. L'IA est qualifiée de générative, car elle permet de générer de nouveaux contenus sous la forme de texte, d'image, de son, de vidéo ou de code. Cette capacité de production constitue un tournant majeur de l'IA, et ce à plusieurs titres.

Premièrement, l'utilisation des modèles d'IA générative peut être d'une grande simplicité. Il est effectivement possible d'utiliser des interfaces de dialogue, par lesquelles un humain s'exprime par écrit ou par oral pour commander la génération d'un contenu. Cette possibilité crée l'impression d'une conversation ou d'un dialogue avec la machine.

Deuxièmement, la génération de contenu est rapide. Il suffit de quelques secondes aux modèles les plus avancés pour produire une musique ou un long texte, tandis qu'une réalisation semblable par l'homme nécessiterait plusieurs jours ou semaines.

Troisièmement, le contenu généré est réaliste. Il apparaît crédible aux yeux et aux oreilles d'un humain, car il présente des caractéristiques d'humanité : clarté des propos, enchaînement logique des mots, cohérence des images, présence d'intonation, etc.

Quatrièmement, les modèles présentent d'importantes aptitudes. Une série d'expériences a montré en 2023 que des modèles d'IA générative semblent être en mesure de réaliser des tâches humaines complexes. Par exemple, une expérience a conclu que des algorithmes présentent de meilleurs résultats que 90 % des candidats humains à certains examens du barreau aux États-Unis[1]. D'autres expériences, en médecine, ont relevé une précision supérieure de diagnostics effectués par des algorithmes par rapport à ceux réalisés par des médecins[2]. Ces performances sont prometteuses, même si elles sont parfois obtenues dans des situations relativement éloignées des conditions de la vie réelle. Nous devons continuer à évaluer ces performances de façon rigoureuse[3] et éviter d'y projeter de l'intelligence humaine[4].

Réalisme, simplicité, rapidité, aptitudes. Ces caractéristiques de l'IA générative permettent l'automatisation d'un certain nombre de tâches qui étaient difficilement automatisables auparavant. Par exemple, elles facilitent la personnalisation des offres commerciales, simplifient l'analyse de données financières, accélèrent la recherche scientifique, etc.

Ces mêmes caractéristiques laissent penser que l'IA pourrait prendre la suite des ordinateurs personnels, des réseaux sociaux et des smartphones comme « la » plateforme numérique dominante, la couche technologique sur laquelle tous les autres nouveaux services sont construits[5]. A chaque changement de plateforme, les cartes sont rebattues et le pouvoir est redistribué aux entreprises qui contrôlent la nouvelle plateforme. IBM, entreprise toute puissante du temps des *mainframes*[6], n'a pas disparu, mais elle n'a plus la même centralité depuis l'essor des ordinateurs personnels. Cette perspective ouvre toutefois plus de questions qu'elle n'apporte de réponses : si l'IA est la prochaine plateforme, qui la contrôle, les entreprises qui font les modèles ou celles qui font des produits qui intègrent de l'IA ? Si des agents conversationnels deviennent la nouvelle interface centrale avec le monde numérique, comment en définir ensemble les conditions et les comportements ?

De manière générale, les incertitudes sont nombreuses. Qui maîtrise et déploie les systèmes d'IA générative ? Quelles sont les responsabilités de ces acteurs ? Quelles utilisations en seront faites ? Quels seront ses effets sur l'économie, le travail et l'emploi ? Quelles sont les incidences de l'apparence humaine du contenu généré sur notre rapport à la vérité et à l'information ? Comment vont évoluer les technologies ?

Pour répondre à ces interrogations, une description plus fine de la technologie et de la chaîne de valeur économique est nécessaire.

1. Katz D. M., M. J. Bommarito, S. Gao, et P. D. Arredondo (2023), « GPT-4 Passes the Bar Exam », *SSRN eJournal*.
2. Caruccio L., S. Cirillo, G. Polese, G. Solimando, S. Sundaramurthy, et G. Tortora (2024), « Can ChatGPT provide intelligent diagnoses? A comparative study between predictive models and ChatGPT to define a new medical diagnostic bot », *Expert Systems with Applications, Volume 235*.
3. Notamment pour vérifier qu'elles ne viennent pas d'une simple inclusion de ces examens dans les données d'entraînement, ou qu'elles ne se réalisent que dans un environnement très contrôlé.
4. Pour un humain, réussir l'examen du barreau est (raisonnablement) corrélé avec une compétence juridique plus générale. Nous n'avons pas de raison de penser que c'est également le cas pour un modèle de language.
5. OpenAI ne s'y est pas trompé et essaye de devenir aussi rapidement que possible une plateforme indispensable. L'entreprise a rapidement ouvert des outils permettant à des tiers de créer versions personnalisées de ChatGPT, à acheter sur un « store ».
6. Grands ordinateurs répandus du temps où les ordinateurs étaient si gros et chers que seuls quelques organisations en avaient.

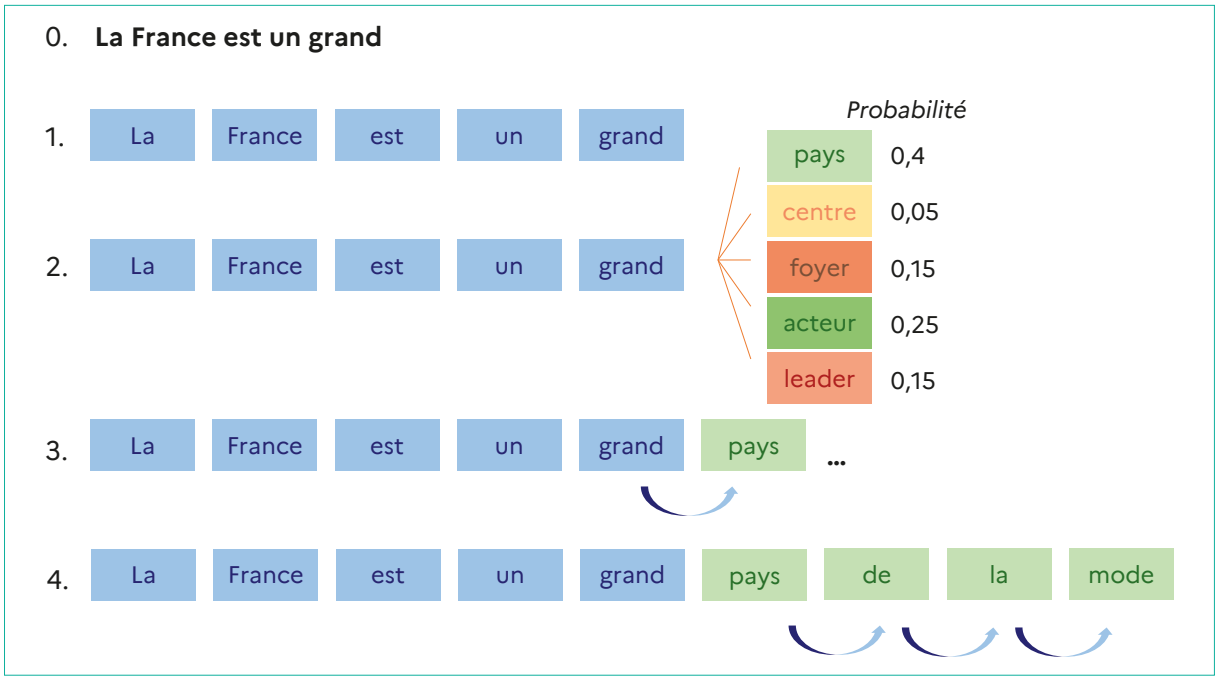
COMMENT FONCTIONNE L'INTELLIGENCE ARTIFICIELLE GÉNÉRATIVE ?

Les systèmes d'IA générative s'appuient, nous l'avons vu, sur les techniques d'apprentissage automatique. Ils sont donc entraînés sur d'importants volumes de données et « apprennent » les relations entre ces données d'entraînement. Ces données peuvent prendre la forme de texte, d'image, de son, de vidéo, de tableaux d'informations, ces catégories pouvant se cumuler.

Les systèmes peuvent être entraînés sur de très grandes quantités de données pour former des modèles de fondation (plus récemment appelés modèles à usage général), qui peuvent être adaptés à beaucoup de tâches différentes. Parmi eux, on trouve notamment les grands modèles de langue qui ont été entraînés sur de vastes corpus de textes.

Une fois entraîné, le modèle peut être sollicité par un utilisateur par le biais de requêtes. Le modèle d'IA générative répond alors à la requête en produisant de nouvelles données (de texte, d'image, de son). Le contenu généré présente une certaine similarité avec les données d'apprentissage, sans être identique.

Prenons le cas particulier de la génération de texte. Lorsque l'on adresse la requête « Complète la phrase suivante : La France est un grand (...) », le modèle de langue commence par décomposer cette requête en une série d'unités élémentaires de texte appelées *token*. Un *token* correspond à une série de quelques lettres qui ne forment pas toujours des mots complets. Pour simplifier, associons un *token* à chaque mot : « La », « France », « est », « un », « grand ».



Après plusieurs étapes techniques, le modèle analyse cette succession de *tokens* à la lumière des données d'entraînement. Il identifie ainsi un ensemble de possibilités pour poursuivre le texte : « pays », « centre », « foyer », « acteur », « leader ». À chaque possibilité de *token* successeur est associée une probabilité. La réponse est générée, *token* par *token*, en fonction de la probabilité de chaque *token*. Le texte généré est alors adressé à l'utilisateur : « pays », « de », « la » …

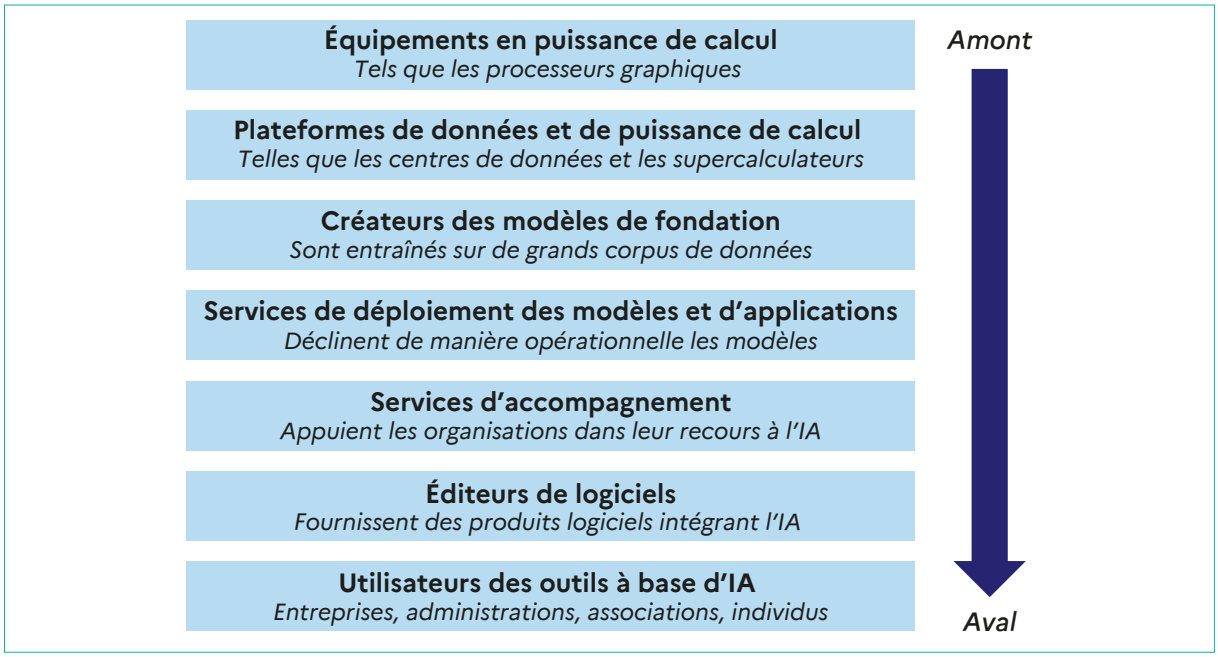
Ce fonctionnement, ici décrit de manière très simplifiée, présente une caractéristique essentielle : les modèles d'IA générative n'appréhendent pas la signification des mots, des images ou du son. La signification de la réponse générée ne provient donc pas de la machine, mais bien des humains. Les humains projettent leur vision du monde sur les résultats générés par la machine. Ils accordent d'autant plus de signification au texte et à la voix générés automatiquement que ceux-ci ressemblent beaucoup à un texte écrit ou énoncé par des êtres humains.

Cette caractéristique suppose donc une certaine vigilance quant à l'utilisation des modèles d'IA générative actuels. En particulier, on ne peut pas les utiliser – à ce jour – comme des sources fiables d'assertions vraies, à l'instar des encyclopédies. Il arrive effectivement que les modèles d'IA génèrent des réponses erronées. Les réponses erronées, souvent appelées hallucinations (également *confabulations*), constituent l'un des axes d'amélioration des systèmes d'IA.

COMMENT SE DÉCOMPOSE LA CHAÎNE DE VALEUR ÉCONOMIQUE DE L'IA GÉNÉRATIVE ?

Il n'est pas besoin d'attendre ces progrès technologiques pour que les applications des systèmes d'IA dans nos sociétés soient innombrables. Ces applications présentent une importante valeur économique, partagée au sein d'une longue chaîne de valeur.

Au milieu de cette chaîne de valeur, se trouvent bien sûr les créateurs de modèles de fondation présentés précédemment. Ces modèles ne pouvant être ni entraînés ni utilisés sans deux ingrédients indispensables — les données et l'infrastructure de calcul —, la chaîne de valeur s'étend en amont aux entreprises qui fournissent des plateformes de données et de puissance de calcul. Ces entreprises se fournissent elles-mêmes en équipements spécifiques aux modèles d'IA générative.



La chaîne de valeur économique se déploie également en aval des modèles de fondation. Ces derniers peuvent être directement utilisés par les consommateurs finaux (entreprises, administrations, associations, individus), qui retirent alors un avantage économique, tels que des gains de productivité ou de qualité. De nombreuses entreprises déploient ainsi cette technologie pour optimiser leurs chaînes de production et leurs services, dans des domaines aussi variés que les médias, la finance, le droit, l'informatique, l'automobile, ou encore l'industrie pharmaceutique…

La plupart du temps, cependant, plusieurs acteurs économiques contribuent à la chaîne de valeur entre les utilisateurs finaux et les créateurs de modèles de fondation. En effet, les modèles de fondation peuvent généralement être optimisés en vue d'applications précises et de tâches particulières. Par ailleurs, le déploiement des outils à base d'IA appelle généralement une transformation des organisations (adaptation des systèmes d'information, évolution de procédures, réaffectation des ressources humaines…), qui nécessite un accompagnement. Les modèles d'IA seront aussi intégrés directement aux produits logiciels choisis par les utilisateurs finaux, en particulier dans la suite bureautique (rédiger des courriels, des comptes rendus de réunions, etc.).

Au total, la chaîne de valeur de l'IA générative se compose d'entreprises dont l'IA est au cœur du modèle d'affaires, mais aussi d'entreprises qui utilisent l'IA et l'intègrent dans un modèle d'affaires préexistant ou adapté. Il y a donc une économie *de* l'IA et une économie *par* l'IA. La valeur économique globale de l'IA générative est considérable. D'après les travaux de notre Commission, le déploiement de l'IA pourrait doubler la croissance économique de notre pays.

Compte tenu de ces avantages économiques, mais aussi des potentiels bénéfices sociaux et de souveraineté de l'IA, il faut s'attendre à une très forte compétition internationale. Les entreprises — et les pays — qui seront en tête en tireront les principaux bénéfices.

EN FRANCE ET EN EUROPE, NOUS ACCUSONS UN NET RETARD…

À ce jour, les acteurs américains dominent très largement l'amont de la chaîne de valeur de l'IA générative. Prenons quelques exemples à propos de la puissance de calcul, des plateformes de données et des modèles de fondation.

Les processeurs graphiques (*GPU* en anglais) sont aujourd'hui les équipements de puissance de calcul les plus indispensables au fonctionnement des systèmes d'IA générative. Or, une seule entreprise — américaine — détient actuellement 80 % des parts de marché mondial de conception des processeurs graphiques.

Les centres de données (*data centers* en anglais) sont des plateformes qui permettent aux organisations et aux particuliers à la fois d'héberger des données et d'utiliser des systèmes d'IA. Dans le monde, trois entreprises — américaines — disposent des deux-tiers des parts de marché.

Le classement des modèles de fondation[7] les plus précis dénombre 30 modèles, créés par douze entreprises. Parmi elles, la majorité est américaine et l'Europe compte seulement trois entreprises classées, deux françaises et une allemande.

Cette domination américaine s'explique notamment par des investissements bien supérieurs à ceux consentis par la France et l'Europe. Les montants investis dans l'IA aux États-Unis sont ainsi 20 fois supérieurs à ceux investis en France. L'économie américaine est certes bien plus grande que la nôtre. Cependant, à richesse comparable, nous investissons environ trois ou quatre fois moins que les Américains.

…MAIS NOUS DISPOSONS D'ATOUTS : RIEN N'EST DONC DÉFINITIVEMENT JOUÉ

Le retard français et européen est d'ampleur et ne doit pas être occulté. Pour autant, notre continent dispose d'atouts qu'il ne faut pas négliger.

Des entreprises européennes sont positionnées sur l'ensemble de la chaîne de valeur et quelques-unes d'entre elles sont de tout premier rang. Par exemple, la première entreprise mondiale de fabrication de machines pour l'industrie des matériaux semi-conducteurs est néerlandaise.

Par ailleurs, l'Europe et particulièrement la France peut compter sur des professionnels précisément formés aux technologies de l'intelligence artificielle. La qualité de la formation supérieure française dans le domaine conduit de nombreuses entreprises étrangères à recruter des ingénieurs et chercheurs formés dans notre pays.

De plus, l'année 2023 a montré que l'avance de certains acteurs américains, dont OpenAI, n'était pas irrattrapable en ce qui concerne la production de modèles d'IA. De nombreux modèles devraient en 2024 rattraper ou dépasser GPT-4. La course à mener n'est pas seulement

7. Centre de recherche sur les modèles de fondation de Stanford (2024).

technologique, elle porte sur les modèles d'affaires, les produits, et la capacité à les servir à un grand nombre d'utilisateurs à bas coût.

Enfin, les innovations technologiques au cœur de l'IA générative sont récentes et la chaîne de valeur économique est très loin d'être arrivée à maturité. Le marché de l'IA est encore en devenir et la plupart des modèles d'affaires restent à inventer. Parallèlement, l'Europe devrait pouvoir s'appuyer sur son tissu économique existant pour se positionner sur les nombreux marchés par l'IA, c'est-à-dire intégrant l'IA dans leurs modèles d'affaires.

Il n'est donc pas trop tard pour (ré)agir. La France et l'Europe peuvent tirer leur épingle du jeu et ainsi retirer de nombreux bénéfices de l'intelligence artificielle. C'est dans cette perspective que notre Commission s'est attachée à élaborer un plan d'action ambitieux.

QUELQUES DÉBATS AU CŒUR DU PLAN D'ACTION QUE NOUS PROPOSONS

Pour parvenir à une proposition de plan d'action, notre Commission a soulevé de nombreuses interrogations et en a débattu au cours des six derniers mois. En voici quelques-unes.

- **Quel poids donner à la maîtrise de la création des systèmes d'IA et quelle place accorder aux technologies étrangères ?**

  S'engager résolument dans la compétition internationale de l'IA représente un investissement significatif. Cependant, faire partie des pionniers de l'IA s'accompagne de nombreux avantages : gains économiques accrus, moindre dépendance à l'égard de l'étranger, maîtrise du référentiel de valeurs sous-jacent au système d'IA, capacité à adapter la technologie à anticiper les effets de son déploiement, etc. Nous savons que les États-Unis tirent des bénéfices majeurs de leur maîtrise des précédentes vagues d'innovations numériques (ordinateurs, internet, téléphones multifonctions).

- **Pour pouvoir participer à la compétition internationale, faut-il des politiques publiques ciblées sur l'IA ou une action plus transverse pour favoriser l'innovation de manière générale ?**

  Il est très difficile pour l'État de cibler son soutien sur les prochaines innovations à succès, parce qu'il n'est pas nécessairement l'acteur le plus clairvoyant, mais aussi parce que nous sommes généralement surpris de l'innovation de rupture, qui naît du croisement inattendu des domaines de recherche et de modèles d'affaires. Cela milite donc pour des politiques publiques de soutien horizontal à l'innovation. Pour autant, l'IA présente un certain nombre de spécificités et de prérequis — les données, la puissance de calcul, des compétences particulières — qui peuvent justifier des actions ciblées. Dans tous les cas, le soutien public ne doit pas viser un rattrapage permanent (c'est-à-dire courir en permanence derrière des pionniers), mais une supériorité sur *certaines* briques de la chaîne de valeur (c'est-à-dire choisir *quelques* domaines de compétition pour les remporter).

- **Quelle confiance accorder aux acteurs privés à l'origine des systèmes d'IA et que faut-il que les pouvoirs publics nationaux, continentaux et internationaux encadrent et régulent ?**

  Les systèmes d'IA auront, dans les années à venir, des effets sur nos sociétés, de l'économie à l'organisation du travail en passant par le rapport à l'information et à la vérité. Ces effets seront de très grande ampleur, que la France soit ou non à l'origine de ces systèmes d'IA. Ces effets justifient l'intervention publique, afin de veiller à ce que l'IA soit déployée au service de l'intérêt général, de favoriser une répartition équitable des gains économiques et sociaux, de limiter les usages néfastes de l'IA ou encore de limiter les positions dominantes de quelques entreprises. Cependant, une intervention publique excessive ou mal dimensionnée pourrait nuire de manière disproportionnée à l'innovation et nous priver de ses bénéfices.

- **Quel équilibre choisir entre la protection des données et leur accessibilité ?**

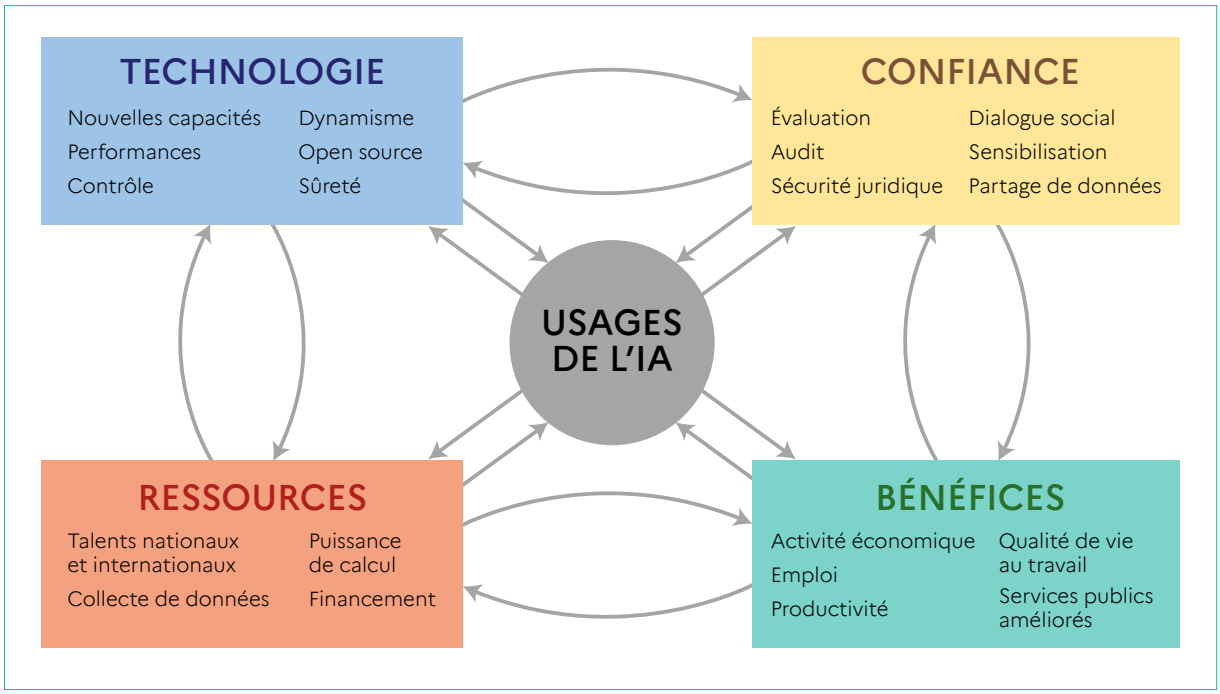
  Les technologies d'apprentissage automatique qui sont au cœur des systèmes d'IA reposent sur d'importantes masses de données. Dès lors, restreindre l'accès aux données conduit à restreindre l'innovation et de ses bénéfices ou à la laisser entre les mains de quelques acteurs capables de collecter plus de données et de supporter le coût de la régulation. Néanmoins, la restriction de l'accès aux données se justifie par de nombreux objectifs de protection : de la vie privée, de la propriété intellectuelle, etc. Par exemple, limiter l'accès aux données de santé des chercheurs assure une certaine confidentialité des données sensibles, mais freine aussi la découverte de nouveaux traitements thérapeutiques.

- **Quels sont les avantages et les risques d'un accès libre aux systèmes d'IA ?**

  L'accroissement de la puissance des modèles d'IA fait craindre une prolifération d'usages malveillants. L'accès libre à ces modèles de plus en plus puissants – auxquels des personnes malveillantes peuvent donc avoir facilement accès – peut apparaître préoccupant. Dans le même temps, l'*open source* facilite nettement le développement d'usages bienveillants, y compris les contre-mesures des usages malveillants. Le libre accès permet également d'élargir la base de contributeurs à l'élaboration des systèmes d'IA, ce qui peut contribuer à les rendre plus sûrs. De même, la protection des créations culturelles permet de faire rayonner la culture française en assurant une certaine indépendance économique aux filières, mais limite aussi la présence de la langue française dans les modèles d'IA.

- **Quels leviers faut-il actionner pour maîtriser l'IA ?**

  Pour que la France et l'Europe tirent pleinement profit de la révolution technologique de l'IA, plusieurs leviers relèvent de l'évidence : l'accessibilité à des données de qualité, la disponibilité de la puissance de calcul, une capacité d'investissements, du personnel expert dans le domaine. Ces quelques leviers apparaissent cependant bien insuffisants. Quelles actions permettront d'orienter l'innovation selon nos objectifs politiques ? Lesquelles accroîtront la confiance dans les outils d'IA ? Faut-il faire évoluer le service public ? Faut-il faire évoluer le cadre juridique ? On le voit, les dimensions sociales, économiques et juridiques ne peuvent pas être prises séparément et il est nécessaire d'actionner simultanément de nombreux leviers pour développer et orienter les projets d'IA…mais selon quelles priorités ?



Le cercle vertueux de l'IA entre ressources, développement technologique, confiance, et bénéfices économiques et sociaux

Le présent rapport restitue les termes du débat et apporte un ensemble de réponses afin de créer un cercle vertueux. Celles-ci reflètent les convictions de notre Commission. Ces convictions, forgées entre septembre 2023 et février 2024, se fondent sur l'expertise des membres et des rapporteurs. Elles sont étayées par la consultation de 7 000 personnes et l'audition de 600 experts et parties prenantes.

LE DÉBAT INCONTOURNABLE : À QUOI RESSEMBLERA LE MONDE DE DEMAIN ?

L'intelligence artificielle a moins d'un siècle. Les récents développements technologiques, qui ont surpris plus d'un, se sont révélés fulgurants. Il est certain que de nouvelles innovations viendront après la vague de l'IA générative que nous connaissons aujourd'hui. Quelles seront ces prochaines innovations ? Viendront-elles du prolongement de l'IA générative ou d'une rupture technologique ? Que doit-on faire pour être dans la course pour les prochaines étapes de l'innovation ?

Les systèmes d'IA sont déjà omniprésents dans nos sociétés et nous devons nous attendre à ce que leurs effets soient de plus en plus prégnants. Mais jusqu'où ces effets iront-ils ? À quel rythme ? Quels pans de la société seront les premiers ou les principaux concernés ? Comment tirer le maximum de bénéfices des innovations, tout en en réduisant les effets négatifs ? Que peut-on faire pour préparer notre société aux ruptures à venir ? Et au-delà, quelle société voulons-nous ?

En 1881, l'Exposition universelle de Paris déploie les premiers éclairages à incandescence. À l'Exposition universelle de 1900, la fée électricité triomphe. Qui pourtant pouvait anticiper l'ampleur des mutations qu'allait connaître le monde sous l'effet de cette nouvelle forme d'énergie ?

Au même moment, la voiture faisait son apparition au domicile de quelques individus fortunés. Qui pouvait anticiper que la démocratisation de la voiture affecterait l'environnement, redessinerait les villes, transformerait notre rapport à la distance et modifierait donc nos interactions sociales ?

Notre Commission, hélas, est dépourvue de boule de cristal. Nous ne prétendons donc pas voir juste, là où tant d'autres ont vu faux par le passé. Nous espérons toutefois avoir posé les bonnes questions, soulevé les bons problèmes. Certaines de nos propositions se révéleront sans doute erronées, tant les prochaines évolutions technologiques et sociétales sont incertaines. Nous croyons cependant que le plan d'action recommandé permettra à la France et à l'Europe d'entrer résolument dans la compétition internationale de l'IA, et de placer l'innovation au service de nos principes, de nos valeurs et de nos intérêts.

1

DÉDIABOLISER L'IA, SANS POUR AUTANT L'IDÉALISER

1.1 L'IA ME CONCERNE-T-ELLE ?

> **Oui !** L'IA nous concerne tous. Comme utilisateurs, car nous utilisons déjà des services intégrant de l'IA dans notre quotidien. Et comme citoyens, car nous aurons à décider comment nous voulons utiliser ces technologies. Entre 1920 et 1930, la plupart des villes et pays occidentaux ont décidé de séparer les flux de piétons des voitures, afin d'éviter les accidents et de pouvoir se déplacer le plus vite possible en voiture. Cette décision a changé nos villes et notre quotidien pour un siècle au moins. Les multiples choix autour de l'IA nous concerneront tout autant.

Le lancement de ChatGPT fin 2022 a placé l'IA au bout de nos doigts. Un an plus tard, 55 % des Français disent bien connaître ChatGPT et 28 % disent en avoir entendu parler sans bien savoir ce dont il s'agit[8]. ChatGPT n'est cependant que la face visible de l'iceberg. De nombreux observateurs soulignent que l'IA n'est pas apparue avec les robots conversationnels comme ChatGPT et ils ont entièrement raison : la recherche en matière d'intelligence artificielle a débuté dès les années 1950 ; elle est, en réalité, presque aussi ancienne que l'informatique.

Au fil des années, les progrès de l'IA se sont traduits par quelques exploits retentissants dans des activités que l'on croyait auparavant réservées à l'humain, comme les échecs et le go. Parallèlement, l'IA s'est immiscée dans de nombreuses activités du quotidien. Nos téléphones l'utilisent pour se déverrouiller en reconnaissant notre visage ou nos empreintes, elle aide à la traduction en langue étrangère, à la reconnaissance d'images, au sous-titrage automatique de vidéos, à la détection des fraudes, à la recommandation de produits… Lorsqu'une publicité s'affiche sur notre écran, c'est souvent une intelligence artificielle qui l'a choisie. Ce que nous permet ChatGPT, c'est notamment de pouvoir interagir (voire jouer !) avec l'IA et d'imaginer des cas d'usage.

8. Étude en ligne réalisée par Ipsos pour Sopra Steria en octobre et novembre 2023 auprès de 1 000 personnes représentatives de la population française âgée de 18 ans ou plus.

L'importance de l'IA n'est pas uniquement liée à ce qu'elle sait faire, quoique battre les meilleurs joueurs d'échecs ou de go reste impressionnant. C'est sa diffusion très rapide dans de très nombreux produits et services qui lui confère un pouvoir transformateur. En ce sens, elle est souvent comparée à d'autres innovations, qui, en leur temps, ont profondément transformé nos vies, nos économies et le fonctionnement même de nos sociétés, comme l'électricité ou le téléphone. L'IA, intégrée dans nos systèmes de communication et d'information, transformera aussi nos démocraties.

Il nous a fallu quelques décennies pour nous approprier les précédentes révolutions technologiques, nous devons donc dessiner sans attendre une société avec l'IA.

Recommandation n° 1

Créer les conditions d'une appropriation collective de l'IA et de ses enjeux afin de définir collectivement les conditions dans lesquelles elle s'insère dans nos vies quotidiennes.

1.2 FAUT-IL AVOIR PEUR DE L'IA ?

> **Non, mais il faut être vigilant comme avec tout outil.** L'IA actuelle ne va pas conduire à la fin de l'humanité. En revanche, les systèmes d'IA s'accompagnent déjà d'un ensemble de risques qui nécessitent d'être gérés.

Au printemps 2023, 60 experts de l'IA et personnalités mondialement connues ont signé une déclaration qui a fait grand bruit[9] : ils avertissaient que prévenir le risque d'une extinction de l'humanité causée par une IA hors de contrôle devait être une priorité mondiale, au même titre que prévenir les pandémies ou les conflits nucléaires. De fait, les discours entourant l'IA portent autant sur ses risques et ses dangers que sur son potentiel pour améliorer le quotidien de l'humanité, et souvent ce sont les risques les plus extrêmes qui sont mis en avant. L'extinction de l'humanité, ce n'est quand même pas rien !

Cette ambivalence dans le traitement de l'IA influe sur la manière dont nous la percevons. La fascination côtoie la crainte. Un an après le lancement de ChatGPT, 77 % des Français déclarent que l'IA est une vraie révolution[10], mais 68 % pensent qu'il faudrait faire une pause dans le développement de l'IA[11]. Nous ne sommes d'ailleurs pas les seuls, puisque 79 % des Chinois et 74 % des Américains y sont favorables.

Dessin anti-électricité de 1900.

Sur ce plan, la situation de l'IA n'a rien de très original : toutes les technologies qui ont bouleversé notre quotidien ont, en leur temps, suscité des peurs, certaines imaginaires, d'autres bien réelles. La crainte que la vitesse des trains rende aveugles leurs passagers s'est avérée entièrement infondée, mais le développement du chemin de fer a aussi été la source d'incidents, parfois sérieux, qui ont imposé une réponse des pouvoirs publics : les tunnels ferroviaires ont ainsi longtemps été considérés comme un milieu insalubre, voire dangereux. Même la « fée électricité » suscitait en 1900 des craintes sur les risques d'électrocution dans la rue.

9. « *Pause Giant AI Experiments: An Open Letter* » publiée par le Future of Life Institute le 22 mars 2023

10. Étude en ligne réalisée par Ipsos pour Sopra Steria en octobre et novembre 2023 auprès de 1 000 personnes représentatives de la population française âgée de 18 ans ou plus.

11. Étude en ligne réalisée par Ipsos pour AXA en mai et juin 2023 auprès de 3 226 experts de 50 pays et de 19 000 personnes dans 15 pays, représentatives de la population nationale âgée de 18 ans et plus.

De la même manière que le train ou l'électricité, qui nous rendent d'immenses services et font partie de notre vie de tous les jours, l'IA présente des risques. Ces risques ne doivent pas être ignorés et ils appellent une réponse. Cette réponse doit être proportionnée et ne doit pas nous priver des bénéfices de l'IA. D'autant qu'il existera non pas « une IA », mais de nombreux outils intégrant des fonctionnalités d'IA. Par exemple, un outil d'IA dans la santé n'implique pas les mêmes risques qu'un outil d'IA dans la publicité en ligne, les précautions doivent donc être différentes.

Les risques liés à la diffusion de l'IA peuvent être regroupés en trois grandes catégories.

Des risques d'imperfection. De nombreux systèmes d'IA fonctionnent avec des probabilités. C'est ce qui leur donne leur flexibilité, mais aussi leur capacité de se « tromper ». Quand je demande à un modèle d'IA générative de m'expliquer « comment fonctionne une voiture », il peut me donner une réponse très bien formulée, mais fausse. Quand je lui demande de me dessiner une voiture, il peut oublier d'ajouter les portes de la voiture. C'est ce que l'on appelle les « hallucinations ». Par exemple, un avocat américain ne s'est pas rendu compte que ChatGPT lui avait fourni des exemples de cas totalement fictifs. Certaines de ces imperfections disparaitront avec le progrès technologique, d'autres n'auront pas d'incidence notable. Certaines perdureront, notamment quand les données d'entraînement seront biaisées ou fausses. Dans les domaines sensibles, comme la santé ou le maintien de l'ordre, l'utilisation de l'IA doit donc être évaluée soigneusement et encadrée.

Des risques d'utilisation malveillante. Les cybercriminels n'ont pas attendu longtemps pour utiliser l'IA générative afin de produire des faux particulièrement convaincants, tout comme leurs prédécesseurs se sont saisis de la voiture il y a 100 ans pour s'enfuir après un braquage. Rien n'indique toutefois que l'IA changera durablement le rapport de force entre cybercriminels et ceux chargés de nous protéger, à condition que ces derniers puissent s'emparer de ces technologies. L'IA générative n'apparaît pas non plus rendre la production d'armes physiques, chimiques ou biologiques plus facile qu'une recherche en ligne[12].

Risques liés aux imperfections	Risques liés à une utilisation malveillante	Risques systémiques
Discrimination et reproduction de stéréotypes	Cyber-criminalité	Concentration du pouvoir
Mésinformation	Cyber-terrorisme	Disruption du marché du travail
Violation de la vie privée/divulgation d'informations confidentielles	Bio-sécurité	Affaiblissement de la diversité culturelle et linguistique
Accidents	Désinformation	Diversité culturelle et normative
Production de contenus illicites ou préjudiciables	Surveillance de masse	Consommation d'électricité et émissions de gaz à effet de serre
Violation des droits de propriété intellectuelle		Accident systémique
		Comportement émergent critique

Risques liés aux systèmes d'IA générative
Source : Commission de l'IA

12. Mouton, C. A., C. Lucas, et E. Guest (2024) « The Operational Risks of AI in Large-Scale Biological Attacks: Results of a Red-Team Study », *RAND Corporation*.

Des risques systémiques. Le développement de l'IA peut être source de risques pour l'ensemble de la société voire de l'humanité. Si le risque que l'IA se traduise par une destruction massive d'emplois semble limité, nous devons, cependant, nous préparer à ce que certains métiers soient fortement transformés, voire disparaissent (voir 1.4 *L'IA : créatrice ou destructrice d'emplois ?*). D'autres risques découlent de la concentration du développement des systèmes d'IA les plus en pointe entre les mains d'un petit nombre de pays, d'entreprises et de personnes. Ils appellent une réponse déterminée en termes de politique industrielle et de concurrence. En revanche, aucune IA n'est à ce jour en mesure de dépasser l'intelligence humaine dans toutes les tâches et moins encore de représenter une menace existentielle pour l'ensemble de l'humanité. Ces perspectives, encore hypothétiques et qui ne se concrétiseront peut-être jamais, ne peuvent constituer l'alpha et l'oméga de notre approche de l'IA. Elles imposent néanmoins de rester vigilant. Il est donc nécessaire que notre pays se dote d'une capacité d'évaluation des systèmes d'IA les plus avancés qui permette d'anticiper l'apparition de nouveaux risques et d'une gouvernance qui permette d'y répondre, au plan national comme au plan international.

1.3 L'IA NOUS RENDRA-T-ELLE PLUS PROSPÈRES ?

Certainement, car l'IA nous rendra plus productifs. L'ampleur de ces gains et la façon dont ils seront répartis dans la société sont incertaines et ne sont pas définies a priori. En 10 ans, ils pourraient augmenter le PIB de 250 à 420 milliards d'euros, soit autant que la valeur ajoutée de toute l'industrie.

Depuis les Trente Glorieuses, l'économie française, et plus généralement les économies de la majorité des pays développés, ont connu une baisse de leur taux de croissance. Certains économistes en ont conclu que cette faible croissance était inéluctable, parlant même de « stagnation séculaire ».

L'IA pourrait nous aider à retrouver des taux de croissance élevés dans nos économies développées par deux effets : (i) en augmentant notre productivité, c'est-à-dire la vitesse à laquelle nous produisons des biens et des services ; (ii) en augmentant notre capacité à générer de nouvelles idées, et donc de nouvelles innovations, de nouveaux produits ou de nouvelles formes d'organisation.

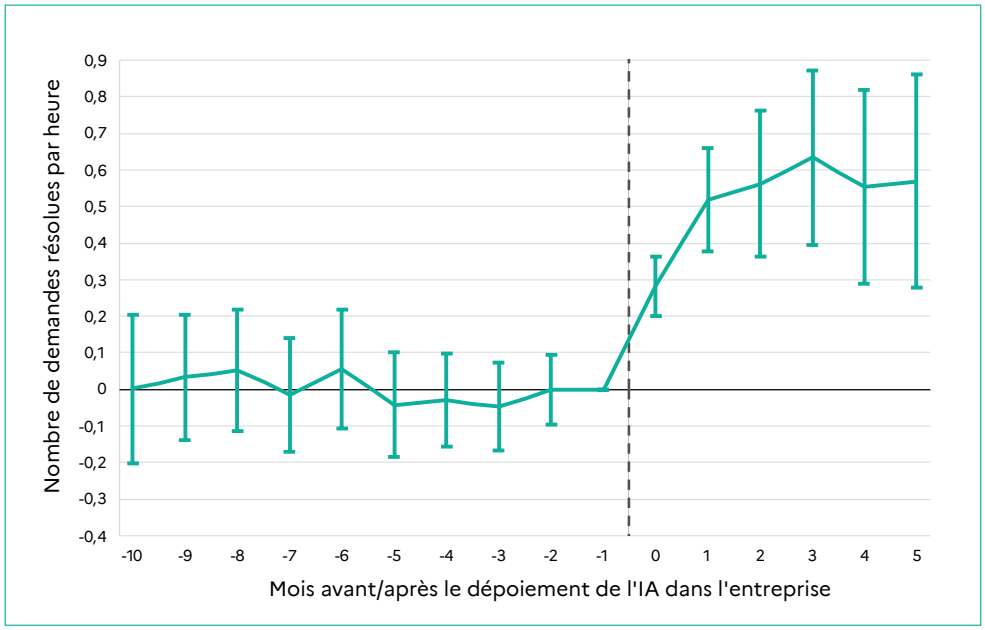
L'IA AUGMENTE SIGNIFICATIVEMENT LA PRODUCTIVITÉ

L'IA peut augmenter notre potentiel de croissance en automatisant des tâches dans la production de biens et services. Elle contribue ainsi à l'augmentation de la productivité, comme cela s'est produit avec la mécanisation dans le secteur agricole, l'invention de la chaîne de montage dans l'industrie, ou plus récemment la numérisation d'une part importante de l'économie. Les gains de productivité sur ces tâches contribueront à augmenter le taux de croissance.

Une récente étude américaine[13] s'intéresse justement aux effets de l'adoption de l'IA générative sur la productivité d'employés du service client d'une entreprise. L'entreprise a progressivement déployé un outil d'IA aidant les employés chargés de répondre à la clientèle via un *chat* en ligne en leur proposant des réponses générées automatiquement. Dans le cadre de cette étude, l'effet du déploiement de cet outil est substantiel : la productivité des employés ayant eu accès

13. Brynjolfsson, E., D. Li, et L. Raymond (2023), « Generative AI at Work », *NBER Working paper.*

à l'assistant IA a crû de 25 %, dont 14 % dès le premier mois d'utilisation. L'effet est donc immédiat et persistant sur les cinq mois d'étude.



Graphique 1 : Effet de l'adoption de l'IA générative sur la productivité d'employés d'un service client.
Source : Brynjolfsson, Li et Raymond (2023)
Lecture : Les employés ayant accès à l'IA voient leur productivité augmenter davantage que ceux n'y ayant pas accès, alors que leurs productivités évoluaient de façon similaire dans les 10 mois précédant l'introduction de l'IA.

Bien entendu, ces résultats ne concernent qu'un type précis de profession au sein d'une entreprise donnée. Toutefois, deux études[14] se concentrant sur des individus hautement qualifiés aux États-Unis (consultants, managers, etc.) montrent que l'utilisation de ChatGPT permet une augmentation de la productivité des tâches typiques de ces métiers entre 25 % et 35 %.

Cela semble montrer que les gains de productivité sont observés au sein d'un panel de professions large, avec des niveaux de qualification différents. Ces effets sont exprimés par les travailleurs eux-mêmes. En France, une vaste enquête de Pôle emploi « Les employeurs face à l'intelligence artificielle » (juin 2023[15]) met aussi en avant un effet positif sur la productivité : 72 % des employeurs recourant à l'IA mentionnent un impact positif sur la performance de leurs salariés, en permettant notamment de réduire les tâches fastidieuses (63 %) ou le risque d'erreur (51 %).

Si l'on cherche à dépasser le cadre d'une entreprise pour estimer l'impact global de l'IA sur l'économie, deux questions se posent : quand verra-t-on les gains économiques de l'IA, et quelle sera leur ampleur ? Nous pouvons faire le parallèle avec les effets sur la productivité des révolutions technologiques passées. Aux États-Unis comme en Europe, dans le cas de l'électricité,

14. Noy, S., et W. Zhang (2023), « Experimental Evidence on the Productivity Effects of Generative Artificial Intelligence », Science. // Dell'Aqua F., E. McFowland, E. Mollick, H. Lifshitz-Assaf, K. Kellogg, S. Rajendran, L. Krayer, F. Candelon, et Lakhani (2023), « Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality », *Working paper.*

15. Enquête menée par Pôle emploi auprès de 3 000 établissements de 10 salariés ou plus en juin 2023. L'étude ne précise pas si l'IA utilisée est générative ou non, alors que les études américaines se concentrent sur l'IA générative.

les gains de productivité se sont matérialisés une vingtaine d'années après l'invention de la technologie.

Pour bien comprendre ce décalage temporel, examinons le cas de l'électricité. Au début du XXe siècle, son adoption dans les entreprises était encore limitée. Les usines maintenaient une organisation interne semblable à celle qu'elles avaient adoptée lorsqu'elles étaient alimentées par des moulins à eau : avec un arbre de transmission central que faisait tourner l'énergie hydraulique. Ni l'apparition de la vapeur avec la première révolution industrielle ni celle de la dynamo au début de la seconde révolution industrielle n'ont conduit à une modification de l'organisation interne des usines. Or, la présence de cet arbre de transmission obligeait à placer les machines similaires côte à côte. Il a fallu attendre les années 1910 pour voir les gains de productivité liés à l'électricité, grâce à l'invention du fil électrique et la miniaturisation des moteurs électriques. Chaque machine devient alors alimentée de manière autonome par l'électricité. Cette innovation élimine l'arbre de transmission et permet un agencement plus efficace des machines : c'est l'invention de la chaîne de montage, qui marque la hausse de la productivité des usines.

Le décalage temporel existera également pour l'IA, car son adoption nécessitera aussi bien une évolution de l'organisation du travail au sein des entreprises que des investissements complémentaires. Toutefois le décalage devrait être moins marqué avec l'IA générative, qui est bien plus aisée à déployer dans l'économie, et s'applique bien aux emplois de service de notre économie. L'industrie du jeu vidéo, par exemple, déploie l'IA générative rapidement pour générer l'esquisse d'un jeu en 2 mois au lieu de 6 mois, et générer beaucoup d'idées différentes d'un personnage. Par ailleurs, le foisonnement d'innovations dans l'IA date d'une dizaine d'années. Nous pourrions donc commencer à observer bientôt des gains de productivité.

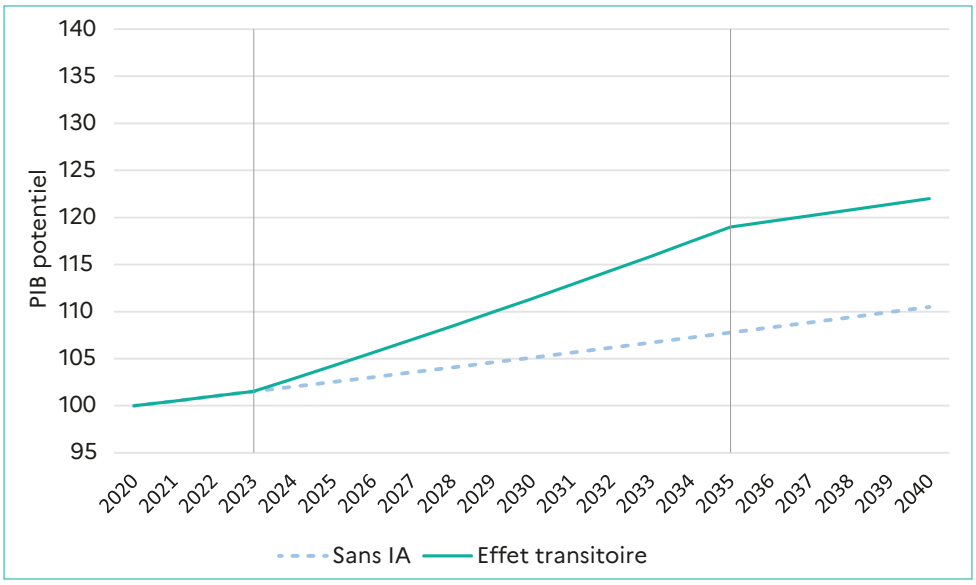
Cela nous conduit à une seconde question : quelle est l'ampleur des gains économiques à attendre ? Si l'on considère que les gains de productivité permis par la vague de l'IA au cours de la prochaine décennie seront comparables à ceux de la vague de l'électricité dans les années 1920 en Europe, la croissance de la productivité augmenterait alors de 1,3 point de pourcentage par an à partir de 2024[16].

Si l'on préfère prendre comme point de comparaison la vague des technologies numériques de la fin des années 1990 et du début des années 2000 aux États-Unis, la hausse de la croissance de la productivité serait d'environ 0,8 point de pourcentage par an. En comparaison, la productivité potentielle de la France[17] est aujourd'hui estimée à 0,5 % par an à moyen terme. En étudiant plus en détail le différentiel de croissance entre la France et les États-Unis sur la période 1997-2006, on constate effectivement que ce sont les secteurs producteurs ou fortement utilisateurs de technologies numériques qui expliquent la quasi-totalité de l'écart observé : le développement et le déploiement d'une nouvelle technologie aussi déterminante que le numérique sont le principal facteur explicatif des écarts de prospérité entre les États-Unis et la France sur cette période.

Toutes choses égales par ailleurs[18], les gains générés par l'IA augmenteraient significativement le taux de croissance de la France, estimé à 1,35 % par an à moyen terme. De tels gains de productivité pendant dix ans conduiraient à une hausse du PIB comprise dans une fourchette allant de 250 Md€ à 420 Md€ en 2034, soit l'équivalent de la valeur ajoutée de l'industrie dans son ensemble !

16. Estimation réalisée à partir des données de Bergeaud, A., G. Cette, et R. Lecat (2016), « Productivity Trends in Advanced Countries between 1890 and 2012 », *Review of Income and Wealth*, 62(3), pp 420–444.
17. Programme de stabilité pour la période 2022-2027.
18. En transposant ces gains de productivité au PIB potentiel ; sans évolution conjointe de l'emploi, point discuté en détail au 1.4. L'*IA : créatrice ou destructrice d'emplois ?*

En revanche, cette hausse de la croissance ne serait que transitoire : une fois l'IA adoptée par l'ensemble du tissu économique, les gains de productivités liés à cette adoption et les transformations engendrées cessent, c'est ce que présente le graphique ci-dessous.



Graphique 2 : Effet transitoire attendu de l'adoption de l'IA sur la croissance.
Source : Commission IA.

Cette prédiction sera cependant considérée comme trop pessimiste pour certains, trop optimiste pour d'autres. Les premiers feront valoir que l'IA permet d'automatiser également la production d'idées, ce qui permet de générer un surcroît de croissance, et ce, de façon permanente. Les seconds mentionneront l'existence d'obstacles à la croissance, notamment l'absence de concurrence sur différents segments de la chaîne de valeur de l'IA.

L'IA DEVRAIT ÉGALEMENT FACILITER LA PRODUCTION DE NOUVELLES IDÉES

L'IA pourrait automatiser la production de nouvelles idées, ou du moins la faciliter. Elle nous aidera donc à générer de nouvelles inventions et à résoudre des problèmes complexes, comme dans l'exemple d'AlphaFold qui permet de trouver de nouvelles protéines[19], ou de GNoMe qui suggère de nouveaux matériaux qui pourraient équiper nos véhicules ou nos objets du quotidien[20].

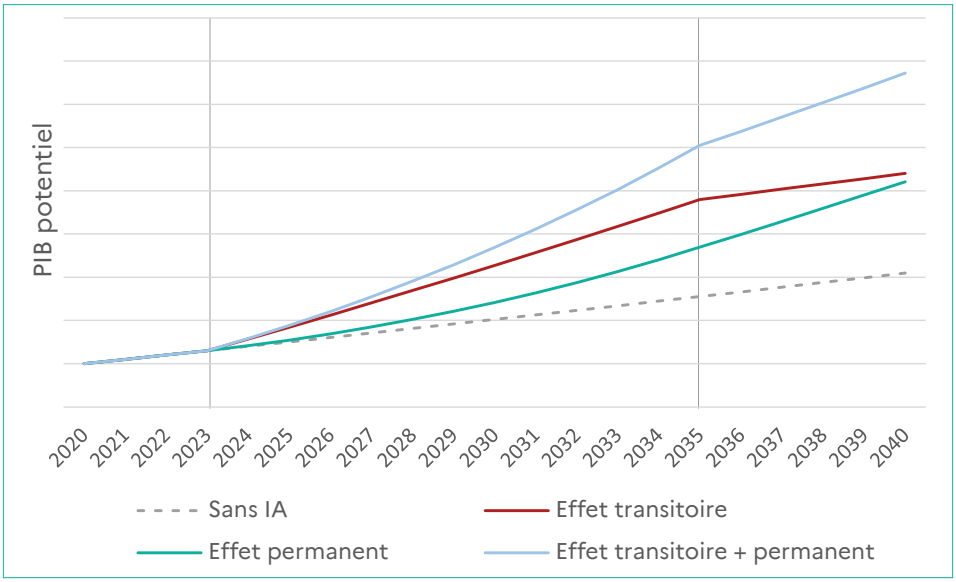
L'impact de l'IA sur la science et l'innovation est difficile à quantifier. D'autant que la capacité de l'IA à générer de nouvelles idées pourrait se heurter à des difficultés pratiques : il ne suffit pas d'identifier 2,2 millions de nouveaux matériaux potentiels pour les produire, il faut encore les valider expérimentalement. Au minimum, l'IA facilitera le travail des chercheurs. Si des outils d'IA appuient progressivement l'humain pour identifier de nouvelles hypothèses, créer des protocoles et réaliser des expériences, alors la production d'idées pertinentes augmentera.

19. Jumper, J., Evans, R., Pritzel, A., *et al*. (2021) « Highly accurate protein structure prediction with AlphaFold ». *Nature 596*
20. Merchant, A., Batzner, S., Schoenholz, S.S. *et al*. (2023) « Scaling deep learning for materials discovery ». *Nature 624*

Cette perspective, incertaine, se profile toutefois à l'horizon. Près d'un article de recherche sur dix mentionne déjà l'utilisation de l'IA[21].

Quel serait alors l'effet de cette capacité de l'IA à générer de nouvelles idées sur la prospérité ? De nouveau, faisons un parallèle historique pour illustrer en quoi une innovation peut impacter sur le long terme le taux de croissance de la productivité. Au XVIIe siècle, l'invention du calcul infinitésimal a rendu possibles de gigantesques progrès en physique, notamment dans la compréhension des mouvements des projectiles ou des planètes. De même, les progrès effectués dans les techniques de polissage du verre ont conduit l'humanité à voir de plus en plus petit et de découvrir le monde des microbes, inconnu jusqu'alors. Ainsi, le microscope a permis des avancées cruciales en médecine. Au même titre, l'IA ouvre un champ des possibles difficile à imaginer. Ces effets induisent donc une augmentation permanente du taux de croissance de la productivité. La magnitude de cet effet n'est toutefois pas quantifiable.

Projetons-nous à présent dans un avenir cumulant les effets transitoires sur la croissance, liés à l'automatisation de la production des biens et services, et les effets permanents sur la croissance, liés à l'automatisation de la production de nouvelles idées. C'est le scénario bleu dans le graphique ci-dessous, avec des gains potentiels plus importants encore que les 250 à 420 milliards d'euros d'ici 2034 présentés plus haut.



Graphique 3 : Effets totaux attendus de l'adoption de l'IA sur la croissance.
Source : Commission IA.

21. Analyse du journal Nature sur la base de données Scopus.

L'IMPORTANCE DES INSTITUTIONS ET DE LA CONCURRENCE

Ne sommes-nous pas trop optimistes dans nos prédictions de croissance : après tout, la révolution numérique, elle aussi, devait conduire à une accélération de la croissance. Or depuis le début des années 2000 les pays développés à commencer par les États-Unis, ont connu un fort déclin de leurs taux de croissance. Comment expliquer nos faibles taux de croissance malgré des innovations majeures qui ont considérablement changé notre quotidien : l'ordinateur, le smartphone, les réseaux sociaux, etc.

Pour certains économistes il s'agit d'un simple problème de mesure, pour d'autres le signe que ces innovations numériques ont surtout amélioré notre divertissement[22]. Une explication plus convaincante est que la révolution des technologies de l'information et de la communication (TIC) a favorisé l'émergence d'entreprises « *superstar* », notamment les Gafam (Google, Amazon, Facebook, Apple, Microsoft). Issues de la révolution des TIC, les Gafam ont dans un premier temps contribué à l'augmentation observée du taux de croissance de la productivité pendant la décennie 1995-2005. Cependant, une politique de concurrence trop laxiste a permis aux Gafam de grossir jusqu'à contrôler certains pans entiers de l'économie américaine, ce qui a fini par décourager l'entrée de nouvelles entreprises innovantes avec des effets négatifs sur la croissance de l'économie dans son ensemble[23].

La différence entre la révolution des TIC et celle de l'IA est que cette fois les Gafam sont dominants dès le début, et donc peuvent immédiatement décourager l'entrée de nouvelles entreprises innovantes. L'absence de concurrence est particulièrement prononcée dans les segments amont de la chaîne de production de l'IA, à savoir l'accès aux données et l'accès à la puissance de calcul : ces segments sont dominés par un petit nombre de géants dont les Gafam. D'où l'importance d'adapter nos institutions et en particulier nos politiques de concurrence, pour que la révolution de l'IA puisse pleinement agir comme facteur de croissance (voir 2.3.3. *Éviter les positions concurrentielles dominantes*).

À L'INVERSE, UN RISQUE DE DÉCLASSEMENT ÉCONOMIQUE ET GÉOPOLITIQUE

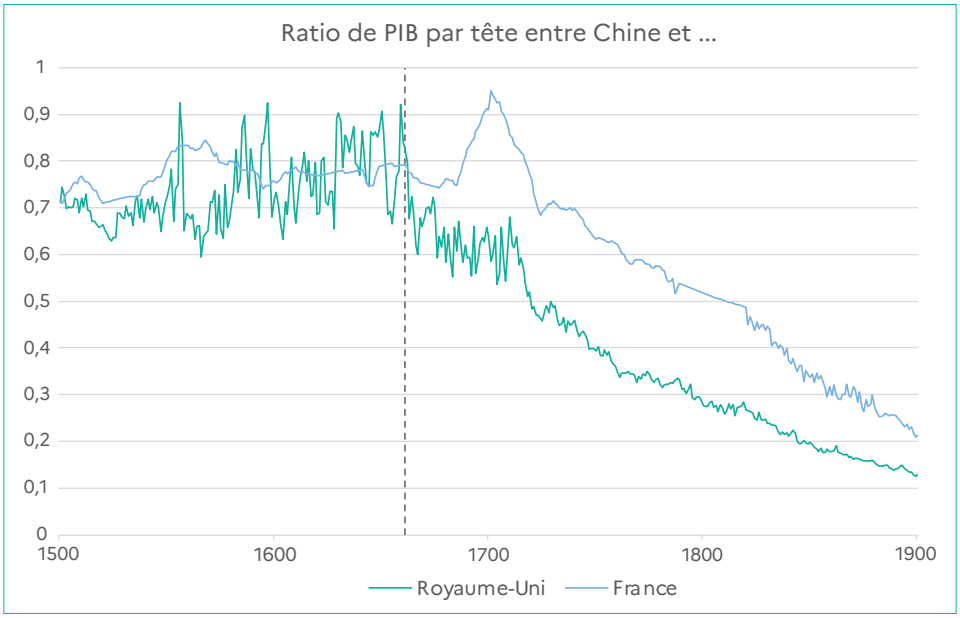
Le déploiement des systèmes d'intelligence artificielle dans notre économie et notre société n'est pas sans risques. Que se passerait-il donc si la France et l'Europe choisissaient de rester à l'écart de cette révolution technologique ? Il faut d'abord souligner qu'empêcher totalement la diffusion des systèmes d'IA dans la société est quasiment impossible. Cela supposerait en particulier d'isoler notre pays de la circulation mondiale des informations et d'imposer un contrôle très strict de la population. À part la Corée du Nord, peu de pays empruntent cette voie.

Si l'isolement total est donc peu enviable, que se passerait-il si l'on freinait la diffusion de l'innovation ? Pour y répondre, tournons-nous vers l'histoire des révolutions industrielles. Le cas

22. Comme le disait l'investisseur Peter Thiel en 2022 sur les réseaux sociaux : « On nous avait promis des voitures volantes, et n'avons eu droit qu'à 140 caractères ».
23. Aghion P., A. Bergeaud, T. Boppart, P. Klenow et H. Li (2023), « A Theory of Falling Growth and Rising Rents », *Review of Economics Studies*, 90(6).

de la Chine est très éclairant. Durant des siècles, *l'empire du Milieu* est de très loin la première puissance mondiale. À partir de la seconde moitié du XVIIe siècle, les dynasties Ming et Qing font le choix de s'opposer au commerce international par crainte des nouvelles technologies et de ses effets sur l'économie et la société[24]. Le contrôle de l'État sur la société et l'économie se renforce, allant même jusqu'à ordonner aux habitants vivant le long de la côte sud, de se déplacer de 30 kilomètres à l'intérieur des terres.

Cet isolement vis-à-vis de l'innovation a changé le cours de l'histoire de la Chine. Au mitan du XVIIe siècle, le PIB par habitant en Chine est seulement inférieur de 10 % à 20 % à celui du Royaume-Uni ou de la France[25]. Puis la Chine ne bénéficie pas des effets des révolutions industrielles qui se déploient en Europe. Deux siècles plus tard, le PIB par habitant en Chine est cinq fois plus faible ; dix fois plus faible en 1900 (graphique 4). Le décrochage économique conduit à un affaissement de la souveraineté de l'Empire, qui se voit imposer des traités inégaux par les puissances étrangères : c'est la « ruée sur la Chine ».



Graphique 4 : Comparaison du PIB par tête de la Chine à ceux du Royaume-Uni et de la France
Source : Commission IA à partir des données du Maddison Project Database 2020.

De nombreux autres exemples montrent que l'absence de participation à une révolution technologique conduit à un déclassement économique et sociétal, qui hâte la domination étrangère et attise la convoitise internationale. Prévenir cette trajectoire délétère ne nécessite pas de s'emparer sans boussole de l'IA, mais de s'organiser pour les maîtriser et les orienter selon nos objectifs politiques.

24. Acemoglu D. et J. Robinson (2012), « Why Nations Fail: The Origins of Power, Prosperity, and Poverty », Crown Business.
25. Selon les données 2020 du Maddison Project Database, base de données de référence pour les estimations historiques du PIB.

1.4 L'IA : CRÉATRICE OU DESTRUCTRICE D'EMPLOIS ?

> **Notre propre analyse empirique suggère un effet positif de l'IA sur l'emploi dans les entreprises qui adoptent l'IA, car celle-ci remplace des tâches, et non des emplois.** Dans 19 emplois sur 20, il existe des tâches que l'IA ne peut pas accomplir. Les emplois directement remplaçables par l'IA ne représenteraient donc que 5 % des emplois d'un pays comme la France. Par ailleurs, la diffusion de l'IA va créer des emplois, dans de nouveaux métiers, mais aussi dans d'anciens métiers. Au total, certains secteurs ou certains domaines pourraient connaître des baisses nettes d'emplois, qui doivent être accompagnées par les pouvoirs publics, mais cela n'implique pas que l'IA aura un effet négatif sur l'emploi national en France.

Des scénaristes et des acteurs d'Hollywood qui craignent pour leurs emplois, une entreprise française de veille médiatique qui déclenche un plan de sauvegarde de l'emploi sous couvert de l'IA : les actualités et les prises de position font redouter une fin du travail et un chômage technologique de masse.

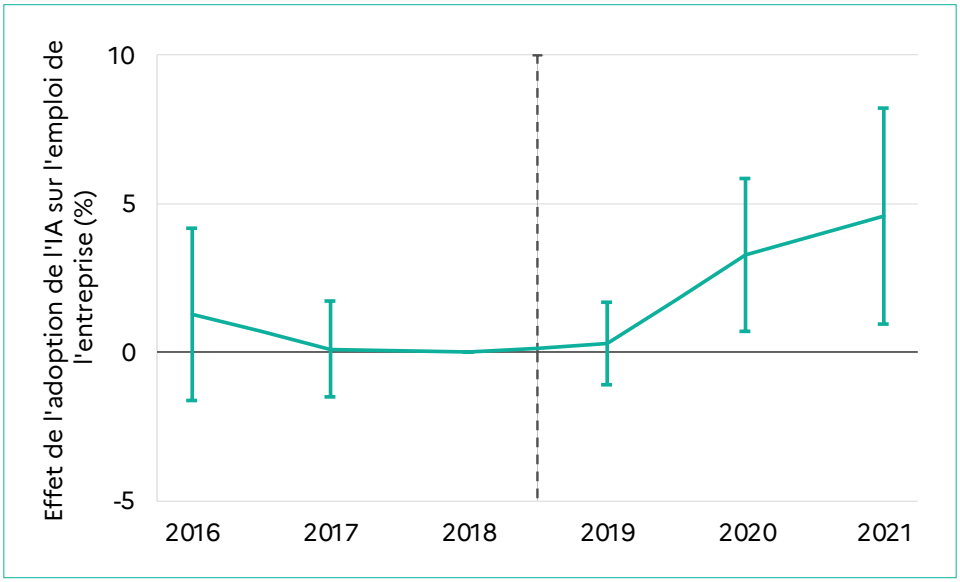
L'intelligence artificielle permet en particulier l'automatisation de tâches, qui est moteur essentiel de la croissance économique (voir *1.3. L'IA nous rendra-t-elle plus prospères ?*), implique deux effets contraires sur l'emploi. D'un côté, l'automatisation déplace certaines tâches du travail humain vers les machines, ce qui tend à détruire des emplois : c'est l'**effet d'éviction**. De l'autre, l'automatisation augmente la productivité des individus, ce qui conduit à une augmentation du rapport qualité/prix des produits proposés aux consommateurs, donc à une demande plus élevée et, *in fine*, davantage d'embauches et la création de nouvelles tâches : c'est l'**effet de productivité**.

Pour comprendre l'effet d'une nouvelle technologie sur l'emploi et le marché du travail, il faut donc comprendre qui de ces deux effets l'emporte. Pour cela, deux principales approches ont été retenues. Premièrement, en étudiant directement les effets de l'adoption de l'IA au sein des entreprises ou des secteurs. Deuxièmement, en étudiant les effets de l'IA sur les différentes tâches composant l'économie.

PREMIÈRE APPROCHE : LES EFFETS SUR L'EMPLOI DE L'ADOPTION DE L'IA PAR LES ENTREPRISES

L'approche n'est pas nouvelle et a déjà été utilisée pour mesurer les effets sur l'emploi d'autres révolutions technologiques. Par exemple, une étude[26] réalisée à partir de données françaises et s'intéressant à l'adoption des machines industrielles (machines-outils, robots, etc.) montre que les entreprises qui adoptent davantage ce type de nouvelles technologies baissent leurs prix, augmentent leurs ventes et leur emploi davantage que leurs concurrentes n'ayant pas adopté les technologies d'automatisation.

Qu'en est-il dans le cas de l'IA ? Une enquête réalisée annuellement par l'Insee permet d'étudier les effets de l'adoption de l'IA par les entreprises en France[27]. On constate que l'emploi total des entreprises ayant adopté l'IA augmente davantage que dans les entreprises ne l'ayant pas adoptée, alors que ces deux groupes suivaient une tendance antérieure similaire (graphique 5). L'effet résulte principalement de la création de nouveaux emplois, plutôt qu'un maintien plus important d'emplois existants. On constate aussi qu'il n'y a pas d'effets différenciés sur les emplois occupés par des hommes par rapport à ceux occupés par des femmes : l'adoption de l'IA a des effets comparables pour l'emploi masculin et féminin.

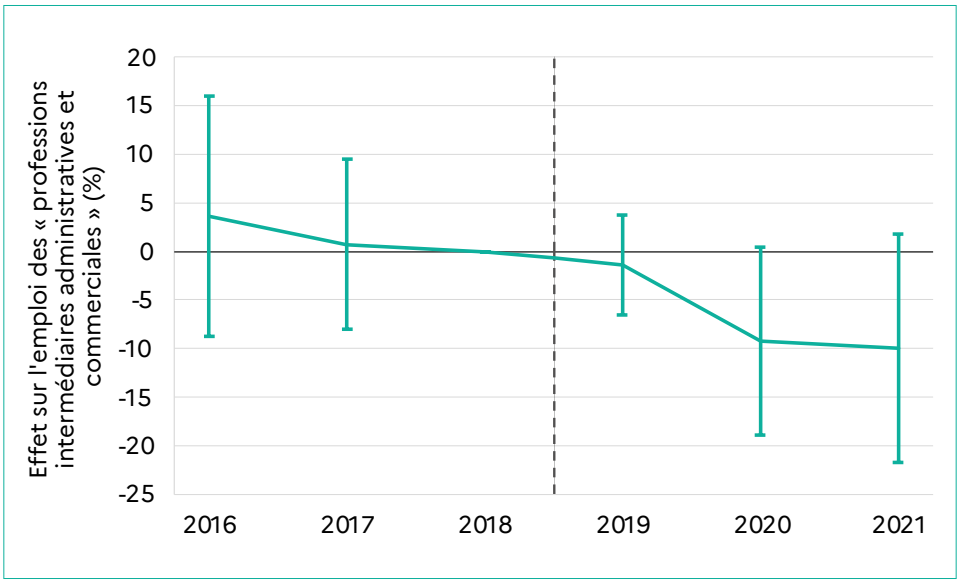


Graphique 5 : Effet de l'adoption de l'IA sur l'emploi total au sein des entreprises en France
Source : Commission IA.
Lecture : Les entreprises adoptant l'IA augmentent leur emploi davantage que celles ne l'adoptant pas, alors qu'elles évoluaient de façon similaire dans les 3 années précédentes.

Cependant, cet effet de l'IA sur l'emploi total n'est pas uniforme d'un métier à l'autre. En particulier, certains échelons au sein de l'entreprise ou certains métiers risquent de subir des réductions nettes d'emplois. De fait, le graphique 6 montre que les entreprises qui adoptent

26. Aghion, P., C. Antonin, S. Bunel, et, X. Jaravel (2023), « Capital industriel moderne, demande de travail et dynamique des marchés de produits : le cas de la France », *Document de travail Insee*
27. Parmi les entreprises n'ayant pas déjà adopté l'IA en 2018, on peut comparer l'évolution de l'emploi entre les 321 entreprises l'ayant fait entre 2018 et 2020 et les 897 entreprises ne l'ayant pas fait.

l'IA pour la gestion administrative ou le marketing voient leur emploi en « professions intermédiaires administratives et commerciales » diminuer.



Graphique 6 : Effet de l'adoption d'IA pour le marketing ou la gestion administrative sur l'emploi des professions intermédiaires administratives et commerciales au sein des entreprises en France.
Source : Commission IA.
Lecture : Les entreprises adoptant l'IA pour le marketing ou la gestion administrative baissent leur emploi davantage que celles ne l'adoptant pas, alors qu'elles évoluaient de façon similaire dans les 3 années précédentes.

Ces résultats sont nouveaux et importants, mais ils ne suffisent pas à appréhender l'ensemble des effets potentiels de l'IA sur le marché du travail. Il est notamment probable qu'une entreprise innovante qui adopte l'IA deviendra plus productive que les entreprises du même secteur qui ne l'adoptent pas, et que par suite l'entreprise innovante gagnera des parts de marché au détriment des entreprises concurrentes qui n'ont pas adopté l'IA. En termes d'emploi, cela se traduira par une création nette au sein de l'entreprise innovante mais au détriment de entreprises qui n'ont pas adopté l'IA : ces entreprises concurrentes risqueront de subir des destructions d'emplois. Au total, quel est l'effet agrégé de l'IA sur l'emploi lorsque l'on considère le marché du travail dans sa globalité ?

Un premier élément de réponse à cette question est fourni par des études récentes, focalisées sur les États-Unis et qui analysent les effets sur l'emploi de l'adoption de l'IA non générative par les entreprises. Ces études suggèrent un effet positif de l'IA sur l'emploi. En particulier, une étude américaine[28] publiée en 2023 suggère que l'adoption de l'IA est associée à une augmentation de l'emploi et du chiffre d'affaires au niveau sectoriel, pas juste au niveau de l'entreprise qui se convertit à l'IA.

Jusqu'à présent nous nous sommes concentrés sur les effets de l'IA non générative. Or, ce qui est nouveau avec l'IA générative, c'est que *certains* métiers de la connaissance, de la stratégie et de la créativité (médecins, enseignants, avocats, journalistes, artistes…), autrefois perçus comme des creusets de l'intelligence humaine, pourraient être concernés par une réduction du nombre total d'emplois.

28. Babina, T., A. Fedyk, A. He, et J. Hodson (2024), « Artificial Intelligence, Firm Growth, and Product Innovation », *Journal of Financial Economics.*

Des travaux récents, réalisés à un niveau très microéconomique, s'intéressent précisément aux effets de l'adoption de l'IA générative. En particulier l'étude de Brynjolfsson et al.[29] (voir *2.3. L'IA nous rendra-t-elle plus prospères*) se concentre sur le cas de l'adoption d'un outil à base d'IA générative permettant d'assister les agents du service clientèle. Si cette étude ne permet pas de conclure à un effet sur l'emploi total, elle met néanmoins en avant le fait que la probabilité qu'un employé quitte son emploi pendant le mois en cours diminue de 8,6 points de pourcentage après l'adoption de l'IA, tout en augmentant la productivité des employés. L'effet est encore plus fort pour les employés dont l'embauche au sein de l'entreprise est récente (moins de six mois). Au total, cette étude suggère un effet positif de l'IA générative sur l'emploi.

Cependant, une autre étude[30] montre que l'arrivée de ChatGPT a eu un effet négatif sur l'emploi et la rémunération des travailleurs américains *free-lance* (travail rémunéré à la tâche). L'étude s'intéresse plus précisément à une plateforme proposant des missions ciblant des tâches de petite ou moyenne envergure au sein de plusieurs métiers (saisie de données, conception graphique, développement de logiciels, marketing, etc.) et met en évidence une baisse du nombre d'emplois et des revenus, y compris pour les actifs ayant une expérience plus importante, suite à la mise en service de ChatGPT en novembre 2022.

En résumé, ces premières études suggèrent que l'effet de productivité domine en moyenne pour les salariés en entreprise, tandis que l'effet d'éviction semble être plus important pour les individus indépendants devant effectuer pour l'essentiel des tâches plus facilement remplaçables par l'IA.

Au-delà de l'effet sur l'emploi, qu'en est-il de l'effet sur les inégalités ? La première série d'étude se concentrant sur l'IA non générative[31] met en avant le fait que les entreprises qui adoptent l'IA embauchent par la suite des profils plus diplômés et plus techniques, en particulier en faveur des emplois dits « STIM » (science, technologie, ingénierie et mathématiques). Ceci conduirait à une hausse des inégalités, car ces profils complémentaires de l'adoption de l'IA ont des salaires plus élevés que la moyenne. Toutefois, la seconde vague d'étude sur l'IA générative[32] bat en brèche cette vision en mettant en évidence le fait que les employés initialement les moins qualifiés ou les moins productifs sont ceux pour lesquels les gains de productivité offerts par l'utilisation de l'IA sont les plus importants. Ceci pourrait alors donner à ces employés des leviers pour renégocier à la hausse leur rémunération, et ainsi réduire les inégalités au sein des entreprises.

Enfin, l'approche directe consistant à comparer *a posteriori* les entreprises ou secteurs ayant adopté l'IA par rapport à ceux l'ayant peu ou pas fait ne permet pas d'avoir un recul temporel très important, en particulier dans le cas de l'IA générative : la sortie du premier outil d'IA générative grand public date de novembre 2022, et ces études ne permettent pas encore de conclure à un effet sur un horizon de quelques années. En complément de l'approche directe, il est intéressant de s'appuyer sur une approche plus prospective par les tâches.

29. Brynjolfsson, E., D. Li, et L. Raymond (2023), « Generative AI at Work », NBER Working paper.
30. Hui, X., O. Reshef, et L. Zhou (2023), « The Short-Term Effects of Generative Artificial Intelligence on Employment: Evidence from an Online Labor Market », *Working paper.*
31. Babina, T., A. Fedyk, A. He, et J. Hodson (2023), « Firm Investments in Artificial Intelligence Technologies and Changes in Workforce Composition », *NBER Working paper.*
32. Brynjolfsson, E., D. Li, et L. Raymond (2023), op. cit. // Noy, S., et W. Zhang (2023), « Experimental Evidence on the Productivity Effects of Generative Artificial Intelligence », *Science.*

SECONDE APPROCHE : LES EFFETS DE L'ADOPTION DE L'IA SUR LES TÂCHES

Comme la précédente, cette approche conduit à prédire un effet positif de l'IA sur l'emploi. C'est ce qu'a notamment montré une étude récente de l'Organisation internationale du travail (OIT)[33]. Celle-ci donne à chacune des tâches réalisées au sein de l'économie une probabilité de remplacement par l'IA. L'étude distingue deux cas. Si une part importante des tâches qui constituent une profession peuvent être effectuées par l'IA, cette profession a un potentiel de *remplacement par l'IA*. En revanche, si une profession est composée de quelques tâches automatisables, mais d'une majorité de tâches difficiles à automatiser, elle a un potentiel *d'amélioration par l'IA* : l'automatisation de certaines tâches permet de libérer du temps pour d'autres. Par exemple, les employés de bureau feraient face à un risque de remplacement, tandis que les managers auraient un potentiel d'amélioration.

L'étude conclut que, dans l'ensemble du monde et y compris dans les pays développés, le nombre d'emplois ayant un potentiel *d'amélioration par l'IA* (13,4 %) est bien plus élevé que celui ayant un potentiel *de remplacement par l'IA* (5,1 %). L'étude note par ailleurs des inégalités de genre : 3,5 % des emplois principalement tenus par des femmes ont un potentiel de remplacement, contre 1,6 % des emplois principalement tenus par des hommes[34]. Une étude du Fonds monétaire international (FMI)[35] adoptant une méthodologie relativement proche aboutit à des chiffres plus élevés : 60 % des emplois seraient fortement exposés à l'IA, environ la moitié de ces emplois pouvant profiter d'un potentiel *d'amélioration par l'IA* tandis que l'autre moitié ferait face à un potentiel *de remplacement par l'IA*.

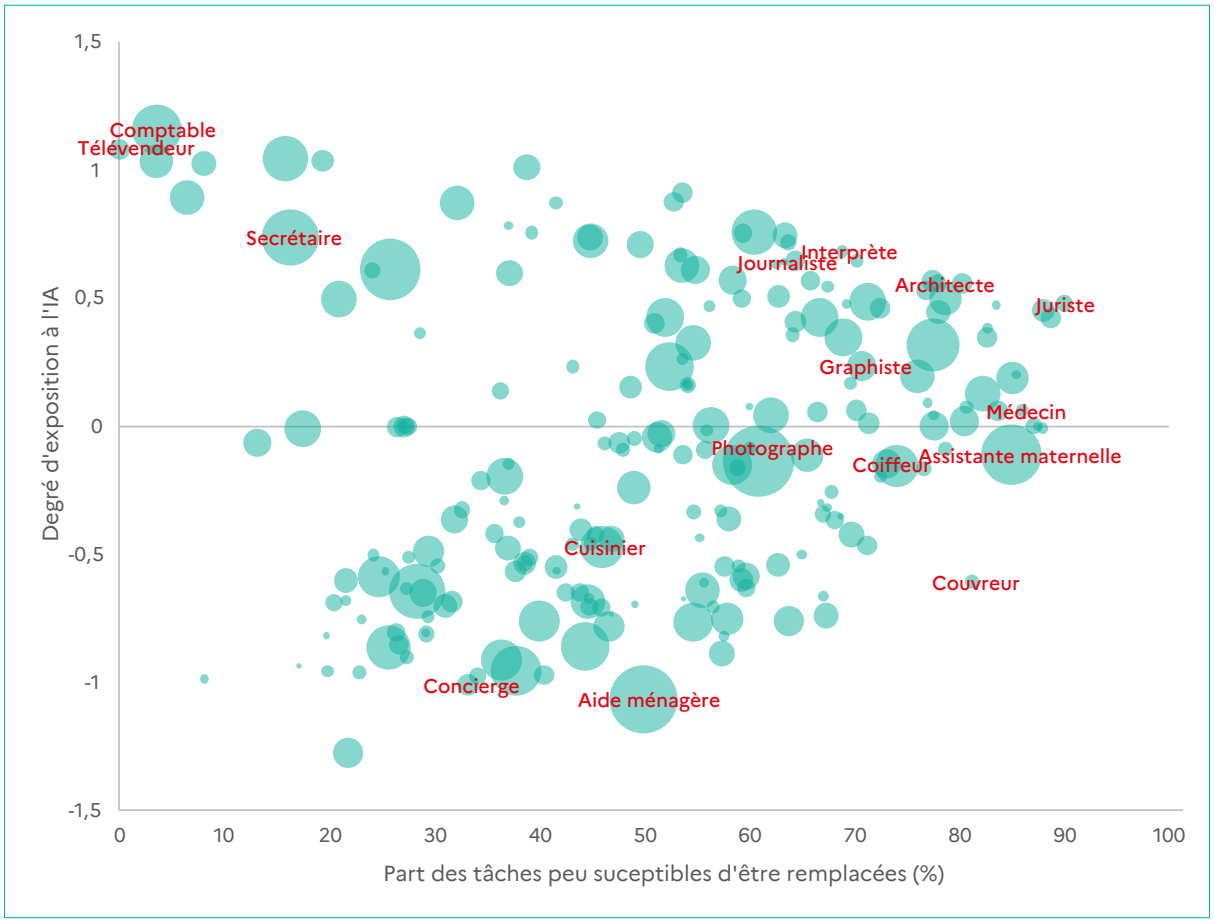
Concernant les inégalités de genre, l'étude note que dans les pays développés, les femmes sont plus exposées que les hommes à la fois pour les emplois ayant un potentiel *de remplacement par l'IA* mais aussi pour ceux ayant un potentiel *d'amélioration par l'IA*. Elle conclut que les femmes sont confrontées à la fois à de plus grands risques et à de plus grandes opportunités face à l'IA.

Une récente note[36] adapte la méthodologie de l'OIT au cas de la France afin de dresser un panorama des effets attendus de l'IA sur les 222 professions recensées par la nomenclature du ministère du travail. L'axe vertical (voir graphique 7 ci-après) représente l'exposition d'une profession à l'IA en général : plus un métier est haut sur ce graphique, plus il est globalement exposé à l'IA. L'axe horizontal représente quant à lui la part des tâches jugées difficiles à automatiser : plus une profession est située à droite sur ce graphique, plus une part importante des tâches ne semblent pas facilement remplaçables par l'IA. La taille des cercles est proportionnelle au nombre de personnes effectuant cette profession.

Les métiers localisés au haut à gauche sont les plus vulnérables à l'IA : à la fois fortement exposés et avec peu de tâches très difficiles à remplacer par l'IA. Ceux en haut à droite combinent une forte exposition à l'IA mais une part importante des tâches peu susceptibles d'être remplacées. On peut donc s'attendre à des transformations majeures de ces emplois : certains travailleurs plus formés pourraient être en mesure de tirer parti des nouvelles possibilités offertes par l'IA

33. Gmyrek, P., J. Berg, et D. Bescond (2023), « Generative AI and jobs: A global analysis of potential effects on job quantity and quality », *Document de travail OIT 96*
34. Les professions de service à la clientèle, ou ceux de tâches administratives et de communication sont par exemple des professions très féminisées.
35. Pizzinelli, C., A. Panton, M. Mendes Tavares, M. Cazzaniga, et L. Li (2023), « Labor Market Exposure to AI: Cross-country Differences and Distributional Implications », *Document de travail FMI*.
36. A. Bergeaud (2024), « Exposition à l'intelligence artificielle générative et emploi : une application à la classification socio-professionnelle française », *Document de travail*.

générative tandis que d'autres travailleurs pourraient faire face à une concurrence plus importante, augmentant ainsi les inégalités salariales au sein de ces professions. Les métiers se trouvant dans la partie inférieure du graphique, en particulier dans le quadrant inférieur droit, semblent globalement à l'écart de l'impact de l'intelligence artificielle.



Graphique 7 : Effet attendu de l'IA sur les métiers en France.
Source : Bergeaud (2024)
Lecture : Plus les métiers sont situés sur le haut du graphique, plus ils sont exposés à l'IA. Plus les métiers sont situés sur la droite du graphique, moins la part de leurs tâches susceptibles d'être remplacées est importante.

Cette approche par l'exposition des tâches à l'IA a l'avantage de permettre d'estimer des effets agrégés au niveau de l'économie dans son ensemble, et de permettre des comparaisons entre pays. Toutefois, elle recèle plusieurs limites. Citons les deux principales. D'une part, c'est une approche statique : les études s'appuient sur les tâches existantes et ne tiennent donc pas compte des tâches qui pourraient être créées à la suite du développement de l'IA. Pour faire un parallèle, il était difficile d'imaginer que le métier de *data scientist* puisse prendre une telle place au début de la révolution numérique dans les années 2000, ou que le métier de dépanneur en électroménager puisse exister avant l'adoption massive de l'électricité au milieu du XXe siècle. D'autre part, cette approche repose sur une estimation de la probabilité de remplacement par l'IA des différentes tâches (voir l'encadré).

Pourquoi une étude de 2013 sur les effets de l'automatisation se trompait-elle ?

Une étude[37] parue en 2013, avait fait grand bruit en concluant que 47 % des emplois américains étaient menacés par l'automatisation sur un horizon de dix ou vingt ans. Dix ans plus tard, force est de constater que cette prédiction était erronée.

Cette étude partait de l'idée subjective que 70 professions seraient entièrement automatisées, c'est-à-dire que l'ensemble des tâches effectuées par ces professions pourraient être effectuées par des machines. Puis, à partir des caractéristiques de ces professions (types de tâches effectuées, compétences requises) l'étude calculait, pour les 632 professions restantes, une probabilité d'automatisation comprise entre 0 et 1. En choisissant le seuil de 0,7 pour conclure qu'une profession était menacée par l'automatisation, les auteurs en déduisaient le chiffre de 47 % d'emplois menacés.

Plusieurs erreurs se sont glissées dans ce raisonnement. La première a été d'accorder une trop grande importance à la subjectivité, et de conclure un peu hâtivement à l'automatisation de certaines professions, comme les chauffeurs de poids lourds et de taxis, en pariant sur le développement rapide de la voiture autonome comme alternative parfaite aux chauffeurs. La seconde a été de confondre exposition à l'automatisation et risque de remplacement par l'automatisation. Les études précédemment mentionnées de l'OIT et du FMI évitent ces deux écueils.

On constate que les deux approches convergent vers des conclusions similaires : au total, le déploiement de l'IA dans l'économie devrait avoir un effet global positif sur le nombre d'emplois. Les prévisions catastrophistes sur la fin du travail ne sont pas plus crédibles que les prévisions du même type faites dans le passé. D'autant que même l'approche par tâche représente une borne haute pour l'impact de l'IA, puisqu'elle fait l'hypothèse qu'il est rentable d'automatiser toutes les tâches automatisables. Or cette hypothèse est loin d'être vraie aujourd'hui. La baisse du coût des systèmes d'IA ainsi que la possibilité de distribuer le même système IA à de très nombreux utilisateurs seront des facteurs clés pour déterminer l'impact de l'IA sur les tâches et les emplois.

Par ailleurs, cet effet global recouvrera des situations variées. Certaines professions pourraient connaître des réductions nettes d'emploi. Par exemple, dans le domaine de la culture et des médias, la part d'emplois exposés à l'IA est plus importante que d'autres secteurs. Pour le plus grand nombre de travailleurs, il s'agira d'évolutions des compétences et des tâches. Pour d'autres, une forte exposition à l'IA rime avec forte complémentarité avec l'IA, qui conduit à des spécialisations nouvelles, des formes d'expression ou de technicités nouvelles ou accrues. Cependant, il faut aussi s'attendre à ce que certains métiers disparaissent ou voient leurs effectifs nettement réduits. La complexité des situations et des métiers ne permet pas de fournir une réponse uniforme et globale aux défis de l'IA. Des études plus précises sont nécessaires pour aborder la variété des secteurs, des chaînes de valeur et prendre en considération les statuts. Au-delà, il est essentiel d'organiser l'accès et le développement de la formation initiale et tout au long de la vie.

37. Frey, C. B., et M. A. Osborne (2013), « The future of employment: How susceptible are jobs to computerisation? », *Document de travail Oxford Martin School.*

1.5 L'IA VA-T-ELLE DÉGRADER OU AMÉLIORER LA QUALITÉ DE VIE AU TRAVAIL ?

L'IA peut augmenter la qualité de vie au travail y compris pour des travailleurs de la classe moyenne. Certains utilisateurs de l'IA se déclarent effectivement plus épanouis et plus performants[38], car ils peuvent se débarrasser de tâches routinières et améliorer la qualité de leur travail. Cependant, des risques existent (surveillance, discrimination, intensification du stress, etc.). Les conséquences de l'IA sur la qualité de vie au travail dépendront de nos choix collectifs et de la qualité du dialogue social à son égard.

L'impact de l'IA sur la qualité de vie au travail peut être vu sous trois prismes : le contenu du travail, le risque de perdre son emploi, et les conditions de travail.

Concernant le contenu du travail, il est certain que la majorité des métiers va évoluer. La plupart des tâches seront transformées, d'autres seront supprimées, et de nouvelles tâches apparaîtront[39]. Ces transformations pourront concerner aussi bien des tâches relativement annexes, à faible valeur ajoutée (comme le fait pour un manager de rédiger une fiche de poste pour un recrutement), comme des tâches à forte valeur ajoutée, qui composent le cœur de métier.

Au sein d'une même profession, les gains de productivité semblent à ce jour bénéficier aux travailleurs les moins productifs. Par exemple, l'introduction d'une IA qui aide les chauffeurs de taxi à trouver des clients en leur suggérant des itinéraires sur lesquels la demande sera élevée permet d'accroître les recettes des chauffeurs les moins compétents, sans améliorer celle des plus expérimentés. Ce résultat se retrouve également dans le cas d'employés d'un service client[40] et de consultants d'un grand cabinet de conseil[41].

38. OCDE (2023) « Perspectives de l'emploi » : 63 % se déclarent plus épanouis et 80 % font état d'une amélioration de leurs résultats
39. Voir notamment l'étude de l'OIT et du FMI citées dans la partie 1.4.
40. Brynjolfsson E., D. Li, et L. Raymond (2023), « Generative AI at Work », *NBER Working paper.*
41. Dell'Aqua F., E. McFowland, E. Mollick, H. Lifshitz-Assaf, K. Kellogg, S. Rajendran, L. Krayer, F. Candelon, et Lakhani (2023), « Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality », *Working paper.*

D'une certaine manière, là où les machines industrielles et les programmes informatiques « classiques » sont un condensé de savoir explicite, un système d'IA peut être un condensé de savoir implicite. Les machines et programmes suivent en effet des règles bien définies et elles ont conduit à l'automatisation et la disparition de métiers qui consistaient à suivre précisément ces règles (verser telle quantité de matière quand la température atteint tel seuil, découper une planche en acier de façon régulière, etc.). L'IA peut détecter des régularités là où nous ne savons pas les expliciter. Elle peut dans certains cas donner accès à chaque travailleur à l'intuition du meilleur expert, même quand celui-ci serait incapable d'expliquer d'où vient son intuition et donc de la transmettre. L'IA peut donc à la fois venir soulager des tâches routinières et irrégulières, mais aussi enrichir le travail et améliorer la qualité du travail que tout un chacun peut réaliser en autonomie.

À l'inverse, un mauvais usage des systèmes d'IA peut également introduire une surcharge mentale qui peut mener à un épuisement cognitif si le temps libéré par la machine se traduit pour le travailleur par du stress et par une hausse excessive de tâches complexes. Ces risques appellent une réflexion sur la redistribution des gains permis par l'IA, pouvant notamment permettre une réduction du temps de travail et un meilleur équilibre vie privée et vie professionnelle.

Les risques justifient aussi l'établissement de règles de gouvernance dans l'interaction entre travailleurs et machines. Le projet de règlement sur l'intelligence artificielle (« *AI Act* ») identifie précisément parmi les applications à haut risque l'IA destinée au recrutement, notamment le filtrage et la sélection des candidatures, et prévoit des obligations strictes pour ces applications. Les systèmes d'IA à haut risque devront également être conçus de manière à ce que les personnes physiques puissent contrôler leur fonctionnement : c'est l'opérateur humain qui devra rester en charge de la décision finale.

Que se passe-t-il dans le reste du monde ? L'exemple des États-Unis[42]

Aux États-Unis, l'usage de l'IA par les employeurs est réglementé par diverses législations de niveau fédéré, qui encadrent l'utilisation de l'IA dans les processus de recrutement et visent à lutter contre les discriminations. Par exemple, la loi 144 à New York, entrée en vigueur en juillet 2023, exige la réalisation d'audits pour vérifier l'absence de biais dans les outils d'IA de recrutement et l'information des candidats quant à l'utilisation de tels outils.

Par ailleurs, on commence seulement à mieux comprendre les différentes interactions des travailleurs et du travail avec l'IA et les conditions augmentant les aptitudes et les compétences des travailleurs et améliorant les organisations de travail. Le déploiement des systèmes d'IA dans les organisations n'est pas le point d'aboutissement des processus d'innovation, mais plutôt un nouveau point de départ[43].

42. Source : direction générale du Trésor, service économique aux Etats-Unis.
43. LaborIA (2023) « Etude des impacts de l'IA sur le travail. Synthèse générale du rapport d'enquête du LaborIA Explorer ».

Concernant le fait de garder son travail, la diffusion de l'IA présente des atouts. Elle peut effectivement avoir un effet positif sur l'envie des travailleurs : les employés du service client et des professions qualifiées de services[44] (consultants, managers, etc.) qui utilisent de l'IA sont plus susceptibles de rester en poste, car l'IA augmente leur sentiment positif vis-à-vis de leur travail. Cependant, les consultants et les managers interrogés utilisant l'IA ont aussi une crainte plus élevée de perdre leur emploi. Ils voient eux-mêmes la technologie et les tâches qu'elle pourra réaliser à leur place. Il faut y ajouter la crainte de voir son entreprise disparaître si son secteur était bouleversé.

Concernant les conditions de travail (organisations de travail, pratiques managériales, relations de travail, rémunération, santé, sécurité), il fait peu de doute que le déploiement de l'IA conduira à l'émergence de nouvelles formes d'organisation et de coordination, comme les révolutions technologiques précédentes. La machine à vapeur a concentré l'industrie dans quelques usines et a conduit à rassembler un grand nombre de travailleurs sur un même lieu. Le moteur électrique a permis l'installation de nombreuses lignes et a facilité l'organisation du travail à la chaîne. Les technologies de communication ont facilité le télétravail et l'organisation de chaînes de production internationales, notamment par le biais de délocalisations d'activité.

Dans le domaine de la santé et de la sécurité au travail, les avancées de l'intelligence artificielle ouvrent des perspectives intéressantes en épidémiologie et en accidentologie et des possibilités nouvelles de supervision d'un environnement de travail, d'un chantier ou d'un site industriel par exemple, notamment par le recours à la maintenance prédictive[45].

En sens inverse, un mauvais usage des outils d'IA peut exacerber les risques professionnels et psychosociaux. Le développement du « management algorithmique » risquerait d'engendrer une perte d'autonomie au travail, une subordination déshumanisante à la machine, une surveillance excessive des travailleurs, un isolement des travailleurs et une perte du sens du collectif.

Une inconnue de taille reste l'impact que l'IA aura sur les salaires. Car les technologies ne font pas que transformer les métiers, elles modifient les pouvoirs de négociation et la valeur de certaines expertises, qui deviennent moins nécessaires ou moins rares. La révolution industrielle a facilité la production de certains biens et a ainsi réduit la rémunération des artisans. Elle a en revanche créé de nouveaux besoins, notamment pour tous ceux capables de faire fonctionner les machines et les organisations. Ces nouveaux emplois ont été au cœur de la « classe moyenne » des pays industrialisés. A leur tour, les technologies de l'information et de la communication ont bouleversé les conditions de travail. D'un côté elles ont facilité l'automatisation de certaines tâches routinières manuelles ou cognitives et ont ainsi menacé certains emplois de la classe moyenne. De l'autre elles ont facilité la circulation et le traitement de l'information, augmentant la centralité des personnes les plus qualifiées et concentrant le pouvoir de décision entre leurs mains. Les trente dernières années ont ainsi vue une polarisation des salaires et plus largement des conditions de travail, y compris pour les personnes titulaires d'un diplôme du supérieur.

L'IA va-t-elle prolonger la même dynamique ? Les premiers résultats permettent d'espérer le contraire, puisqu'ils montrent que l'IA générative augmente plus la productivité des personnes les moins qualifiées. Au-delà des résultats empiriques, c'est un espoir que l'on peut formuler[46] : en condensant l'intuition, le savoir implicite, l'IA pourrait aider les travailleurs ayant une certaine expérience et une formation de base à réaliser du travail à plus forte valeur ajoutée. Cette évolution permettrait de regarnir les rangs des emplois offrant de bonnes conditions de travail.

44. Noy S., et W. Zhang (2023), « Experimental Evidence on the Productivity Effects of Generative Artificial Intelligence », *Science*.
45. INRS (2022) « L'intelligence artificielle au service de la santé et de la sécurité au travail. Enjeux et perspectives à l'horizon 2035 ».
46. Autor, D. (2024), « Applying ai to rebuild middle class jobs », *NBER Working paper*.

Elle ne viendrait pas pour autant détruire la valeur de l'expertise. Après tout, ce n'est pas parce que chacun peut acheter des outils de plombiers et de chauffagiste que nous ne manquons pas de ces compétences-là. Un outil n'a de valeur que combinée à certaines compétences.

Les conséquences positives de l'IA pour la qualité de vie au travail sont un scénario, et non une prévision. Elles dépendront de la manière dont elle est introduite dans les organisations.

Pour orienter le déploiement de l'IA dans des directions positives, une série d'initiatives ont été récemment lancées : engagement de développement de l'emploi et des compétences à l'échelle régionale dans les Hauts-de-France (Cité de l'IA), et à l'échelle nationale (« Perspectives IA »), élaboration de guides d'accompagnement, d'outils d'autodiagnostic, etc. Mais de nombreuses administrations, branches d'activité et entreprises n'ont cependant pas encore pris le temps d'analyser les conséquences de l'IA sur les processus de production et leur organisation du travail. Les études menées par le LaborIA[47] ou d'autres équipes de recherche sont précieuses, mais encore trop peu fournies.

Recommandation n° 2

Investir dans l'observation, les études et la recherche sur les impacts des systèmes d'IA sur la quantité et la qualité de l'emploi.

Pour éviter certaines conséquences négatives, le cadre juridique définit un socle incontournable de droits (droit du travail, droit de la protection des données personnelles, etc.), qui apparaît pour le moment suffisant pour assurer un déploiement de l'IA respectueux des travailleurs. La Cnil a ainsi pu avec le cadre actuel condamner Amazon à une amende de 32 M€ pour « surveillance des salariés ». Le 23 janvier 2024, elle a considéré que mesurer chaque interruption de quelques secondes du scanner des salariés ou mesurer la vitesse d'utilisation du scanner lors du rangement était excessif, même au regard des enjeux de délais de livraison. Le même encadrement pourrait permettre d'empêcher tout « management algorithmique » abusif. Il s'agit donc surtout de veiller à son caractère effectif.

Le développement rapide des systèmes d'IA, notamment générative, dans les organisations du travail ainsi que le recours non encadré du « management algorithmique » devront être abordés par les partenaires sociaux et par un dialogue social spécifique dans les entreprises et dans les administrations. En outre, les services d'inspection du travail devront être modernisés et renforcés. La formation de ses agents, ainsi que celle des agents qui veillent à la prévention des risques en matière de santé et sécurité au travail, doit être organisée.

Recommandation n° 3

Faire du dialogue social et professionnel un outil de co-construction des usages et de régulation des risques des systèmes d'IA.

47. Initiative commune du ministère du travail, du plein emploi et de l'insertion et de l'INRIA créée en 2021

1.6 L'IA MET-ELLE EN DANGER LA CRÉATION ARTISTIQUE ?

Comme d'autres technologies avant elle, l'IA s'intègre dans les processus de création pour servir la création humaine. Elle fragilise cependant le secteur car elle constitue un immense défi pour les créateurs, leurs compétences, carrières et rémunérations.

De Hollywood à Paris, de l'Europe à l'Inde, les créateurs et leurs organisations sont inquiets. Les modèles d'IA générative font penser que la création artistique est en danger, tant leur adoption a été rapide pour la génération de texte, d'image et de vidéo. La création est placée face au défi de l'IA qui, de requête en requête, demande à un logiciel (Dall-E, MidJourney, Adobe Firefly…) d'affiner une création de texte ou d'images, de créer ou résumer un texte, voire d'écrire ou figurer « à la manière de… ».

Cette pulvérisation d'« IA créatives » semble dresser un immense défi pour les créateurs, comme deux décennies plus tôt la démocratisation des publications personnelles sur le Web à côté des médias dits traditionnels. Cette précédente mutation a rendu moins étanches les frontières de l'amateur au professionnel. Les IA, notamment génératives, abaissent elles aussi les barrières de la création artistiques. Mais elles font davantage que cela.

En termes d'emplois, de carrières, de revenus, l'apparition de l'IA n'est pas seulement une nouvelle concurrence. Pour bien des créateurs, auteurs, artistes, traducteurs ou comédiens, elle pose un enjeu existentiel, pouvant se traduire par une réduction d'activité, parfois par une substitution d'emplois. La baisse de revenus se profile, particulièrement pour les revenus complémentaires, ce qui pose la question des débuts de carrière dans la création. Même si le mouvement n'est pas général, c'est précisément pour cet enjeu social que les effets de l'IA sont débattus. Dès maintenant, sans formation appropriée, les débuts et les évolutions de carrière paraissent très délicats.

Par ailleurs, l'utilisation des créations humaines protégées pour l'entraînement des IA génératives, lesquelles sont ensuite susceptibles de produire des contenus concurrents, pose des questions d'autorisation et de juste rémunération. De nombreux procès ont été engagés aux États-Unis pour l'utilisation non autorisée (et donc non rémunérée) de contenus protégés par le droit d'auteur lors de l'entraînement des IA génératives.

L'IA ne met cependant pas en danger l'originalité de la création en elle-même et ses processus de sélection. La production et la capacité de production des systèmes d'IA ne constituent qu'un matériau nouveau, accompagné des évolutions logicielles devenues habituelles pour la musique ou l'image. La panoplie des techniques de création s'enrichit. Des artistes emploient déjà et parfois depuis longtemps des IA qui viennent stimuler réflexion et imagination, permettre des expressions qui n'étaient pas possibles sans elles. Certains (pour le moment peu nombreux) ont déjà choisi ces chemins de découverte : Robbie Barrat, Justine Emard, Gregory Chatonsky, Pierre Giner, Benoît Carré, etc. De ce point de vue, la formation sera essentielle, pour ne pas laisser de côté une part majoritaire du secteur créatif.

Dans plusieurs domaines — musique, architecture, multimédia… —, l'IA générative ne constitue pas une rupture. Elle s'ajoute à la longue histoire d'adoption de strates technologiques pour la composition, les arrangements, la conception, la production. Elle s'intègre dans des processus de création, qui passent de tous les univers possibles, par de multiples sélections et itérations, à une œuvre. Des usages d'IA peuvent accélérer des temps de création, ouvrir de nouveaux espaces, favoriser, nourrir, augmenter la créativité. D'ailleurs, le développement d'IA spécifiques, singulières et personnelles pour les artistes, pourrait permettre d'identifier et d'approfondir un style propre.

Pouvant réduire des tâches répétitives ou de faible valeur et abaissant les barrières à l'entrée, l'IA permet de se concentrer sur les éléments les plus essentiels, de l'imaginaire ou de sa sortie, où la part de l'humain, sa création, devient encore plus rare et plus distinctive dans un régime de « mégabondance » de productions par IA. S'ouvrent ainsi deux chantiers. D'une part la nécessaire reconnaissance de la création humaine, son identification, son caractère irréductible qui tient à son originalité et mérite distinction et protection. D'autre part, la juste rémunération de la création humaine, indispensable à une réception harmonieuse de l'IA dans le secteur culturel.

1.7 L'IA PEUT-ELLE NUIRE À LA QUALITÉ DE L'INFORMATION ?

L'information, au cœur de la démocratie, est chahutée par l'essor de l'IA. Des actions sont nécessaires pour préserver la confiance et la qualité de l'information.

L'information n'est pas un bien ou un service comme un autre. Elle est un élément essentiel des démocraties, suppose la liberté de communication, fabrique l'opinion, le jugement et la citoyenneté. Pour que l'espace informationnel soit de confiance, les démocraties ont conçu, établi et protégé des responsabilités majeures : des publications de presse (diffamation, injures…), des sources (présence et éthique journalistiques, fiabilité et vérification des informations…), du lectorat.

Or, l'IA a le potentiel de mettre en cause l'activité des entreprises de presse et par conséquent le rôle fondamental qu'elles jouent dans la production d'une information fiable et pluraliste. De nouveaux médias à base exclusive d'IA, sans respect de ces responsabilités, cherchent déjà à se positionner comme concurrents des médias « traditionnels ». Des sites fournissent des informations non fiables, générées par des IA et portant souvent des noms destinés à faire croire qu'il s'agit de contenus produits par des journalistes. S'ajoutent des robots plagiaires qui utilisent de manière irrégulière les contenus publiés par les médias traditionnels pour produire des articles, sans créditer leurs sources et sans rémunération.

Le risque démocratique tient à des enjeux techniques et de disponibilité des données de qualité : à mesure que la production d'information de qualité s'affaiblira dans un univers où les IA favorisent la multiplication de médias, il y a un risque que les modèles d'IA s'entraînent de manière croissante sur les *posts* de réseaux sociaux ou de sites d'information manipulée. Le risque est aujourd'hui perçu comme celui d'un cercle vicieux reposant sur l'éviction progressive de l'information produite et vérifiée par des professionnels au profit d'une information réalisée à faibles coûts sur de fausses informations.

Par ailleurs, la polarisation de la qualité de l'information est à l'œuvre. D'un côté, on observe que les plus gros acteurs de l'information (agences de presse, titres à notoriété mondiale, groupes de médias) peuvent passer des accords avec des fournisseurs d'IA, ou bien les poursuivent en justice pour les usages antérieurs non autorisés (en violation du *copyright*) à des fins d'entraînement d'IA (ex. Getty Images, New York Times…) parce qu'ils mesurent la valeur

économique de leurs contenus et de leur mission d'information. De l'autre, bien des médias ne peuvent pas installer des IA spécialisées sur leurs propres contenus. Le risque consiste donc à accentuer une information à deux vitesses : d'une part, une information de qualité, et payante, issue de grands groupes de médias ; et d'autre part, une information médiocre, parfois inexacte, fondée sur des modèles d'IA généralistes.

Enfin, l'IA renforce la propagation d'informations qui n'en sont pas et permet la création d'assistants personnels, favorisant la fragmentation de l'information et la création de « bulles » d'informations.

Face à de tels défis, les acteurs de l'information du monde entier se sont mis en marche : création de chartes (Reporters sans frontières), identification des images et des textes produits par l'IA (Google, Meta et d'autres), etc. De même, la lutte contre le parasitisme et la violation des droits, l'identification des informations non fiabilisées et la lutte contre les infox (*deepfakes*) suppose des investissements continus. Il faut aussi assurer une traçabilité de l'information, conformément aux responsabilités fondamentales de l'information dans une démocratie. Ces exigences portent leurs lots de difficulté, à commencer par l'interaction entre l'exigence de véracité et de traçabilité, puisqu'un contenu labellisé comme « généré par une IA » est automatiquement vu comme moins crédible[48]. C'est parce qu'elles sont difficiles qu'elles nécessitent un investissement collectif.

48. Tous les contenus générés par une IA ne sont pas faux et tous les contenus faux ne sont pas générés par une IA. Wittenberg, C., Z. Epstein, A. J. Berinsky, et D. G. Rand (2023) « Labeling AI-Generated Content » *MIT Topical Policy Brief*

1.8 FAUT-IL DIFFUSER L'IA DES AUTRES OU CRÉER LA NÔTRE ?

> **La France et l'Union européenne doivent se mobiliser pour diffuser les systèmes d'IA mais ne peuvent pas se contenter d'utiliser ceux développés à l'étranger.** Pour tirer les bénéfices de l'IA et en maîtriser les risques, nous devons être acteurs de la révolution technologique en cours.

Ce sont les Babyloniens qui se sont servi les premiers des tables de multiplication. Fallait-il utiliser leurs tables de multiplication ou créer les nôtres ? Cette question nous semble farfelue et personne en France ne songe à annoncer la création de tables de multiplication souveraines. En quoi l'IA serait différente ? Les systèmes d'IA sont loin d'être une technologie neutre comme peuvent l'être les tables de multiplication, qui sont les mêmes pour tout le monde.

D'abord, les modèles d'IA sont effectivement imprégnés des données qui servent à les entraîner, et des référentiels culturels présents dans ces données. Un Français vous dira que Clément Ader a réalisé le premier vol en avion, alors qu'un Américain répondra que ce sont les frères Wright. ChatGPT, pour sa part, répondra les frères Wright, qu'il soit interrogé en France ou aux États-Unis : 93 % des données d'entraînement de GPT-3 proviennent de textes en langue anglaise[49]. Les outils à base d'IA auront une influence croissante sur notre société, il est donc important de maîtriser leur référentiel culturel.

Ensuite, le déploiement des systèmes d'IA dans les secteurs sensibles (défense, énergie, recherche, etc.) pose une question de souveraineté. Connaître les données d'entraînement des modèles, leur fonctionnement, leurs fragilités et leurs atouts est un prérequis pour y déployer avec confiance l'IA.

Enfin, si toutes les administrations, les entreprises et les particuliers en France recourent à des outils étrangers, une partie croissante de notre richesse bénéficiera à ces fournisseurs, dégradant notamment notre balance commerciale. C'est ce qui s'est passé avec la précédente vague d'innovation technologique (ordinateurs, logiciels, internet) : notre balance commerciale dans

49. Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry G. and A. Askell, (2020) « Language models are few-shot learners » *Advances in Neural Information Processing Systems*

le numérique affiche un déficit d'au moins 22 Md€ (en 2019, année pré-Covid[50]), à mettre en regard du déficit commercial français global, toutes filières confondues, de 23 Md€ !

En sens inverse, c'est dans le déploiement très rapide des outils d'IA que nous tirerons beaucoup de valeur, pour nos services publics comme pour nos entreprises. La France ne peut pas attendre de disposer d'outils européens pour tirer les bénéfices de l'IA, en particulier le gain annuel de 1,5 point de productivité présenté précédemment (voir *1.3 L'IA nous rendra-t-elle plus prospères ?*). Il est même peu probable que notre continent puisse à moyen terme disposer d'une maîtrise complète de la technologie.

Au total, sans choisir une approche autarcique de l'IA qui serait délétère, la France et l'Europe doivent donc être des acteurs de la révolution technologique en cours et de sa chaîne de valeur. Autrement dit, ne nous fermons pas aux IA créées par d'autres, mais engageons-nous en même temps dans la création de nos propres maillons technologiques et modèles d'affaires, différenciés et compétitifs au niveau international. Nous devons avoir une vision dynamique : au fil des ans et du renforcement de notre expertise, notre approvisionnement pourra être de plus en plus européen.

50. Données provenant de la « Boîte à outils de l'OCDE sur la transformation numérique »

1.9 L'IA DEVRAIT-ELLE RESTER ENTRE LES MAINS DE QUELQUES ACTEURS ?

> **On ne peut pas se satisfaire que seules quelques entreprises maîtrisent l'IA.** Soutenir un écosystème ouvert de développeurs d'IA a des bénéfices de transparence, de pluralisme et de concurrence, sans causer de risques spécifiques.

Les progrès en IA sont le produit à la fois d'efforts massifs d'entreprises privées et d'une communauté de recherche très ouverte où chercheurs et développeurs contribuent ensemble aux briques qui permettent de construire ces systèmes. En ce qui concerne les modèles de fondation qui sont la base de l'IA générative, un article fondateur a été publié par une équipe de Google en 2017[51], tandis qu'OpenAI a été créée avec le but explicite de faire progresser la recherche ouverte en IA, mais ni l'une ni l'autre ne publie plus leurs modèles de façon ouverte. Pour leur part, Meta et Mistral ont fait le choix d'une ouverture partielle de leurs modèles. En Chine, Alibaba et 01.AI continuent de publier leurs modèles, tandis que Baidu a arrêté de publier son modèle Ernie[52].

Ce panorama révèle une tension au sein de l'écosystème : d'un côté l'ouverture permet d'élargir le cercle des contributeurs, de démocratiser l'IA et sa diffusion ; de l'autre, elle réduit le contrôle de la technologie et augmente le risque d'une utilisation inappropriée ou malicieuse. Certains plaident même pour un « registre » des systèmes d'IA afin de limiter leur « prolifération ».

Cette réflexion n'est pas nouvelle en informatique. Elle forme le cœur du débat autour de l'*open source*. Un logiciel *open source* est un logiciel dont le code est public, permettant à n'importe qui de l'utiliser, de l'inspecter, de le modifier ou de le partager. Des protocoles *open source* permettent de lire et recevoir des emails, et la plupart des serveurs de la planète ne sont pas équipés de Windows, mais du système d'opération *open source* Linux.

L'*open source* est donc présent dans beaucoup des outils numériques que nous utilisons. Outlook ou Gmail s'appuient sur des briques *open source*, mais ils ne sont pas eux-mêmes *open source*, à la fois pour des raisons économiques et pour assurer un certain degré de contrôle sur

51. Vaswani A. et al., Attention Is All You Need (2017), 31st *Conference on Neural Information Processing Systems.*
52. Baidu est le moteur de recherche chinois, Alibaba un leader du e-commerce et 01.AI une start-up fondée par Kai-Fu Lee.

l'expérience utilisateur. L'*open source* et les solutions propriétaires sont donc complémentaires à bien des égards.

Si le débat n'est donc pas récent, l'IA présente deux caractéristiques particulières. D'abord l'IA pourrait causer des risques importants, ce qui justifierait d'être plus précautionneux qu'avec une messagerie. Ensuite, l'ouverture d'un système d'IA n'est pas binaire, mais se fait le long d'un « gradient » d'ouverture.

Faut-il encourager le développement de modèles de fondation ouverts dans l'IA ? Pour y répondre, il faut regarder précisément les risques mentionnés, la comparaison avec les modèles fermés, et les bénéfices de l'ouverture. Il faut également garder en tête que certains des garde-fous intégrés dans les modèles fermés peuvent aujourd'hui être contournés[53].

Pour les risques biologiques et cyber, rien n'indique que les modèles ouverts posent plus de risques que des modèles fermés. Un modèle donne accès à des informations en virologie et bactériologie, mais il n'apparaît pas davantage faciliter la production d'une arme biologique qu'un simple moteur de recherche[54]. En revanche, les modèles ouverts réduisent de 70 % le coût de production de désinformation[55] et facilitent la création d'images non consenties. Ces capacités sont néanmoins déjà à la portée des modèles actuels, diffusés largement sur internet. Restreindre le développement de modèles ouverts ne diminuera donc pas le risque.

Dans l'ensemble, notre Commission considère que les modèles de fondation ouverts ne posent pas de risque supplémentaire significatif par rapport aux modèles fermés. En revanche ils appellent, comme les modèles fermés, à investir dans des contre-mesures, comme sur la cybersécurité et la désinformation.

Par ailleurs, il ne faut pas négliger les bénéfices des modèles ouverts pour la société, y compris en matière de gestion des risques. La transparence facilite l'évaluation par des tiers et permet de mobiliser la communauté pour améliorer et sécuriser les modèles. L'ouverture contribue aussi la diffusion de l'IA[56], puisqu'elle facilite la personnalisation des modèles, et donc leur adaptation à différents contextes. Elle réduit l'empreinte environnementale, puisqu'elle évite que chacun ne réentraîne son propre modèle de fondation. Enfin, l'ouverture abaisse les barrières à l'entrée et permet l'entrée de nouvelles entreprises. Les pays qui contribuent plus à l'*open source* créent plus de *start-up*, et elles sont de meilleure qualité[57]. Plusieurs grandes entreprises de technologie (MongoDB, Huggingface, Confluent, etc.) ont ainsi fait de l'*open source* le cœur de leur stratégie, tout en monétisant leur technologie par ailleurs.

Comment réaliser les bénéfices de l'ouverture des modèles d'IA ? Il faut s'assurer que l'écosystème puisse réellement utiliser, inspecter, modifier et partager des systèmes. Cela dépendra d'abord des modalités d'accès aux modèles, de leur degré de transparence et des capacités d'évaluation des tiers.

Pour qu'un système soit réutilisable et partageable, il faut pouvoir télécharger le code qui constitue le modèle, mais aussi les poids qui sont les paramètres résultant de l'entraînement, et d'une licence permissive[58]. Pour pouvoir inspecter le système ou le modifier à bon escient, le code du modèle ne suffit pas. Il faut disposer des outils pour le modifier, connaître aussi les

53. Henderson P. *et al.*, « Safety Risks from Customizing Foundation Models via Fine-Tuning » (2024), *Stanford HAL Policy Brief*
54. Mouton C. A., Caleb L. et E. Guest (2024) « The Operational Risks of AI in Large-Scale Biological Attacks », *RAND Research Report*
55. Musser A. (2023) « A Cost Analysis of Generative Language Models and Influence Operations », *Computers and Society*
56. Ding, J. (2023) « The Diffusion Deficit in Scientific and Technological Power: Re-assessing China's Rise », *Working paper*
57. Wright N. L. et al. (2023) « Open source software and global entrepreneurship », *Research Policy*
58. PaLM 2 de Google est ainsi fermé, et GPT-4 n'est accessible que de façon limitée *via* une API. Llama 2 de Meta est téléchargeable avec une licence qui autorise la réutilisation commerciale sauf dans les services ayant plus de 700 millions d'utilisateurs.

techniques et les données utilisées pour l'entraîner[59]. L'accès et la transparence dessinent ainsi un gradient d'ouverture des systèmes d'IA.

Être en faveur de l'ouverture ne signifie pas s'opposer aux modèles fermés, mais plutôt soutenir l'écosystème ouvert pour qu'il produise des briques et des systèmes d'IA de qualité, réutilisés du fait de leur qualité. Il est probable que demain cohabiteront systèmes d'IA ouverts, systèmes partiellement ouverts, systèmes fermés, systèmes d'IA ouverts intégrés dans des interfaces propriétaires, etc.

Ce dont l'écosystème d'IA ouverte a besoin, ce n'est pas d'un jardin protégé, mais d'une sécurité juridique (éviter que des contributeurs individuels soient responsables de l'utilisation de leur modèle par une entreprise), de denrées rares comme des données riches, et de certaines briques qu'aucune entreprise ne voudra publier ouvertement. Le soutien public devrait se concentrer sur des briques peu intenses en capital ou sans intérêt pour des acteurs privés comme l'affinage dans des langues rares.

Recommandation n° 4

Porter une stratégie de soutien à l'écosystème d'IA ouverte au niveau international en soutenant l'utilisation et le développement de systèmes d'IA ouverts et les capacités d'inspection et d'évaluation par des tiers.

59. Le *Stanford Transparency Index* liste ainsi 100 indicateurs de transparence des modèles d'IA.

1.10 L'IA MET-ELLE EN DANGER LA PLANÈTE ?

Tout dépend de l'utilisation que l'on fait de l'IA. L'utilisation des modèles implique une hausse de la consommation énergétique, même si elle est en partie compensée par les bonds dans leur efficacité énergétique. Cependant, les modèles d'IA peuvent accélérer l'innovation verte et ainsi lutter contre le réchauffement climatique.

La consommation énergétique nécessaire à l'entraînement des grands modèles de langue a retenu l'attention. En 2019, une étude a montré que les émissions de gaz à effet de serre liées à l'entraînement aux États-Unis d'un des premiers grands modèles de langue étaient du même ordre de grandeur que celles d'un vol entre New-York et San Francisco[60]. L'augmentation de la taille des modèles accroît rapidement ces émissions dès lors que l'approvisionnement énergétique est polluant. En sens inverse, les gains croissants d'efficience dans la microélectronique et l'entraînement des modèles réduisent ces émissions à performance égale[61].

L'empreinte carbone des modèles d'IA ne vient pas que de leur entraînement mais de l'ensemble de leur cycle de vie : la fabrication et le transport des équipements (20 à 30 % de l'empreinte) ; le développement et l'utilisation du modèle (50 à 60 % : entraînement et inférence) et les déchets (10 à 20 %). L'utilisation croissante de l'IA augmente par ailleurs son impact énergétique. Dans le cas d'un modèle comme ChatGPT utilisé par plus de 10 millions d'utilisateurs quotidiens, si la consommation énergétique de chaque requête reste relativement faible, la consommation énergétique totale de l'inférence dépasse celle de l'entraînement au bout de quelques semaines d'utilisation[62].

Au total, l'IA pourrait consommer 85 à 134 TWh d'électricité en 2027, soit une consommation équivalente à celle de l'Argentine ou de la Suède[63]. Ces chiffres, qui reposent sur une approche indirecte, doivent être considérés avec prudence. Surtout, la consommation énergétique représente un tel coût dans l'entraînement et l'inférence que tous les acteurs cherchent à améliorer leur efficience énergétique. En 2023, une requête Google augmentée à l'IA coûterait 10 fois plus en énergie qu'une requête classique. Il est difficile d'imaginer cette technologie

60. Strubell, E., Ganesh, A. & McCallum, A. (2019) « Energy and Policy Considerations for Deep Learning in NLP ».
61. Luccioni, S., (2023) « Towards Measuring and Mitigating the Environmental Impacts of Large Language Models ».
62. Luccioni, S. et al. (2023) « Power Hungry Processing: Watts Driving the Cost of AI Deployment? ».
63. A. Vries, (2023), "The growing energy footprint of artificial intelligence", *Joule*.

déployée ainsi aux 5 milliards d'utilisateurs de Google. Il est même certain que le coût de chaque requête sera un champ central de la compétition autour des systèmes d'IA. A ce titre, les modèles spécialisés consomment aujourd'hui moins d'énergie que les modèles à usage général, et de nombreuses pistes sont explorées, que ce soit pour améliorer les modèles, trouver de nouvelles architectures ou de nouveaux modèles. Augmenter la transparence sur l'impact énergétique des modèles d'IA permettrait aux utilisateurs de faire des choix éclairés et de peser au-delà des incitations économiques.

En ce qui concerne l'impact environnemental non énergétique, il vient principalement de la production des processeurs utilisés dans la puissance de calcul, et notamment de l'extraction et de l'utilisation d'eau, de silice et de terres rares dans la production. Le prix élevé des processeurs spécialisés pour l'IA ne doit toutefois pas faire oublier qu'ils constituent une part très faible des processeurs produits mondialement. En 2022, ils représentaient moins de 1 % des processeurs de moins de 7 nm, et moins de 0,00026 % de tous les processeurs produits[64].

L'impact environnemental de l'IA doit être mis en regard de ses potentiels bénéfices. Grâce à la capacité d'optimiser des processus complexes, l'IA pourrait permettre de réduire fortement les émissions de gaz à effet de serre dans de nombreux secteurs : énergie, transports, agriculture, logement… Par ailleurs, l'IA pourrait accélérer la transition écologique en diminuant la dépendance au sentier de l'innovation[65]. Une récente étude[66], réalisée à partir de données de brevets, met justement en évidence cet effet. Des cas concrets d'accélération de l'innovation apparaissent[67], mais doivent encore être soutenus.

Recommandation n° 5

Faire de la France un pionnier de l'IA pour la planète en renforçant la transparence environnementale, la recherche dans des modèles à faible impact, et l'utilisation de l'IA au service des transitions énergétique et environnementales.

64. Heim, L., and K. Pilz (2024). « What Share of All Chips Are High-End Data Center AI Chips? » *blog.heim.xyz*
65. Une entreprise qui a innové dans les technologies polluantes dans le passé va plus aisément continuer à innover dans ces technologies, car elle y a acquis des avantages (compétences de l'équipe de chercheurs, maîtrise de l'industrialisation, …).
66. Andres, P., E. Dugoua et M. Dumas (2022), « Directed Technological Change and General Purpose Technologies: Can AI Accelerate Clean Energy Innovation ? », Document de travail LSE
67. L'IA permettrait par exemple aux avions de réduire de 54 % les traînées qu'ils laissent dans leur sillage, elles-mêmes responsables de 35 % des émissions en volume. Google (2023), « How AI is helping airlines mitigate the climate impact of contrails »

1.11 Y A-T-IL UNE BULLE DANS L'IA GÉNÉRATIVE ?

Probablement, mais la bulle contribue à attirer des investissements sur des projets risqués. Lorsque la bulle éclatera, cela ne sera pas la preuve que l'IA générative ne sert à rien, mais qu'elle ne sert pas à tout. Il s'agira alors de maintenir le cap des investissements dans l'IA pour faire émerger un solide écosystème européen.

En matière de financement, 2023 a été une année exceptionnelle pour les *start-up* d'IA générative. Celles-ci ont levé 22 Md$ en 2023, contre 4,3 Md$ en 2022, 4,5 Md$ en 2021, et moins de 2 Md$ par an auparavant[68]. En d'autres termes, les deux tiers de tous les investissements reçus par les *start-up* de l'IA générative ont eu lieu en 2023 ! Parallèlement, l'investissement en capital-risque dans les *start-up* continuait de chuter, atteignant 224 Md$ en 2023, contre 655 Md$ en 2021.

Cet afflux de financement reflète bien l'effervescence de ce secteur, qui s'accompagne de son lot de *start-up* aux produits ou aux modèles d'affaires incertains. Le réflexe est d'y voir une « bulle », c'est-à-dire un emballement déraisonnable d'investisseurs et d'entrepreneurs guidés par le mimétisme.

Du point de vue financier, une bulle existe lorsque le prix (d'une action, d'une entreprise, d'un secteur) excède largement sa « valeur fondamentale ». Pour savoir s'il y a une bulle dans l'IA générative, il faudrait donc estimer la valeur future des entreprises et des produits qui sont financés aujourd'hui. Or, nous sommes au début du déploiement de l'IA générative. Il est probable que beaucoup des produits d'IA générative financés aujourd'hui ne seront ni utiles ni rentables.

Cette euphorie est toutefois normale et nécessaire au début d'une révolution technologique. Elle correspond à une « phase d'installation » d'une nouvelle technologie[69], celle où tout semble possible, où les bénéfices semblent sans limites, où les secteurs craignent d'être disruptés et où il y a beaucoup à apprendre. Cette phase permet de financer les premières innovations en dépit de l'incertitude.

68. Données Dealroom
69. Carlota Perez (2002), *Technological Revolutions and Financial Capital*

L'effervescence autour de l'IA générative a été comparée plusieurs fois en 2023 à une « ruée vers l'or ». Elle correspond bien plus à l'emballement qui a accompagné le développement du chemin de fer au XIXe siècle, ou au développement des usages en ligne durant la bulle internet.

Cette phase d'installation est souvent suivie d'un « tournant », à partir duquel les usages pertinents de la technologie seront plus clairs. La bulle éclate alors, les valorisations des entreprises diminuent, et l'investissement est redirigé vers les entreprises et les projets utiles et rentables. Nous verrons probablement dans les prochaines années quelques faillites spectaculaires dans le domaine de l'IA générative. Des entreprises très bien financées ne trouveront pas leur marché et perdront leur pari. En parallèle, d'autres entreprises réussiront leur pari et trouveront leur marché.

L'éclatement de la bulle ne sera pas la preuve que l'IA générative ne marche pas, mais que certains usages ne marchent pas, ou pas encore. Après l'éclatement de la bulle internet, les Européens se sont détournés du numérique. Les diplômés sont partis dans le secteur financier et les gouvernements n'ont pas fait du numérique une priorité, quand la *Silicon Valley* continuait d'investir dans le numérique. Il y a probablement une bulle dans l'IA générative, mais ne reproduisons pas la même erreur à son éclatement : tout en laissant les entreprises non viables disparaître, gardons le cap des investissements en IA.

1.12 DOIT-ON SE PRÉPARER À UNE IA PLUS INTELLIGENTE QUE NOUS ?

Oui, nous pouvons nous préparer à un avenir où les machines dépasseront les humains dans de très nombreux domaines. Cette supériorité arrivera progressivement et à un horizon temporel incertain. Il nous revient d'agir collectivement pour limiter les risques associés à cette évolution et en tirer tous les bénéfices pour l'humanité.

Avec l'intelligence artificielle générative, une étape importante dans l'histoire de l'innovation a été franchie. Cette étape est loin d'être la dernière. Dans les mois, années et décennies à venir, nous devrions connaître de nouvelles avancées rapides et de grande ampleur. Les modèles seront progressivement capables d'être factuels, de s'adapter aisément à des cas d'application de plus en plus pointus, de générer de la voix et des vidéos dans toutes les langues avec précision, de mener des raisonnements, de faire des mathématiques, de comprendre le monde physique autour de nous.

Ces évolutions seraient permises par la poursuite de la massification des données disponibles, sous l'effet notamment de la multiplication des capteurs sur les objets connectés et embarqués (voitures, robots…), mais aussi dans l'espace (constellations de satellites) et dans les océans. Elles seraient également accélérées par la baisse du coût de la puissance de calcul, contribuant à la fois à la généralisation de l'utilisation des systèmes d'IA et à la précision croissante des modèles. Elles seraient probablement facilitées par une combinaison d'approches technologiques, mêlant l'apprentissage automatique et l'approche symbolique de l'IA.

D'ici la fin de la décennie, il est vraisemblable que les systèmes d'IA accompagnent les humains en continu et dans toutes les tâches, personnelles ou professionnelles. Cet appui pourrait en particulier prendre la forme de puissants assistants personnalisés, qui accompliront des tâches rébarbatives, appuieront la réflexion et la prise de décision, et accéléreront le travail de groupe.

Nous pouvons également nous attendre à ce que la robotique fasse des bonds de géant. La complexité du monde réel en trois dimensions, les difficultés des interactions avec l'environnement et les inerties sociétales ne font néanmoins pas entrevoir une immédiate massification de l'usage des robots. La voiture à conduite automatisée, dont la démocratisation est sans cesse reportée, illustre la double difficulté de lever les verrous technologiques et de déployer effectivement les systèmes automatisés en situation réelle. Ce n'est pas parce que les

machines réalisent aujourd'hui des tâches qui nous semblent complexes qu'elles sauront demain faire aussi ce qui nous est simple. C'est tout le paradoxe souligné depuis 1988 par Hans Moravec : « le plus difficile en robotique est souvent ce qui est le plus facile pour l'homme ».

Au fur et à mesure des progrès technologiques, les machines dépasseront les humains dans des domaines sans cesse plus nombreux. Ce dépassement par la machine est parfois appelé intelligence artificielle *générale*. La notion est très débattue et aujourd'hui loin de faire consensus. Elle cherche à décrire des systèmes d'IA très performants, dotés de capacités à la fois étendues (d'où l'expression d'IA générale) et pointues[70]. Selon toute vraisemblance, il y aura une amélioration progressive des systèmes d'IA et la supériorité ne sera pas soudaine. Nos assistants personnalisés devraient être progressivement de plus en plus compétents. Leur utilisation sera de plus en plus aisée et discrète, par exemple par la disparition des interfaces que nous connaissons aujourd'hui (écran et clavier) au profit d'interfaces plus naturelles.

Les transformations sociétales que susciteront ces innovations dépendront de notre ambition et de notre engagement. L'IA peut être mise au profit de la réduction des inégalités sociales, de la prospérité collective, de la qualité du travail par la suppression des tâches les plus ingrates, du progrès scientifique au bénéfice de l'humain (de son intelligence, de sa santé, de son alimentation, de sa vie démocratique…) et de son environnement (optimisation des procédés industriels, nouvelles formes d'énergies, nouvelles technologies de décarbonation, etc.). Dans cette version de notre avenir, l'humain aurait progressivement accès – grâce à des outils numériques personnalisés et puissants – à un ensemble de connaissances, de biens et de services qui lui paraissaient jusqu'alors hors de portée.

Ces bénéfices ne seront pas spontanément obtenus. Seuls un projet politique et un engagement collectif permettront de les obtenir. Car, à l'inverse des gains présentés précédemment, un avenir dystopique peut aussi se dessiner. Un avenir où les bulles informationnelles et les influences cognitives affaiblissent notre démocratie. Un avenir où de nombreux travailleurs ne trouvent pas leur place vis-à-vis de machines toujours plus compétentes. Un avenir où la concentration des technologies les plus avancées entre les mains de quelques acteurs altère notre souveraineté et absorbe la majeure partie de la valeur produite par notre économie.

Le chemin que la France et que l'Europe emprunteront n'est pas encore tracé. C'est à nous qu'il revient de définir un projet politique de société et de forger les applications de l'IA qui y sont conformes. C'est à nous qu'il revient de tirer parti de l'IA, en investissant stratégiquement dans la maîtrise de la technologie et de sa chaîne de valeur. Le plan d'action recommandé par notre Commission s'attache à baliser les premiers kilomètres du chemin à parcourir. Aller au-delà nécessitera de la constance dans l'action, de la plasticité de nos institutions (publiques et privées), ainsi qu'un travail continu d'anticipation et de préparation.

70. Ringel Morris M., Sohl-dickstein J., Fiedel N. *et al.* (2024) « Levels of AGI: Operationalizing Progress on the Path to AGI ».

2

HUMANISME, SOUVERAINETÉ, RESPONSABILITÉ : INNOVONS, DÉPLOYONS ET MAÎTRISONS L'IA

2.1 HUMANISME : PLAÇONS L'IA À NOTRE SERVICE

2.1.1. FAIRE DU DIALOGUE SOCIAL ET DE LA CO-CONSTRUCTION LA PIERRE ANGULAIRE DU RECOURS À L'IA

La diffusion de l'IA est prise dans un face-à-face entre « techno-prudents » et « techno-enthousiastes » qu'il est facile de caricaturer. Les premiers craignent que l'IA ne renforce les inégalités, dégrade la qualité de vie au travail et ne profite, *in fine*, qu'à une minorité. Ils exigent donc de pouvoir définir comment les systèmes d'IA seront conçus et déployés. Les seconds disent avoir besoin d'aller vite, d'expérimenter pour trouver les bons usages, les bons produits, d'avoir le champ libre en somme[71].

À ce stade, deux choses sont certaines. D'une part, nous avons besoin d'expérimenter, de tâtonner avec l'IA afin de trouver la façon d'en tirer tous les bénéfices. Et ce d'autant plus si nous voulons trouver les « bonnes » façons de déployer l'IA, celle qui viendra améliorer le quotidien des travailleurs. D'autre part, cette expérimentation n'est pas qu'une question d'outils, mais aussi de formation et d'organisation du travail. Le filage du coton mécanique n'a augmenté la productivité dans le secteur textile qu'à partir du moment où le secteur s'est réorganisé autour de grandes usines textiles pouvant tirer parti de ces nouvelles machines[72]. L'électricité n'a permis d'augmenter la productivité dans les usines qu'à partir du moment où elles se sont réorganisées (voir *1.3 L'IA nous rendra-t-elle plus prospères ?*).

Pour tirer tout le potentiel de l'IA, nous devons trouver comment concilier le rythme rapide de l'expérimentation et le rythme nécessairement moins soutenu de nos compétences et de nos organisations. Le dialogue social est indispensable pour encourager l'usage de l'IA, pour discuter des finalités et du sens des transformations technologiques, pour développer la capacité d'apprentissage des organisations et pour concevoir des plans de formation adaptés. La participation de l'ensemble des parties prenantes est une condition incontournable du déploiement des nouvelles technologies dans une perspective d'émancipation, d'autonomisation et d'amélioration des conditions de travail, notamment par la réduction de tâches ingrates[73].

71. Le fondateur d'Uber T. Kalanick était allé jusqu'à en faire un mantra : « move fast and break things » (« aller vite et tout casser »).
72. Juhász, R., M. P. Squicciarini et N. Voigtländer (2020), « Technology Adoption and Productivity Growth: Evidence from Industrialization in France, » *National Bureau of Economic Research Working Papers*
73. Commission internationale Olivier Blanchard et Jean Tirole (2021), « Les grands défis économiques »

La diffusion de l'IA sera un atout clé pour renforcer la compétitivité et l'emploi des entreprises (voir *1.4 L'IA, créatrice ou destructrice d'emplois*). Elle ne se fera pas sans un dialogue social, basé sur la confiance réciproque, l'expérimentation et la co-construction. Pourtant, alors que les effets sur le monde du travail des précédentes vagues numériques sont profonds, les travailleurs et leurs représentants sont aujourd'hui peu associés aux choix technologiques et organisationnels sur les lieux de travail comme au plan national.

Le droit à l'information[74] et l'avis éclairé des représentants des salariés sur les transformations au travail sont en pratique peu mobilisés par les entreprises et les administrations. Cela tient à ce que l'intelligence artificielle et le numérique en général sont présentés comme des sujets techniques avant tout, que les travailleurs et les employeurs ont du mal à appréhender. Les acteurs sociaux ne sont pas assez informés ni formés à ces enjeux et à ces outils. Les directions des systèmes d'information (DSI), d'une part, et des ressources humaines et des relations sociales, d'autre part, agissent souvent en silos. La faiblesse de la co-construction peut devenir une source d'inquiétude voire de rejet pour les travailleurs et accroître le sentiment de précarité et la crainte du déclassement.

Que se passe-t-il dans le reste du monde ? L'exemple du Canada[75]

Au Canada, l'usage des systèmes d'IA par les employeurs est de plus en plus sensible et a conduit à des revendications dans le cadre de mouvements sociaux d'ampleur au cours de l'année 2023 : grève des fonctionnaires fédéraux, grève d'employés portuaires… En particulier, les dockers des ports de Colombie-Britannique et les employés de la chaîne de distribution Metro ont revendiqué un plus fort encadrement du recours à l'automatisation. Les salariés des trois grands constructeurs automobiles (Ford, General Motors, Stellantis) portent également, au-delà de leurs demandes salariales, des revendications en matière d'automatisation et de sous-traitance.

Le gouvernement fédéral a présenté en septembre 2023 un code de conduite volontaire visant le développement et une gestion responsable des systèmes d'IA générative avancés. Le code de conduite s'articule autour de six principes : la responsabilisation, la sécurité, la justice, la transparence, la surveillance humaine et la fiabilité. L'application de ce code est volontaire, plusieurs entreprises du monde l'ont déjà signé.

Pour que le dialogue social puisse intégrer les enjeux de l'IA, faciliter son expérimentation et sa diffusion, deux caractéristiques doivent être réunies. Premièrement, les partenaires sociaux doivent être des interlocuteurs formés[76] et actifs dans les instances où sera discuté le déploiement de l'IA. Secondement, ce dialogue social technologique doit s'insérer dans un processus itératif qui caractérise les projets d'IA.

74. Accords européens sur les transitions numériques signés par les partenaires sociaux en 2020 pour les entreprises et en 2022 pour les administrations.
75. Source : direction générale du Trésor, service économique au Canada.
76. Voir par exemple le projet « Dial IA », conduit par l'Institut de recherches économiques et sociales (Ires), qui vise à déployer une méthodologie faisant du dialogue social technologique au travail un levier opérationnel de la transformation numérique.

Pour aller plus loin, l'IA elle-même pourrait être mise au service du dialogue social. Des outils à base d'IA générative peuvent être développés avec les partenaires sociaux pour aider les salariés à mieux comprendre des débats techniques, qu'il s'agisse de technique informatique, financière ou juridique. Ces outils d'IA, par exemple sous la forme d'une instance simple de dialogue, pourraient intégrer une part commune à toutes les entreprises (code du travail, etc.) et une part spécifique à l'organisation dans laquelle est placé chaque travailleur (convention collective, règlement intérieur de l'organisation, etc.) et à ses représentants syndicaux. L'outil pourrait contribuer à accroître la connaissance des droits et la compréhension des transformations en cours ou encore à améliorer la préparation des réunions (conseils d'administration, comité social d'entreprise, élections des représentants…). Lors des négociations sociales, l'IA peut aussi contribuer à analyser et exploiter d'immenses quantités des données et ainsi appuyer la négociation.

2.1.2 FORMER : SANS DÉLAI, MASSIVEMENT ET EN CONTINU

2.1.2.1 La formation initiale

Les besoins actuels et à venir en IA nécessitent un vaste plan de formation pour tous et à tous les âges de la vie. Plus précisément, les enjeux de formation recouvrent trois besoins différents : former des personnes en mesure de concevoir et développer des solutions d'IA, former des personnes capables de déployer ces solutions d'IA au sein de leurs entreprises, et sensibiliser plus généralement l'ensemble de la population à la culture et la compréhension des grands principes de fonctionnement de l'IA.

Comme nous l'avons évoqué précédemment (voir *1.4. L'IA : créatrice ou destructrice d'emplois ?*), les entreprises investissant dans des compétences en IA embauchent davantage de profils plus diplômés et plus techniques, en particulier en faveur des emplois dits « STIM » (science, technologie, ingénierie et mathématiques). Cela concerne aussi bien les entreprises qui conçoivent les solutions d'IA que les entreprises qui les adoptent. Plus précisément, une étude[77] réalisée à partir de données danoises examine les besoins de formations adéquates pour la production d'IA d'un côté et l'adoption d'IA de l'autre. Elle montre que les entreprises productrices d'IA, qui vendent un produit ou service « contenant de l'IA », recrutent davantage des étudiants ayant des formations en informatique, mathématiques ou physique, tandis que les entreprises qui déploient l'IA sans la développer en interne ciblent des profils « STIM » plus appliqués ayant notamment des formations en chimie, biologie ou biotechnologie.

Du côté de la demande de compétences spécifiques à l'IA, une récente étude de l'OCDE[78] permet de fournir un ordre de grandeur des besoins. Cette étude constate que les offres d'emploi en ligne qui requièrent des compétences en IA représentent 0,35 % des offres postées en France. Parmi ces « offres d'emploi IA », 69 % d'entre elles concernent les secteurs de l'informatique et des activités spécialisées, et se rattachent donc majoritairement au développement de solutions d'IA, tandis que les 31 % restant se rattachent au déploiement de l'IA au sein des autres secteurs. Cette seconde catégorie recouvre les personnes ayant une formation permettant d'adapter l'IA aux usages spécifiques d'une discipline comme la santé, le

77. Humlum, A. et B. Meyer (2020) « Artificial Intelligence and College Majors", *Document de travail.*
78. Borgonovi, F. *et al.* (2023), « Emerging trends in AI skill demand across 14 OECD countries », *OECD Artificial Intelligence Papers*

droit ou la physique, ce qu'on nomme parfois les profils « X + IA ». Par ailleurs, le nombre total « d'offres d'emploi IA » est en progression d'environ 45 % entre 2019 et 2022.

Que peut-on en déduire pour la France ? Si l'on projette une évolution similaire pour les dix années à venir et que l'on suppose que la répartition entre développement et déploiement restera similaire, les offres d'emploi en développement d'IA et en déploiement d'IA devraient représenter devraient représenter respectivement 1 % et 0,5 % de l'ensemble des offres en 2034. En projetant une évolution similaire des besoins de mains d'œuvre estimés par Pôle Emploi pour les dix prochaines années à celle connue au cours des dix dernières, on aboutirait à un besoin de l'ordre de 56 000 postes par an en développement d'IA et 25 000 postes par an en déploiement d'IA (« X + IA »).

Il est donc nécessaire de calibrer l'offre de formation initiale aux besoins en compétence en IA d'aujourd'hui et de demain. En 2021, un rapport de la Cour des comptes estimait à 16 687 le nombre de places dans des formations spécialisées en IA au niveau bac+3. Pour répondre aux besoins en développement d'IA, il faudrait donc au moins tripler ce chiffre au cours de la décennie à venir.

Pour déployer l'IA dans l'économie, le besoin de 25 000 personnes par an en 2034, correspond à former chaque année environ 1,5 % de l'ensemble des étudiants du supérieur à des compétences « X + IA », soit par la création de filières spécifiques, soit par la création d'un module IA avancé au sein des différentes filières, comme le préconisait déjà le rapport Villani.

Un second type de profil, indirectement lié à l'IA semble nécessaire au déploiement de l'IA au sein des entreprises : celui de personnes chargées de l'infrastructure du système d'information disposant de connaissances indispensables à l'IA, notamment sur la collecte et le traitement des données, pour tirer le meilleur parti du déploiement des solutions d'IA au sein des entreprises. On utilise parfois le terme de « MLOps » pour décrire ce type de profils. En 2023, alors que Pôle Emploi indique un besoin de main d'œuvre de 12 180 postes dans les services informatiques qualifiés, 16 959 étudiants sont inscrits en cycle ingénieur dans le domaine « informatique et sciences informatiques ». Afin de répondre au besoin de formation, il serait donc nécessaire au minimum de former l'ensemble des étudiants des filières informatiques spécialisées aux enjeux de l'IA en lien avec leur activité. Étant donné que les besoins dans les services informatiques ne concernent pas tous l'IA, et qu'ils ne vont pas tous disparaître, il faudrait probablement viser une augmentation du nombre d'inscrits dans ces formations de 25%, pour atteindre 20 000 étudiants.

À tous les besoins spécifiques de formation de développement et de déploiement de l'IA s'ajoute le besoin de sensibilisation de tous, adultes et jeunes générations. Par ailleurs, la faible part des femmes au sein des diplômés des filières STIM (31 % en 2019), mais plus encore au sein de la filière « informatique et sciences informatiques » (19 % des inscrits en 2023) soulève de forts enjeux en termes d'emploi et d'inégalités de rémunération entre hommes et femmes en parallèle de la diffusion de l'IA. La sensibilisation en amont à tous les publics en est d'autant plus importante, afin d'attirer des étudiantes au sein de ces filières.

L'État a encouragé les établissements d'enseignement supérieur à développer des formations dans ce sens grâce à deux appels à manifestation d'intérêt (AMI). Les AMI « compétences et métiers d'avenir » et « IA Cluster » (ce dernier étant actuellement ouvert) visent à structurer la filière de formation de l'IA afin de consolider une dizaine de pôles d'excellences et de tripler le nombre d'étudiants formés à l'IA. Ces AMI incluent également une part de formation continue et de déploiement de modules d'initiation à l'IA permettant de toucher les étudiants de diverses disciplines. Ces investissements vont permettre une montée en puissance significative

du nombre de profils formés à l'IA dans les prochaines années, mais cela ne sera effectif que si les formations proposées sont attractives et trouvent leur public.

Les formations financées par les AMI présentent toutefois une limite : elles ne concernent pas tous les étudiants d'une génération notamment en raison d'une répartition qui ne couvre pas tout le territoire et d'un manque de formateurs spécialisés en IA. Une évaluation à trois ans des résultats de ces AMI sera nécessaire pour vérifier la réalité de la montée en puissance. Pour autant, il apparaît dès à présent essentiel de poursuivre les efforts engagés pour toucher un public le plus large possible. Des pistes peuvent être envisagées via notamment le partage de cours en ligne ou encore par un plan de formation ambitieux d'enseignants-chercheurs de toutes les disciplines. La Commission recommande par ailleurs de s'appuyer plus largement sur les professionnels du secteur, probablement prêts à participer à l'effort de formation, à condition qu'il fasse une part importante à la pratique, cruciale dans ce secteur.

En outre, pour s'assurer que les formations proposées attirent suffisamment d'étudiants, il est nécessaire d'agir en amont de l'enseignement supérieur pour acculturer les élèves aux enjeux de l'IA au fur et à mesure de leur scolarité. Certains programmes scolaires incluent déjà des apprentissages explicites de l'IA : le projet de programme de technologie de collège, l'enseignement scientifique du tronc commun du bac général ou encore le programme de la spécialité sciences de l'ingénieur du baccalauréat général. Toutefois, ces éléments d'apprentissage ne forment pas un cursus progressif d'apprentissage touchant tous les élèves. Une réflexion doit être menée pour identifier la contribution de chaque discipline enseignée et permettre à tous les élèves (qu'ils soient de la voie professionnelle, technologique ou générale) de bénéficier d'un apprentissage en lien avec l'IA.

Recommandation n° 6

Généraliser le déploiement de l'IA dans toutes les formations d'enseignement supérieur et acculturer les élèves dans l'enseignement secondaire pour rendre accessibles et attractives les formations spécialisées.

2.1.2.2 La formation continue

En complément de la formation initiale et de l'enseignement supérieur, mais aussi des politiques de recrutement et de réorganisation du travail, la formation professionnelle continue sera un outil indispensable pour faire face à la transition profonde des métiers que l'intelligence artificielle impliquera. Ce mouvement est renforcé par l'IA générative mais n'est pas nouveau. À titre d'illustration, le groupe La Poste a démarré en 2015 la formation au numérique et l'a inscrite dans le pacte social de l'entreprise. Un budget spécifique de formation (500 M€ sur 5 ans) a été établi et le catalogue de formation s'est progressivement enrichi sur l'IA. Le nouveau plan stratégique 2020-2030 prolonge cette action. Le travail paritaire doit être engagé afin d'anticiper les transformations des métiers et les évolutions du contenu des métiers, mais aussi d'identifier les formations nécessaires.

L'enseignement majeur de notre consultation citoyenne : le besoin de formation

De mi-décembre 2023 à mi-janvier 2024, notre Commission a mené une consultation citoyenne en ligne, afin de mieux comprendre les attentes et les craintes en matière d'IA. Nous avons recueilli 6 917 réponses[79]. Parmi les résultats clés de l'enquête, apparaît en particulier le besoin d'information et de formation à l'IA dans l'environnement professionnel.

La majorité des participants ne craint pas de voir leur emploi disparaître ou dévalorisé par l'émergence de l'IA (notons qu'une part plus importante de femmes, d'ouvriers et de professions intermédiaires partage cette inquiétude). Cependant, les répondants ont majoritairement indiqué des difficultés à s'exprimer sur les effets tangibles de l'IA dans leur métier. Ils sont ainsi nombreux à avoir exprimé leur besoin de mieux comprendre les bénéfices qu'ils pourraient tirer de l'IA et les cas d'application concrète pour utiliser ces outils au travail. Après la formation, les répondants ont souligné l'importance du dialogue social.

L'IA semble également promise à un rôle important, voire prépondérant, dans le secteur de la formation professionnelle continue. En matière de conception pédagogique d'abord, afin de structurer des contenus et d'organiser des idées pour aller vers un plan de formation ou aider à concevoir des outils pédagogiques. Pendant la formation où l'intelligence artificielle pourra permettre, comme assistant virtuel au service des apprenants, d'individualiser les parcours. Pour l'accompagnement et le tutorat pédagogique post-formation enfin. Par exemple, certaines formations d'ingénieur logiciel intègrent dès aujourd'hui les modèles de langue pour apprendre à coder.

Pour l'instant, selon l'OCDE, les offres d'emploi qui requièrent des compétences en matière d'intelligence artificielle ne représentent qu'un nombre limité de toutes les offres d'emploi (OCDE, 14 pays), dont moins de 1 % aux États-Unis en 2022, le pays où cette proportion est la plus élevée. Néanmoins, leur croissance est très rapide dans tous les pays ou presque. L'OCDE

79. La consultation a été conduite sur l'application Agora avec l'appui du centre interministériel de la participation citoyenne. Les profils des participants sont diversifiés mais ne peuvent pas être considérés comme pleinement représentatifs de la population française (par exemple, les hommes ont été plus nombreux à répondre que les femmes).

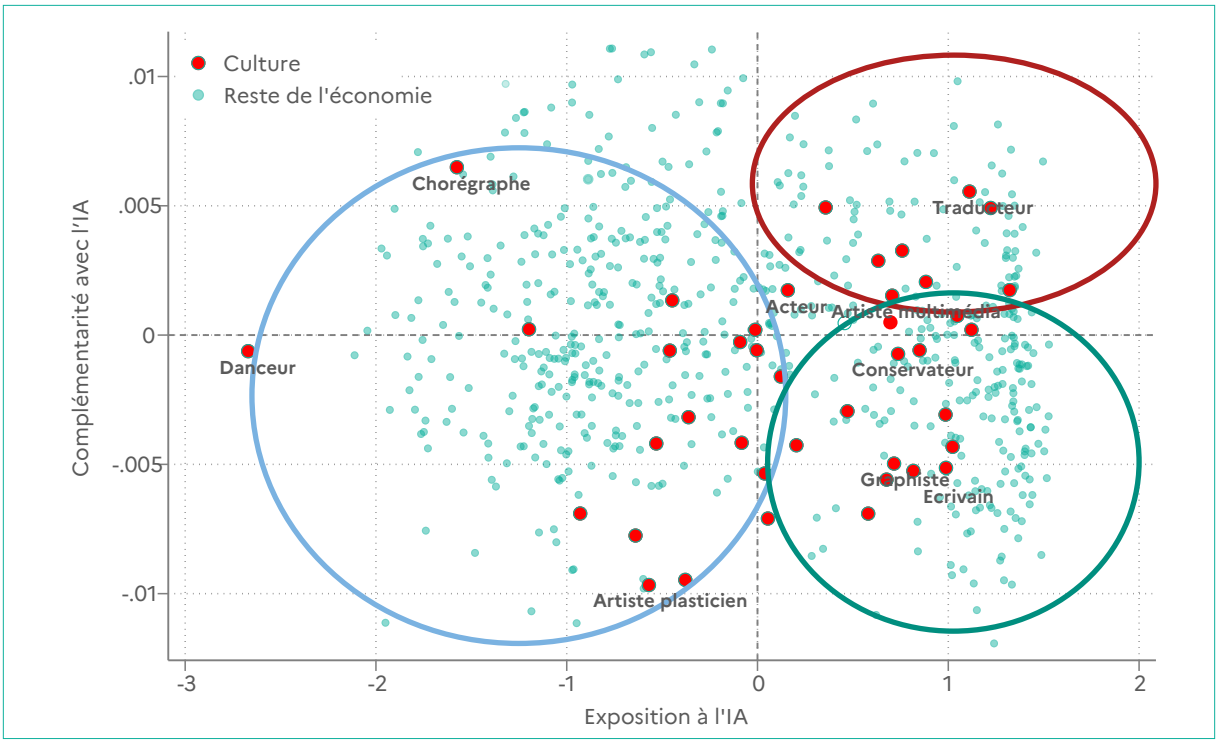
souligne aussi que la demande d'emplois liés à l'intelligence artificielle est très concentrée en termes de secteurs et de métiers, avec des différences entre l'Europe et les pays anglo-saxons.

Tout en étant volontariste, la puissance publique doit faire preuve d'humilité dans ce domaine. Les entreprises ont *a priori* intérêt à former leurs salariés aux nouvelles compétences. Si elles ne le font pas, ou pas encore, c'est aussi que le flou règne encore sur les enjeux de compétences et de formation professionnelle continue pour les métiers transformés par l'introduction de l'IA. Ce flou ne se lèvera que progressivement. Là où le soutien public sera utile, c'est pour appuyer la cartographie continue des besoins et des modalités de formations, pour apporter de la lisibilité dans une offre pléthorique, et pour assurer la formation des demandeurs d'emplois.

Recommandation n° 7

Investir dans la formation professionnelle continue des actifs et les dispositifs de formation autour de l'IA.

Dans le secteur de la culture et des médias, l'impact de l'IA sur l'emploi (4 % de l'emploi total) est déjà analysé comme important, d'autant que 40 % des emplois artistiques et culturels relèvent du reste de l'économie (luxe, publicité, design industriel, automobile, etc.). Le niveau d'exposition à l'IA est globalement plus élevé que pour l'ensemble des professions, en particulier pour ces emplois artistiques et culturels hors des secteurs culturels.



Graphique 8 : Exposition et complémentarité des professions artistiques à l'IA.
Source : calculs d'A. Bergeaud pour la Commission IA (2024).
Lecture : Plus les métiers sont situés sur la droite du graphique, plus ils sont composés de tâches susceptibles d'être automatisées par de l'IA. Plus les métiers sont situés sur le haut du graphique, plus ils sont complémentaires de l'IA, car ils mélangent tâches automatisables et tâches non automatisables.

Les effets de substitution existent déjà (traducteurs, doubleurs) avec des formes de déclassements (pertes du statut d'auteur, mutations d'activités à la baisse de compétences et risque de plus faibles revenus). Ils devraient s'étendre — à une échéance mal connue — à de nombreux emplois où l'automatisation est possible : professions artistiques dans les secteurs non culturels. Dans les secteurs artistiques et culturels, la complémentarité peut jouer un rôle plus important avec un moindre risque de substitution : activités « augmentées » grâce à l'IA et temps dégagé pour des activités plus créatives et distinctives. C'est selon la complémentarité à l'IA que peut se former une polarisation des professions, dans l'univers déjà concurrentiel des talents.

Pour une large part, les distinctions de carrières et de stabilité d'emploi devraient dépendre des statuts (salariés/indépendants ; privé/public) et de la taille des employeurs. S'y déterminent les conditions d'accès à la formation professionnelle pour accompagner les carrières. L'enjeu est immédiat. À plus long terme se jouent les parcours des carrières qui se font selon des durées spécifiques et des statuts variés, où la complémentarité des emplois est essentielle. Déjà, des métiers y sont confrontés (pigistes et *freelancers* graphistes)[80]. Ils sont dans des domaines où l'IA se déploie vite, comme la presse[81] ou la communication, où des tâches sont substituables par les systèmes d'IA qui les réaliseront à meilleur coût.

Favoriser l'adaptation de l'emploi artistique et culturel suppose de réunir très vite les conditions de formation, en particulier des jeunes générations pour faire de la France et de ses créateurs un lieu d'excellence. Dans un univers avec l'IA, la création originale, la spécialité culturelle, ont toute chance de devenir une ressource rare au moment du développement de « petits » modèles ou de l'accès généralisé aux « grands modèles ». Il convient de renforcer la formation initiale professionnalisante dans l'enseignement supérieur spécialisé et la recherche en favorisant les passerelles entre création, recherche, technologie, projets économiques et sociaux et culturels.

La reconfiguration de l'économie dans son ensemble par l'introduction dans un régime technologique dominé par l'IA place la création dans une position cruciale. Déjà très concurrentielle et très recherchée, la création va devoir se placer à la croisée de mutations technologiques rapides et d'évolution des modèles économiques.

Recommandation n° 8

Former les professions créatives à l'IA, dès les premières années de l'enseignement supérieur et en continu.

80. Reshef, O., X. Hui, et L. Zhou (2023) « The Short-Term Effects of Generative Artificial Intelligence on Employment: Evidence from an Online Labor Market », *cesifo Working papers*

81. Beckett, C., et M. Yaseen « Generating Change A global survey of what news organisations are doing with AI », *LSE JournalismAI*.

2.1.3 ÉQUIPER LES AGENTS PUBLICS : UNE OPPORTUNITÉ POUR TRANSFORMER L'ADMINISTRATION

Le numérique volant au secours du service public n'est pas une figure nouvelle. Si la dématérialisation des démarches en ligne a bien progressé, la transformation numérique reste incomplète. Le déploiement de l'IA est une opportunité pour relancer cette transformation, à condition de lever les freins qui continuent d'empêcher une réelle transformation numérique du service public. Car ce sont les mêmes freins qui empêcheront de tirer parti de l'IA et de parier sur un foisonnement d'initiatives.

Le service public devrait être un des premiers bénéficiaires du numérique. De fait, du point de vue du citoyen, la dématérialisation des services publics et de leurs démarches a largement progressé ces dernières années. Plus des trois quarts des 250 procédures les plus utilisées sont dématérialisés[82], et 79 % des Français pensent que la dématérialisation des services publics leur simplifie la vie[83]. Mais elle s'est accompagnée pour certains d'un sentiment de déshumanisation et d'éloignement du service public[84]. Pour les agents publics, cette dématérialisation a des effets ambigus, notamment car ils sont encore 51 % à trouver leur environnement numérique moyen (32 %), mauvais (14 %) ou médiocre (5 %)[85].

Trop souvent, la transformation numérique s'est arrêtée à la dématérialisation des démarches, sans transformer en profondeur la circulation de l'information, ou le traitement des demandes. Les promesses de personnalisation (et donc d'humanisation) du service public, de rapidité de traitement, de simplification du travail des agents n'ont pas été tenues. Les démarches à réaliser à la naissance d'un enfant sont nombreuses et complexes, et il reste frustrant de devoir apporter à chaque service public les pièces justifiant qu'un enfant est né, et qu'il est bien le vôtre. Certes, des initiatives prometteuses sont en cours (« Dites-le nous une fois », « administration proactive », etc.), mais la frustration est d'autant plus grande qu'en comparaison nos vies numériques sont de plus en plus intégrées.

L'IA est l'occasion pour les services publics d'aller plus loin dans leur transformation. Elle promet en effet de personnaliser le service public, de le rendre plus efficient, et l'IA générative promet de fluidifier la communication avec les utilisateurs.

Une IA générative pourrait bientôt réexpliquer plusieurs fois dans un langage accessible quelles sont les démarches à faire pour inscrire son enfant dans une école, ou pour remplir ses déclarations d'impôts. Un agent pourrait même les réaliser pour vous. Au rectorat de l'académie de Lyon, l'IA Cassandre propose depuis septembre 2023 aux 40 gestionnaires des ressources humaines de la direction des personnels enseignants des réponses aux questions posées par les 45 183 enseignants de l'académie sur l'affectation des stagiaires et les mutations à l'intérieur de l'académie. Les enquêtes montrent une très large appréciation positive de l'outil, dont l'utilisation sera étendue en 2024.

82. Observatoire des démarches en ligne.
83. Étude en ligne réalisée par Ipsos pour Sopra Steria en septembre 2019 auprès de 6 000 personnes dans 6 pays, représentatives de la population nationale âgée de 18 ans ou plus.
84. Défenseur des Droits (2019), « Dématérialisation et inégalités d'accès aux services publics ».
85. Direction interministérielle du numérique (2021) « Baromètre numérique de l'agent ».

Que se passe-t-il dans le reste du monde ? L'exemple de la Suède

La plateforme *Platsbanken.se*, le plus grand site d'offres d'emploi en ligne de Suède, et géré par l'équivalent suédois de France travail, a évalué la mise en place grâce à l'IA de recommandations d'emploi ciblées, basées sur l'historique des recherches d'emplois des demandeurs[86]. Les demandeurs d'emploi ayant accès aux recommandations cliquent et postulent davantage aux offres recommandées, conduisant à un effet positif sur le taux de réemploi d'environ 0,6 %. Bien qu'il soit difficile d'extrapoler et de transposer, une augmentation similaire en France conduirait à l'embauche d'un peu plus de 12 000 demandeurs d'emploi. Mais le principal gain réside probablement dans le temps libéré pour les agents, qui peuvent renforcer leur accompagnement sur d'autres axes.

Ainsi, l'IA et singulièrement l'IA générative peut libérer les agents de tâches répétitives, ou chronophages, tout en améliorant la qualité de service. Par sa simplicité d'utilisation, l'IA générative offre l'occasion de libérer la créativité des agents en permettant d'expérimenter la technologie à leur niveau sans avoir toujours besoin d'un système spécifique. Une enseignante peut d'ores et déjà utiliser une IA générative pour l'aider à concevoir un déroulé de formation ou faire varier un exercice.

Les services publics peuvent se fixer deux niveaux d'objectifs vis-à-vis de l'IA. Le premier consiste à déployer des systèmes d'IA remplissant un objectif précis : répondre à des usagers, simplifier un message, résumer une vidéoconférence, faire une analyse financière, etc. Le deuxième vise à repenser le service public à partir de ses missions, des besoins des usagers et des capacités offertes par l'IA. Par exemple, en imaginant un compagnon nous assistant dans toutes nos démarches pour notre enfant.

Le premier niveau correspond à la dématérialisation dans la transformation numérique de ces 20 dernières années. Le second à une profonde transformation de nos services publics.

Pour réussir le déploiement de l'IA dans les services publics, le premier niveau, les services publics auront besoin des mêmes ingrédients que pour la transformation numérique :

- **une vision claire des objectifs du service**, permettant d'évaluer l'apport de l'IA, d'ajuster de façon continue les systèmes face aux retours d'expérience, de gérer les risques et de répartir les rôles entre développements publics et rôle des acteurs privés ;
- **la confiance des agents et des usagers pour expérimenter**, ce qui nécessitera de prendre en charge la responsabilité sans se dédouaner sur l'outil ;
- **des personnes** capables de concevoir, piloter, produire ou acheter ces systèmes d'IA ;
- **des données**, pour entraîner ou réentraîner un modèle, mais aussi plus simplement pour déployer les outils à base d'IA dans le processus de travail ;
- **des infrastructures robustes**, facilitant la circulation des données, des référentiels communs, la sécurité des applications et leur mise à jour régulière.

Des efforts ont été engagés ces dernières années, notamment pour renforcer les compétences numériques de l'État et investir dans une infrastructure *cloud*. Ils doivent être poursuivis, car ils

86. Le Barbanchon, T., L. Hensvik, et R. Rathelot (2023), « Experimental Evidence on the Productivity Effects of Generative Artificial Intelligence », *Working paper.*

ne montreront leur impact que dans la durée. La fragmentation des infrastructures d'hébergement numériques publiques coûte cher, réduit la flexibilité en cas de pic de charge, empêche d'accéder à certains outils de développement, de test, d'intégration, réduit l'attractivité du service public pour les talents techniques et rend plus difficile l'amélioration des applications. Et point crucial pour l'IA, elle freine la collecte et la circulation des données. La transformation du service public par l'IA n'avancera pas sans un saut dans les efforts d'investissement numériques publics, à la fois en qualité et en quantité[87].

Ces ingrédients ne dessinent toutefois pas une stratégie, qui devra éviter deux écueils. D'une part le « grand projet IA », destiné à tout faire, tout remplacer, développé loin des agents, des usagers et de la réalité du service public. D'autre part le « tout ChatGPT », dans lequel un robot conversationnel universel commercial et étranger deviendrait la seule utilisation de l'IA dans le service public.

Entre ces deux écueils, les services publics doivent tracer un chemin combinant maîtrise technologique, maîtrise des coûts, foisonnement des expérimentations et transformation profonde. La Commission propose quatre axes : (i) clarifier les objectifs et la répartition des rôles entre services publics et fournisseurs privés ; (ii) mutualiser des investissements dans l'IA ; (iii) renforcer la capacité de pilotage et d'exécution ; (iv) donner aux agents et citoyens les moyens de s'impliquer dans cette transformation.

Dès 2024, les services publics devront décider s'ils utilisent des solutions d'IA sur étagère, s'ils entrent dans des partenariats avec des entreprises, ou s'ils redéveloppent leurs propres outils. Selon la logique de « plateforme publique » déjà appliquée dans la Santé ou l'Éducation, un travail sur le « pourquoi » (leur mission et leurs objectifs) est indispensable afin d'en déduire le « quoi » (les briques, services et jeux de données à développer) et le « comment » (la répartition des rôles entre différents acteurs publics et privés, l'implication des agents et des usagers, la considération des enjeux éthiques)[88].

Les solutions sur étagère auront l'avantage de la performance, de la simplicité, étant disponibles immédiatement, voire s'intégrant directement dans les outils des agents (suite bureautique, moteur de recherche). Mais elles présentent des risques, au premier rang desquels la fuite de données[89]. Le sujet n'est pas anodin. Certains ministères ont déjà interdit l'utilisation des outils d'aide au développement informatique, qu'il s'agisse du Copilot de Github ou GPT-4. Les développeurs informatiques recrutés (avec peine !) pour réinternaliser et maîtriser le numérique au sein de l'État se voient donc privés d'outils bientôt incontournables pour tout codeur. Les services publics devraient donc rapidement se doter de chartes d'utilisation et de contractualisation de solutions d'IA sur étagère, afin d'encourager les agents publics à se saisir de ces outils[90], d'autant plus quand ils sont gratuits pour un usage occasionnel.

Les solutions sur étagères auront toutefois des limites, soit parce qu'elles ne s'insèrent pas dans des outils existants, soit parce qu'elles ne sont pas adaptées à certains usages sensibles. Il serait délicat de demander à ChatGPT de résumer une note à destination d'un ministre par exemple. Mais il serait absurde que chaque ministère et collectivité redéveloppe ou rachète une IA capable de résumer une note sans en faire fuiter les données. La chaîne de production de l'IA

87. Les investissements numériques ne représentent que 3 % des investissements de l'Etat en 2019 : Cour des Comptes (2020), « La conduite des grands projets numériques de l'État »

88. Feuille de route du numérique en Santé, 2019-2022 et 2023-2026, Stratégie numérique pour l'éducation 2023-2027.

89. Une circulaire de 2021 interdit déjà l'utilisation au sein de l'État de Microsoft 365 sur un cloud de Microsoft, ce qui empêchera d'utiliser le module « Copilot » basé sur la technologie d'OpenAI. Les collectivités locales et les hôpitaux ne sont pas concernés.

90. L'Académie de la transformation numérique du gouvernement du Québec, portée par l'Université de Laval, est un exemple intéressant pour développer les compétences et diffuser une culture numérique au sein des services publics.

générative, faite de réentraînement et de raffinage de modèles se prête particulièrement bien à la mutualisation.

Le projet Albert, lancé à l'automne 2023, illustre bien cette dynamique. Cette IA générative permet notamment de faire du résumé de texte dans un langage administratif et sera bientôt intégrée à des outils collaboratifs existants. Dès 2024, Albert sera expérimenté au sein des guichets de services publics de proximité — les maisons France services — afin d'aider des usagers dans leurs démarches.

Albert est une brique qui peut être ensuite affinée avec des données propriétaires, insérée dans des interfaces spécifiques à chaque service public. Les services publics pourraient identifier ensemble d'autres fonctionnalités à mettre en commun, notamment là où ils auront à la fois un besoin partagé entre plusieurs services, et relativement spécifique par rapport au secteur privé. La simplification du langage administratif, l'accompagnement dans les démarches pourraient être de premières pistes.

Pour donner un accès large à ce service d'IA générative, mettre en place une infrastructure adaptée à l'IA et éviter les doubles investissements, notre Commission recommande également que soit renforcée la capacité de pilotage et d'exécution technique dans les services publics. Au niveau interministériel, une réelle direction technologique devrait pouvoir apporter non seulement de la doctrine, mais aussi des infrastructures de qualité (hébergement, puissance de calcul, usine logicielle, identité numérique[91], etc.), de l'expertise, et du budget de transformation. A l'heure où l'État cherche à réinternaliser des compétences et à les faire circuler entre ministères, il faut qu'il renforce sa capacité à produire du numérique de qualité.

Dans tous les services publics, le numérique et l'IA doivent être portés au bon niveau par les administrations et le politique, au-delà de simples fonctions support ou projets *ad hoc*. De la même façon que tout service public a en tête les contraintes et les possibilités du cadre juridique, le vecteur du numérique et de l'IA doivent devenir incontournable dans l'élaboration et la conduite des politiques publiques. Trop souvent, les décisions à haut niveau se concentrent sur le design d'une interface plutôt que sur la rapidité de l'application, sa résistance aux pics de charge ou la mesure de sa qualité. Trop souvent, des décisions d'élaboration ou de mise en œuvre sont prises sans réfléchir au bon vecteur numérique, alors que de meilleures solutions existent et sont équivalentes du point de vue de l'objectif de politique publique[92]. Les campagnes de tests et de vaccination contre le Covid-19 n'auraient pas été possibles sans prévoir la remontée des données au niveau national et sans coopérer en amont avec les éditeurs de logiciels des professions libérales pour qu'ils puissent communiquer les données.

Au-delà du pilotage, l'arrivée de l'IA pourrait être l'occasion de donner aux agents publics la capacité de transformer leur propre travail, plutôt que d'en subir la transformation d'en haut. Des outils permettent aujourd'hui de concevoir un outil numérique sans coder. « Démarches-simplifiées » a ainsi permis aux services de dématérialiser 32 468 démarches sans coder, dans l'État comme dans les collectivités locales. Le service public gagnerait à équiper ainsi les agents d'IA configurables pour qu'ils les déploient eux-mêmes, qu'il s'agisse de solutions sur étagère ou spécifiques au service public. Si les agents s'approprient l'IA, les usages foisonneront et permettront d'identifier plus rapidement là où l'IA a de la valeur. Le projet Llamendement de la Direction générale des finances publiques qui facilite le traitement des amendements du projet de loi des finances a ainsi été suggérée par un agent du métier impliqué sur les enjeux de transformation numérique.

91. Comme dans le cas du programme France Identité Numérique, il n'est pas nécessaire que le niveau interministériel opère toutes ces infrastructures mutualisées, tant que l'opérateur sait prendre en compte les besoins des autres administrations

92. Ce problème n'est pas spécifique à la France. Voir Pahlka, J. *Recoding America*

Cette prise en main par les agents eux-mêmes permettrait aussi de mieux définir où placer « l'humain dans la boucle ». Il est indispensable qu'un usager du service public puisse toujours s'adresser à une personne ou un service responsable d'une décision. Mais il y a souvent à gagner tant en qualité qu'en efficacité de traitement en automatisant certaines tâches. La bureaucratie, faîte de règles, de processus et d'égalité de traitement, entre parfois en tension avec les objectifs d'humanisation et de personnalisation du service public. En permettant d'automatiser les tâches bureaucratiques et de dégager du temps aux agents pour leur mission de service public, l'IA offre une voie de réconciliation.

Les citoyens eux-mêmes pourraient être amenés à contribuer aux services publics à partir d'IA, que ce soit pour définir leurs modalités de fonctionnement ou pour participer à leur construction. Leur implication est cruciale pour éviter que la transformation des services publics par l'IA ne renforce la centralisation bureaucratique, inexplicable et distante. À Taiwan, des « assemblées d'alignement » ont ainsi été réunies pour définir des règles de déploiement et de comportement d'IA dans le service public. Des initiatives de ce type pourraient être déployées en France, prolongeant les expériences de concertation déjà menées dans le numérique dans la santé et l'éducation par exemple.

La dématérialisation des services publics a parfois été synonyme de déshumanisation, tant pour les usagers que pour les agents forcés de faire rentrer leur service dans les cases précises d'un processus rigide. L'IA, et notamment l'IA générative, peut être l'occasion d'une réhumanisation des services publics, en rapprochant le service public des usagers et en permettant aux agents d'être de l'amélioration de leur travail.

Recommandation n° 9

Renforcer la capacité technique et l'infrastructure du numérique public afin de définir et de passer à l'échelle une réelle transformation des services publics grâce au numérique et à l'IA, pour les agents et au service des usagers.

2.1.4 MIEUX SOIGNER GRÂCE À L'IA : PLUS DE TEMPS AU SOIN

La transformation du système de santé par l'IA n'est pas à venir, elle a déjà commencé. L'IA est déjà présente dans de nombreux dispositifs médicaux afin d'améliorer l'analyse et les procédés cliniques. Les progrès dans la vision par ordinateur ont permis des progrès dans la robotique chirurgicale, dans l'analyse de radiographies au point que certains ont pensé que le métier de radiologue allait disparaître.

L'évolution des performances des systèmes d'IA depuis dix ans améliore déjà la précision des diagnostics, tout comme le ciblage et la vitesse de mise en place des processus de traitements. Dans le domaine de la recherche médicale, l'IA permet la découverte de nouveaux traitements. Ainsi, l'IA a conduit à la découverte récente d'un nouvel antibiotique dirigé contre le staphylocoque doré après 60 ans de recherches infructueuses[93]. Dans la santé publique, l'IA peut améliorer les prédictions d'évolution des épidémies[94].

Ces avancées ont déclenché des réflexions éthiques[95] et une évaluation de ces systèmes en situation réelle, afin de distinguer potentiel et effet réel. Par exemple, les systèmes de diagnostic par IA sont configurés pour être très sensibles et ne rater aucun signal. En contrepartie, ils sont plus susceptibles de repérer des signaux quand il n'y en a pas. Ils peuvent donc nécessiter des doubles contrôles qu'ils supposaient réduire[96],[97].

Avec cette histoire déjà longue, que changent les derniers progrès technologiques ? Par leur capacité à identifier des éléments dans du texte et dans la voix, à générer du texte et de la voix et à adopter des tons très variés, ces nouveaux systèmes changent radicalement la capacité d'interaction avec les soignants et les patients.

Du côté des soignants, la transformation numérique nécessite qu'ils structurent sans cesse leurs notes, réflexions et décisions médicales dans des logiciels qui ne peuvent que décevoir : soit leurs interfaces sont trop complexes, soit les options qu'ils proposent sont trop limitées. La multiplication des destinataires (spécialistes, généralistes, sécurité sociale, mutuelle, etc.) renforce le sentiment d'engorgement. C'est dans ce domaine que l'IA générative offre de grandes promesses.

Du côté des patients, la capacité d'une IA à dialoguer de façon crédible avec un patient soulève de nombreuses questions. Certains utilisateurs se tourneront vers des IA généralistes pour des diagnostics, malgré les mises en garde[98]. Dans le cadre d'une conversation par texte, une IA spécialisée peut être jugée plus empathique qu'un docteur humain[99]. Hors du diagnostic, une IA sera peut-être bientôt plus à même de tenir une longue conversation avec une personne atteinte d'Alzheimer qu'une personne humaine. Alors qu'il y a encore 5 ans, il semblait clair que la partie « relationnelle » du soin serait encore longtemps hors de portée des IA, cette frontière est désormais troublée.

93. Wong, F., Zheng, E.J., Valeri, J. A., *et al.* (2023) « Discovery of a structural class of antibiotics with explainable deep learning » Nature.
94. Olawade D. B., O. J. Wada, A. C. David-Olawade, E. Kunonga, O. Abaire, et J. Ling (2023) « Using artificial intelligence to improve public health: a narrative review » *Front Public Health*
95. Comité consultatif national d'éthique pour les sciences de la vie et de la santé (2023). « Diagnostic Médical et Intelligence Artificielle : Enjeux Ethiques »
96. Antun, V., F. Renna, C. Poon, B. Adcock, et A. C. Hansen (2020) « On instabilities of deep learning in image reconstruction and the potential costs of AI » *Proceedings of the National Academy of Science.*
97. Shen, Y., F. E. Shamout, J. R. Oliver, *et al.* « Artificial intelligence system reduces false-positive findings in the interpretation of breast ultrasound exams » *Nat Commun*
98. Un diagnostic de ChatGPT sur un enfant est aujourd'hui faux 83% du temps selon Barile, J., *et al.* (2024) « Diagnostic Accuracy of a Large Language Model in Pediatric Case Studies » *JAMA Pediatrics*
99. Tu, T. *et al.* (2024) « Towards Conversational Diagnostic AI »

Si l'étude citée précédemment[100] met en avant un meilleur diagnostic de la part de l'IA que des médecins dans la majorité des cas, elle précise que ce résultat a été obtenu dans des conditions particulières : la relation médecin-patient a été réduite à un *chat* en ligne, bien loin de la pratique médicale habituelle, qui aurait pu permettre au médecin de déceler des éléments impossibles à percevoir pour l'IA et donc améliorer son diagnostic. Ceci pose la question du cadre optimal d'utilisation de l'IA par les soignants. Une autre étude[101], qui se concentre cette fois sur le cas de la radiologie aboutit aux conclusions suivantes : si le médecin a une grande confiance dans son diagnostic, alors l'assistance par l'IA nuit à la qualité de ce diagnostic, car elle conduit parfois à faire douter le médecin, et donc à le faire changer d'avis à tort. En revanche, si le médecin est incertain, alors l'assistance par l'IA permet d'améliorer la qualité du diagnostic en moyenne. Ces études illustrent la complexité dans le déploiement de systèmes d'IA au service du soin : à ce stade de la technologie, ils doivent assister les soignants par des recommandations, sans donner l'illusion de fournir le diagnostic le plus adapté dans toutes les situations.

En France, la transformation numérique et par l'IA en santé a connu ces dernières années d'importants progrès, catalysés par la pandémie de la Covid. La feuille de route du numérique en santé a permis de clarifier les objectifs, de renforcer la confiance dans les outils numériques classiques et d'investir dans certaines infrastructures, notamment les référentiels communs.

Pour développer les capacités d'interaction des outils à base d'IA dans le domaine de la santé, il faudra pouvoir mettre à disposition des données riches de dialogue entre soignants et patients, faute de quoi ces IA ne seront entraînées que sur des données non francophones et dans des systèmes de santé différents. Ces mêmes données aideront à développer des IA à même de prendre en charge une partie des tâches administratives.

Pour développer les capacités des IA dans la prévention, le diagnostic et le traitement, les données de santé seront tout aussi cruciales. Certains outils d'IA faciliteront le recueil et la structuration des données, mais il faudra également améliorer les infrastructures numériques des hôpitaux et la mise à disposition de ces données auprès des chercheurs. Les bases de données de santé existent et elles sont d'une grande qualité, du fait du système de soin centralisé. Pourtant, l'accès aux bases de données reste limité tant à cause du mode de financement (chaque hôpital essaye de valoriser ces données individuellement, y compris pour payer ses infrastructures numériques) que du mode d'autorisation (les comités scientifiques et éthiques des établissements hospitaliers ne sont pas dimensionnés pour autoriser chaque demande d'accès aux bases de données).

Jusqu'ici, l'IA a surtout transformé les soins pour une petite partie de la population (ceux atteints de certaines maladies, ou ayant une radiographie complexe à lire). Les nouvelles avancées promettent de transformer l'expérience de soin pour tous, en libérant du temps de soignant. Pour le réaliser, nous devons collectivement accepter une meilleure circulation des données et exiger en retour qu'elles soient protégées, ainsi qu'une discussion ait eu lieu sur l'évolution du système de soin que nous voulons.

Recommandation n° 10

Faciliter la circulation des données, le partage de pratiques et l'évaluation pour tirer les bénéfices de l'IA dans les soins, améliorer l'offre et le quotidien des soignants.

100. Tu, T. et *al.* (2024) « Towards Conversational Diagnostic AI », *arXiv pre-print*
101. N. Agarwal, A. Moehring, P. Rajpurkar, et T. Salz (2023), « Combining Human Expertise with Artificial Intelligence: Experimental Evidence from Radiology », *Working paper.*

2.1.5 MIEUX ÉDUQUER GRÂCE À L'IA : L'ACCOMPAGNEMENT INDIVIDUALISÉ DES ÉLÈVES

L'éducation est l'un des domaines où l'impact de l'IA générative pourrait être le plus grand. L'éducation ne se prête cependant pas bien à la transformation numérique. Elle est faite de relations humaines, où le contenu est peu normé, malgré les programmes, et où le professeur doit disposer d'autonomie pour s'adapter au contexte, aux élèves. Elle ne se prête donc pas bien à la standardisation que requiert le numérique pour collecter et valoriser des données.

C'est pourquoi la plupart des innovations dans l'éducation numérique se sont faites moins dans la salle de classe qu'autour d'elle, pour rendre plus accessibles les manuels, notamment en ligne ou faciliter des démarches administratives. Hors de l'école, des formations totalement numériques ont été créées, mais principalement sans professeurs, pour apprendre des langues notamment. Même dans les pays les plus technophiles, la disruption de l'éducation par le numérique n'a pas vraiment eu lieu. Partout, l'influence des écrans sur la concentration des élèves conduit à la prudence, même s'il faut distinguer très clairement selon l'utilisation qui est faite des écrans.

Dans cet univers peu lisible pour la machine, l'IA générative pourrait changer la donne. Parce que l'IA générative peut produire du texte et des images, les modifier ou les classifier, et parce qu'il est possible d'interagir simplement en langage naturel avec cette IA, elle rouvre le champ des possibles. Ses usages sont multiples : l'appui à la production de contenus et de séquences pédagogiques, notamment en faisant des ponts entre matières, la personnalisation de ces contenus aux élèves, l'accompagnement des élèves sous la forme de tutorat, l'évaluation de l'apprentissage, en automatisant une partie de la correction, l'orientation des élèves, et même la formation des enseignants eux-mêmes.

Pour ne prendre que le cas du tutorat, il est désormais bien démontré que le tutorat personnel ou en petit groupe est une méthode d'apprentissage très efficace[102], car elle aide les élèves à se confronter à la matière. Malgré ses résultats très positifs, elle reste coûteuse. L'IA générative pourrait fournir à chaque élève un tuteur adapté à son niveau, à ses cours, capable d'aider l'élève à raisonner, disponible 24h/24, donnant au professeur des informations sur les éléments mal compris de son cours.

Avant de déployer un tuteur d'IA générative auprès de tous les élèves, quelques marches restent à franchir. En effet, ces systèmes sont aujourd'hui coûteux et trop lents, tous les élèves n'ont pas un terminal permettant de faire tourner une IA générative, les modèles d'IA générative ne savent pas bien repérer des erreurs de raisonnement, et surtout nous ne savons pas encore comment des élèves se saisiraient d'un tel outil. Le tutorat n'est peut-être efficace que lorsqu'un adulte est penché au-dessus de l'épaule de l'élève pour l'inciter à réfléchir. Déployée à l'échelle, cette solution poserait beaucoup de questions : l'IA doit-elle être utilisée pour des devoirs à la maison, dont on essaye de limiter le nombre en France ? Nécessite-t-elle la présence d'un adulte derrière chaque élève pour être utilisée ?

L'industrie technologique travaille déjà à rendre les IA moins chères, plus performantes, capables de raisonner, capables d'identifier des failles de raisonnement. Au-delà de la technologie, l'éducation nationale et l'enseignement supérieur sont les lieux où professeurs et élèves pourront

102. Nickow, A., Oreopoulos, P. et V. Quan (2020) « The impressive effects of tutoring on PreK-12 learning: a systematic review and meteaanalysis of the experimental evidence », *NBER Working Papers*.

expérimenter pour découvrir précisément quand et comment l'IA peut être utile à l'éducation et à la formation.

Notre Commission recommande donc d'avancer sur trois fronts de façon simultanée. À court terme, il est crucial d'encourager et sécuriser l'utilisation de l'IA par les enseignants dans la préparation et l'organisation de leurs cours, mais aussi aider les élèves à utiliser intelligemment l'IA dans leur apprentissage. À moyen terme, il est indispensable expérimenter une intégration profonde de l'IA dans l'éducation, en évaluant rigoureusement son apport, le matériel nécessaire et l'évolution des pédagogies. La stratégie du numérique pour l'éducation 2023-2027 est un bon support pour cet effort, notamment car elle organise la collaboration avec les acteurs privés et prévoit même un « compte ressource » pour que les équipes pédagogiques puissent recourir à des outils commerciaux.

Comme dans les services publics, cinq ingrédients seront nécessaires pour réussir cette transformation : des objectifs clairs, de la confiance, des personnes compétentes, des données et de l'infrastructure.

La confiance des élèves, des enseignants et des parents vis-à-vis d'une éducation avec l'IA nécessitera de la transparence, et de l'évaluation des outils, à commencer par les outils actuellement testés (« MIA seconde » pour l'enseignement des mathématiques ou « Jules » pour l'aide aux devoirs des collégiens)[103]. Pour que les élèves s'approprient l'IA, il faut qu'ils y soient formés dès le collège, dans sa dimension technique, mais aussi sociologique, historique, philosophique. Les collectivités locales pourraient jouer un rôle clé en construisant des projets liants temps scolaire et périscolaire. Cette appropriation, qu'elle se fasse grâce ou en dépit de l'éducation nationale, entraînera nécessairement une discussion sur l'utilisation du numérique et de l'IA à la maison. Après tout, des élèves utilisent déjà des outils commerciaux pour réviser avec des vidéos ou apprendre les mathématiques avec de la reconnaissance d'image. Une prise de position officielle à ce sujet aiderait les élèves et enseignants à s'emparer de ces outils.

En ce qui concerne les personnes, la réussite de l'IA dans l'éducation passera par l'implication et la complémentarité avec les enseignants. Le métier de l'enseignant sera bousculé par le développement de l'IA, qui pourrait faciliter l'évolution d'un enseignant « sachant » (expert disciplinaire) à un enseignant « accompagnant » l'élève, hors du paradigme historique de la transmission monodisciplinaire. Cela impliquera que les enseignants d'aujourd'hui soient formés et s'emparent de ces outils. Cela nécessitera également peut-être de revoir l'organisation actuelle[104] et les modes de recrutement.

Il est indispensable que l'Education nationale encourage, sécurise et valorise les enseignants qui expérimenteront l'utilisation de l'IA. Il est rassurant de voir les réseaux de formation des enseignants, notamment le réseau Canopé s'emparer rapidement du sujet afin de former les enseignants à la création de ressources pédagogiques basées sur les systèmes d'IA et mettre en place des projets pilotes en situation réelle. D'autres systèmes ont publié des guides à destination des parents, élèves et enseignants sur l'utilisation de l'IA générative, qui gagneraient à être émulés en France[105].

Pour ce qui est des données et des infrastructures, les besoins pour l'IA dans l'éducation sont de deux types. D'une part les données de contenu, pour que les IA proposées dans le cadre de l'éducation soient pertinentes et ajustées au contexte, à l'élève et à l'enseignant. D'autre part

103. À ce titre, il est impératif que les équipes pertinentes du ministère aient accès aux données d'utilisation de ces outils.
104. Le modèle danois semble est à ce titre intéressant vis-à-vis de l'IA : un groupe d'enseignants y est chargé de suivre une classe durant tout son premier cycle (9 ans, soit l'équivalent du primaire et du collège en France), d'assurer plusieurs matières et de créer un esprit de groupe.
105. Dès février 2023, le Land de Rhénanie du Nord-Westphalie a ainsi publié un guide pour les écoles et les élèves afin d'incorporer l'IA générative dans leur pratique.

les données relevant de l'activité des élèves et des enseignants (emplois du temps, vie scolaire, etc.). Ces données ne seront acquises et accessibles qu'avec un effort important sur les infrastructures et l'équipement des élèves. Le point bloquant reste effectivement la fragmentation des formats de contenu (manuels papier ou numériques) et des plateformes, tant pour les élèves que pour les enseignants. Des collaborations pourraient être utilement engagées avec les acteurs privés, par exemple en leur mettant à disposition des copies annotées afin de développer des IA pouvant appuyer les enseignants dans la correction.

En ce qui concerne les objectifs, la transformation de l'éducation nécessitera de rappeler et d'interroger les missions assignées à l'éducation. L'évaluation des IA nécessitera des protocoles précis et partagés pour déterminer les conditions où le numérique et l'IA sont utiles à l'apprentissage disciplinaire, mais aussi au développement de l'esprit critique. Nous devrons aussi arriver à dépasser collectivement la peur du « remplacement » de l'enseignant. Il ne fait aucun doute qu'ils resteront centraux dans la transmission du savoir et le développement de l'esprit critique. La complémentarité entre enseignants et IA doit permettre de décupler l'impact de chaque enseignant sur l'apprentissage des élèves,en estimant pour chaque option les bénéfices, les coûts et leur évolution[106].

Au-delà de l'apprentissage, l'IA pourra être mobilisée pour l'accompagnement à l'orientation. En effet, Parcoursup voit passer chaque année 917 000 candidats, pour les orienter sur les 23 000 formations proposant des diplômes reconnus par l'État. Ce moment est un rare exemple où une seule plateforme peut servir de support à l'introduction de l'IA, rassemblant au même endroit processus, information sur les profils et vœux des élèves, et choix des formations.

Recommandation n° 11

Encourager l'utilisation individuelle, l'expérimentation à grande échelle et l'évaluation des outils d'IA pour renforcer le service public de l'éducation et améliorer le quotidien des équipes pédagogiques.

106. Le coût de fonctionnement des IA est tel que généraliser aujourd'hui un système « enseignant + IA » ne serait pas soutenable financièrement. Toutefois les coûts décroissent très rapidement à performance constante. L'évaluation et la transparence sur les bénéfices et le coût à l'échelle sera crucial pour assurer la confiance des parents et des enseignants dans ces outils

2.2 SOUVERAINETÉ : INVESTIR POUR NOTRE AUTONOMIE STRATÉGIQUE

L'économie du numérique est dominée par quelques entreprises, telles qu'Alphabet/Google, Meta/Facebook, Microsoft et Amazon, également très présentes dans le domaine de l'IA. Les entreprises européennes sont spécialisées dans les industries traditionnelles, alors que les États-Unis sont les fers de lance de la révolution technologique du numérique[107].

Cet écart a des conséquences importantes sur les dépenses en recherche et développement (R&D). Chaque année, la Commission européenne établit la liste des 2 500 premiers investisseurs en R&D du monde[108]. Dans le domaine des technologies de l'information et des télécommunications, le dernier classement est saisissant : les États-Unis comptent plus de quatre fois plus d'entreprises que l'Union européenne et ces entreprises américaines investissent plus de six fois plus en R&D que les entreprises européennes. L'Europe est également très nettement distancée par la Chine en la matière[109].

La suprématie économique des acteurs étrangers pose la question de notre souveraineté, une question que renforce le développement et le potentiel de l'IA générative : si l'économie européenne dépend largement des entreprises étrangères, qu'adviendra-t-il si les relations internationales se tendent avec ses États fournisseurs ? Ne risquons-nous pas de voir une partie importante de notre valeur ajoutée économique, voire de notre connaissance, captée par les entreprises étrangères qui fournissent les systèmes d'IA ? Puisque l'IA contribue à la manière dont nous percevons le monde, comment préserver la culture française et européenne si nos outils sont bâtis sur des référentiels étrangers ?

107. Coste, O., et Y. Coatanlem (2023) « Tech : quand l'Europe s'éveillera », *Commentaires*
108. Commission européenne (2020). *The 2020 EU Industrial R&D Investment Scoreboard.*
109. La Chine compte plus de trois fois plus d'entreprises que l'Union européenne et ces entreprises chinoises investissent plus de deux fois plus en R&D que les entreprises européennes.

De nouvelles mesures américaines risquent de freiner le développement des capacités en IA[110] du reste du monde, Europe comprise

Le 30 octobre 2023, l'administration Biden a publié un décret sur le développement et l'emploi sûrs, sécurisés et dignes de confiance de l'IA. Le texte prévoit notamment des obligations de signalement (*reporting*) pour les fournisseurs américains de services de *cloud*. Ceux-ci — et leurs revendeurs — devront notifier au gouvernement américain les transactions avec tout acteur étranger relatives à l'hébergement de grands modèles d'IA dont les capacités pourraient être utilisées à des fins malveillantes dans le domaine cyber. À ce stade, entrerait dans le champ de ces obligations déclaratives l'entraînement de tout modèle de fondation au-dessus d'un certain seuil de complexité.

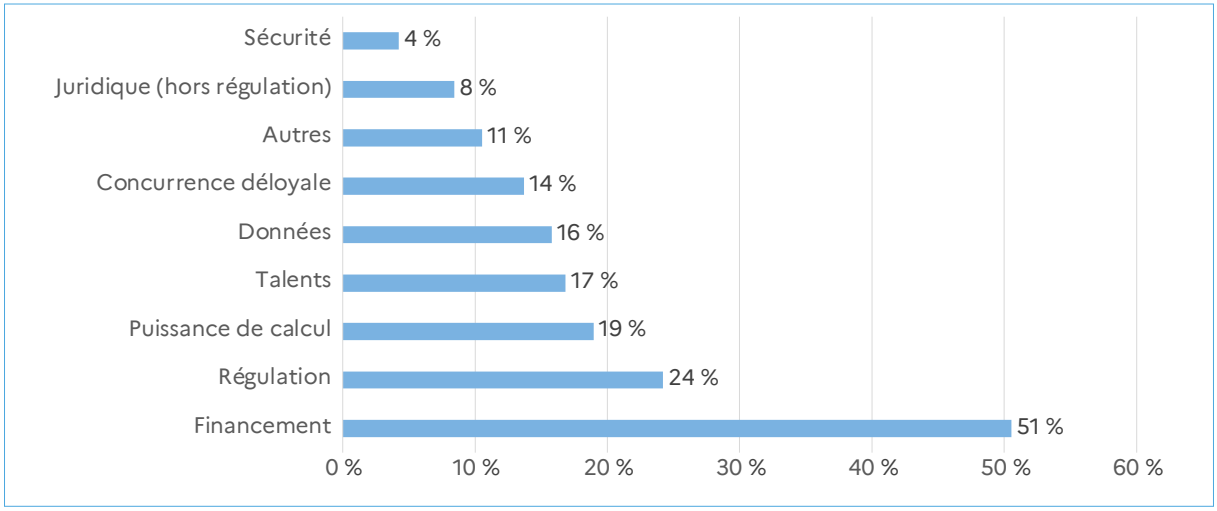
Ces dispositions font bien davantage que soulever des interrogations en matière de confidentialité des données et de protection du secret des affaires pour nos entreprises européennes. Elles contribuent à renforcer la domination américaine, en freinant le développement des capacités étrangères en matière d'IA et en facilitant le renseignement économique.

Pour répondre à ces interrogations, précisons le concept de souveraineté. Est souverain un État qui n'est soumis à aucun autre État. Une approche radicale de la souveraineté pourrait donc conduire à l'autarcie, afin de maîtriser la totalité de la chaîne de valeur de l'IA. Plusieurs pays dans le monde empruntent cette voie, notamment la Russie et la Chine.

Notre Commission considère qu'une approche autarcique de la souveraineté n'est pas opportune, ni au regard de nos valeurs ni au regard de nos intérêts. D'une part, la France promeut un cadre ouvert et démocratique de société. D'autre part, il serait techniquement et financièrement impossible de maîtriser l'ensemble des biens et services composant la chaîne de valeur de l'IA.

Nous estimons que la France devrait choisir la voie de l'*inter*dépendance, juste équilibre entre la dépendance totale (aucune souveraineté) et l'autarcie (aucun lien de dépendance). Concrètement, il s'agit pour la France de se doter d'avantages comparatifs en se positionnant sur quelques briques technologiques et quelques maillons de la chaîne de valeur. Pour y parvenir et renforcer la souveraineté européenne, nous recommandons de développer un écosystème de l'IA dynamique et attractif. Quatre piliers sont nécessaires : le financement, la puissance de calcul, l'accès aux données et les talents.

110. Source : direction générale du Trésor, service économique aux États-Unis.



Graphique 9 : Principaux défis identifiés par les start-up européennes de l'IA générative.
Source : Sondage réalisé dans le cadre du Generative AI in the European start-up landscape 2024.
Note : les start-up peuvent identifier plusieurs défis.

2.2.1 FINANCER DURABLEMENT L'INNOVATION : L'INDISPENSABLE CHANGEMENT D'ÉCHELLE

2.2.1.1 Des investissements dans l'IA au moins trois fois trop faibles

Les innovations technologiques sont portées soit par des entreprises existantes, soit par des *start-up*. L'IA ne fait pas exception. Or, nous venons de montrer que les entreprises françaises et européennes sont peu positionnées sur l'économie numérique et donc moins promptes à investir dans l'émergence de solutions d'IA de rupture.

En France, le développement d'un tissu de nouvelles entreprises spécialisées dans l'IA est donc incontournable. Pour y parvenir, des financements spécifiques doivent être mobilisés. C'est en particulier l'objet du « capital risque », qui finance la création ou le développement d'entreprises risquées mais à fort potentiel.

Or, les montants investis en capital risque dans l'IA sont très insuffisants. En 2022, les investissements s'élèvent à 2,8 Md$[111] en France, contre 56,8 Md$ aux États-Unis. Pour avoir un investissement comparable à celui des États-Unis, il faudrait que la France investisse entre 8,4 et 10 Md$ par an, soit au moins entre 5,6 et 7,2 Md$ supplémentaires chaque année. Ces différences sont souvent expliquées par des différences culturelles et une faible appétence au risque des acteurs français.

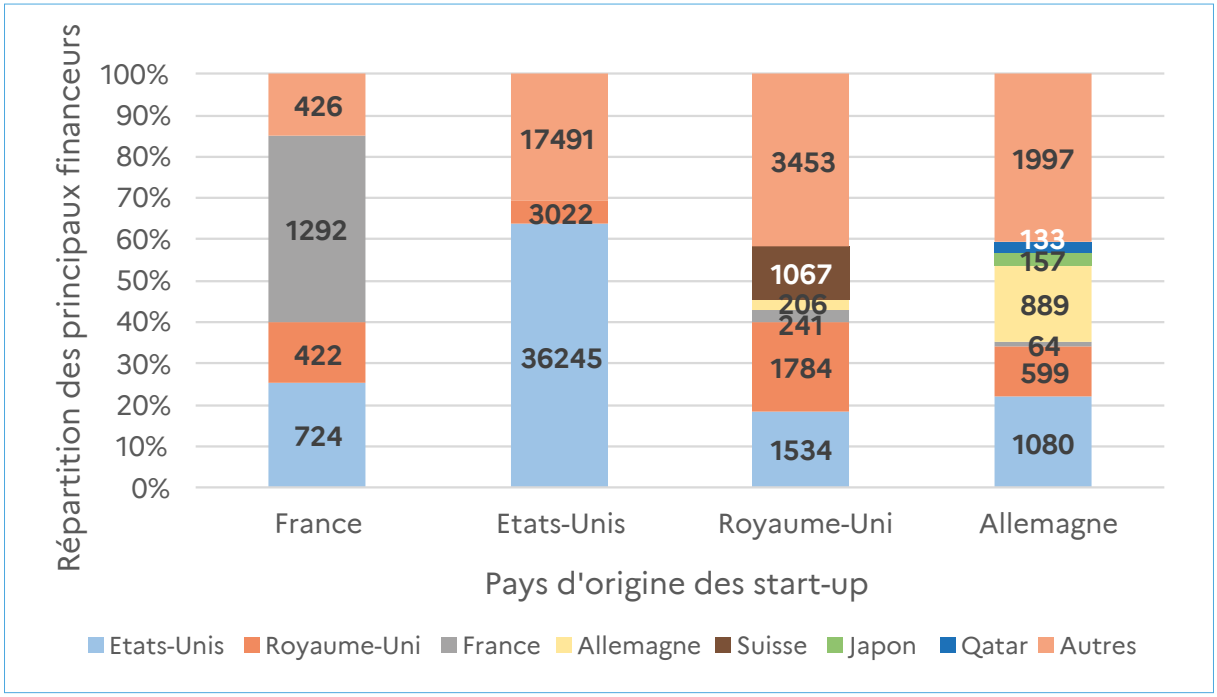
Notre Commission estime donc que la France devrait au moins tripler les investissements dans l'IA. Ce montant constitue un minimum, car il ne tient pas compte de l'évolution des besoins de financement. Cette évolution est atteignable. Entre 2018 et 2022, les investissements en

111. Observatoire IA de l'OCDE.

capital-risque dans l'IA en France ont été multipliés par quatre, en passant de 0,6 à 2,8 Md$. Il faut donc maintenir le cap.

En outre, les investissements des fonds européens sont majoritairement de petite taille (rarement plus de 30 M€), alors que les besoins annoncés pour certaines entreprises se chiffrent en centaines de millions, voire en milliards d'euros. Il est essentiel d'y remédier car, en l'absence d'investissements européens, les financements apportés par des acteurs extra européens contribuent à la délocalisation des activités.

L'augmentation des investissements français et européens dans le domaine de l'IA ne signifie pas qu'il faut viser un financement 100 % national ou européen. En effet, la diversité des financeurs contribue au succès du développement des *start-up*. Par ailleurs, les investissements étrangers témoignent de l'attractivité de l'écosystème national d'innovation. Aux États-Unis, les *start-up* de l'IA sont financées à 63 % par des investisseurs nationaux. Une part qui s'élève à 45 % en France, 25 % au Royaume-Uni et 24 % en Allemagne.



Graphique 10 : Origine des financements des *start-up* par pays en 2022 (tous secteurs)
Source : OCDE.
Lecture : Les start-up américaines sont financées à plus de 60 % par capital venture américain

2.2.1.2 Une ambition d'investissements qui nécessite des choix de société

Notre Commission estime qu'il est nécessaire d'investir dans l'IA entre 5 et 8 Md$ en plus par an, soit un triplement des sommes actuellement engagées. Si cette dynamique est un minimum, nous recommandons que dans cette phase d'émergence de nouvelles technologies, porteuses de très grands potentiels (voir partie 1), où les premiers investissements sont essentiels pour se positionner et prendre des positions sur la chaîne de valeur, la France se fixe des objectifs de financements de l'IA à la hauteur de ses ambitions, à savoir d'être parmi les premières économies de l'IA. Un objectif d'investissements annuels de 15 Md$ pourrait être visé.

Le changement d'échelle suppose de réorienter une partie de l'épargne privée, en mobilisant entre 0,1 % (démarche de rattrapage) à 0,3 % (démarche pionnière) de l'épargne des ménages français[112].

Des réformes sociétales peuvent être envisagées pour permettre un fléchage de l'épargne durablement différent. Les pays où les entreprises technologiques sont les plus nombreuses et les plus dynamiques ouvrent plusieurs chemins possibles. La France pourrait faire évoluer les incitations fiscales des assurances vie, afin de flécher davantage les fonds en direction de l'innovation. Elle pourrait aussi faire évoluer la gestion des retraites complémentaires pour assumer des investissements plus risqués et dans le temps long.

L'approfondissement de la construction européenne constitue une autre voie de financement de l'innovation. Le morcellement du marché des capitaux en Europe réduit la taille des investissements. Pour y remédier, l'Union européenne devrait mettre en place une véritable union des marchés de capitaux, c'est-à-dire permettre et favoriser la libre circulation de l'épargne et des investissements. Le premier plan d'action européen date de 1999… mais il reste tant à faire ! Ce pourrait être un chantier prioritaire à partir de juin 2024, pour la prochaine législature de l'Union européenne.

Un troisième objectif mérite d'être souligné : le renforcement de l'attractivité des fonds d'investissement étrangers, afin qu'ils s'implantent à Paris et pas uniquement à Londres. Il ne s'agit pas d'une solution alternative aux deux objectifs cités précédemment (évolution du fléchage de l'épargne privée française, renforcement de l'intégration du marché européen des capitaux), mais bien d'une solution complémentaire permettant à l'écosystème français d'innovation de se développer plus rapidement.

Notre Commission considère que ces trois chantiers pourraient être menés parallèlement. Ils porteront leurs fruits à long terme et au-delà du seul champ de l'intelligence artificielle.

Pour accélérer à court terme le développement de l'écosystème d'IA, nous recommandons de créer d'ici fin 2024 un fonds d'investissement « France & IA ». Le fonds viserait à la fois à soutenir l'émergence de *start-up* spécialisées dans l'IA appliquée et à faciliter la transformation du tissu économique de PME et ETI. Il mobiliserait 7 Md€ de capital investissement d'entreprise et 3 Md€ de soutien public, selon plusieurs modalités d'intervention (une première enveloppe de fonds de fonds, une seconde de co-investissements, une troisième de financement par dette de projets de transformation numérique). Aux côtés des moyens financiers, le fonds s'accompagnera d'une mise en commun de données d'activité pour conduire certains projets numériques. L'ampleur des moyens est inédite en France et représenterait un investissement moyen élevé pour chaque entreprise. Si chacune des 250 grandes entreprises Françaises[113] hors contrôle étranger y participait, cela représenterait un investissement de l'ordre de 25 M€

112. L'épargne des français représente en 2023 un peu plus de 6 000 Md€.
113. Entreprises d'au moins 5 000 salariés, ou au moins 1,5 Md€ de chiffre d'affaires et plus de 2 Md€ de bilan.

chacune. Il pourrait donc être envisagé de réunir l'enveloppe totale de 10 Md€ en deux temps. Face au risque de déclassement économique, l'audace participera de l'émergence de solutions innovantes performantes et de l'accélération de la modernisation des entreprises françaises.

Recommandation n° 12

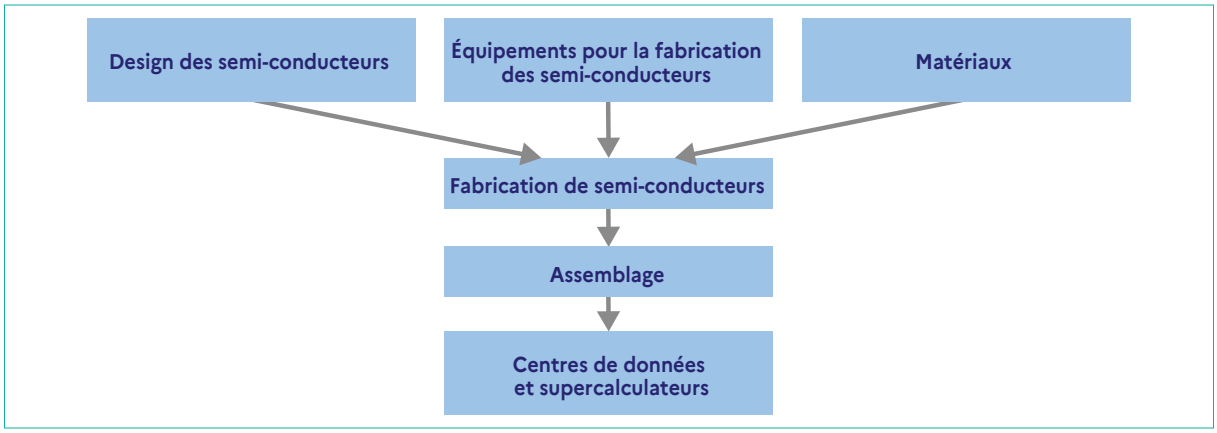
Investir massivement dans les entreprises du numérique et la transformation des entreprises pour soutenir l'écosystème français de l'IA et en faire l'un des premiers mondiaux.

L'objectif de triplement (au moins) des investissements dans le domaine de l'IA ne doit pas s'entendre comme une cible uniquement quantitative. La qualité des investissements est tout aussi importante. Le ciblage et la concentration des moyens sont primordiaux pour fonder la supériorité de la France et de l'Europe sur certains segments de la chaîne de valeur et ainsi être en mesure de parler d'égal à égal avec nos concurrents et nos partenaires. Cette voie de la différenciation pourra notamment porter sur le développement de composantes *open source* mais aussi sur la dimension environnementale, en visant de nouvelles générations d'IA, de l'architecture matérielle au choix des modèles, qui consommeront moins d'énergie. Cette différenciation va de pair avec l'émergence progressive et constante d'écosystèmes d'innovation dans le domaine de l'IA en France et en Europe.

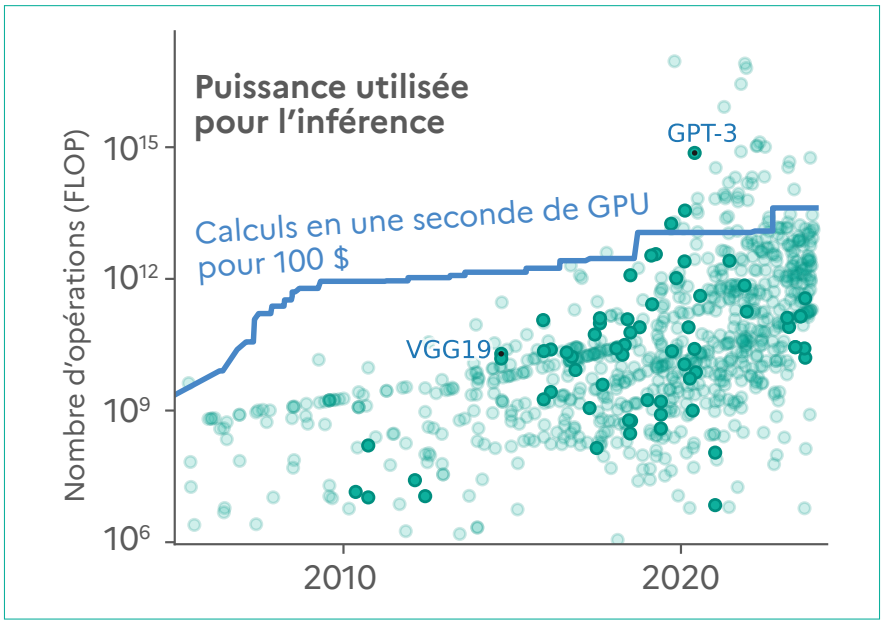
2.2.2 DISPOSER DE CAPACITÉS DE CALCUL SOUVERAINES : UNE CONDITION *SINE QUA NON* DE L'AUTONOMIE STRATÉGIQUE

La puissance de calcul est un ingrédient incontournable de l'IA générative. Elle constitue le socle de la chaîne de valeur de l'IA générative (voir le schéma en introduction du rapport). Concrètement, elle consiste en un mélange de matériel électronique distribué et de services associés.

En effet, les calculs requis pour l'entraînement ou l'utilisation (« l'inférence ») d'un modèle d'IA sont réalisés par des réseaux de semi-conducteurs assemblés, formant des serveurs eux-mêmes rassemblés au sein de supercalculateurs ou de centres de données. Lorsqu'une entreprise, une administration ou un particulier utilise un modèle d'IA, il recourt donc aux services de calcul de plusieurs supercalculateurs ou de centres de données.



Les besoins en puissance de calcul croissent aujourd'hui très rapidement par le produit de trois dynamiques. Tout d'abord, l'entraînement et l'utilisation (« l'inférence ») de modèles de plus en plus grande dimension est gourmand en calcul. En sens inverse, des efforts importants sont réalisés aujourd'hui pour diminuer la puissance de calcul nécessaire à l'entraînement et à l'inférence, car c'est l'un des principaux postes de coûts dans la construction et l'utilisation d'un modèle d'IA, et l'un des obstacles à sa diffusion. La compétition se joue d'ailleurs actuellement autant sur la performance des modèles que sur la capacité à développer des IA dont le coût d'utilisation est acceptable. Depuis 2020, la puissance nécessaire pour l'inférence des modèles d'IA augmente plus vite que ne décroît le coût de l'inférence (*cf. graphique 11*). Enfin, même si l'entraînement ou l'utilisation des modèles devient plus économe à performance égale, cela pourrait accélérer leur diffusion, et donc la puissance de calcul totale nécessaire. En dépit de l'incertitude sur ces trois dynamiques, notre Commission estime que les besoins continueront de croître rapidement à court terme, avec la diffusion de l'IA dans toutes les sphères économiques et sociales.



Graphique 11 : Évolution de la puissance et du coût de l'inférence des modèles d'IA.
Source : calculs de la Commission IA (2024).
Lecture : Chaque point représente un modèle d'IA et la puissance nécessaire pour pour l'inférence. La courbe indique l'évolution du nombre de calculs réalisables en une seconde de GPU pour 100 $.

Or, la France et l'Europe sont aujourd'hui très loin de pouvoir répondre à cette demande croissante. Le marché est dominé par quelques entreprises, principalement américaines, détenant individuellement ou collectivement jusqu'à 80 % des parts de marché mondial de certains segments d'activité, que ce soit dans le *design* de composants semi-conducteurs spécialisés dans l'IA, ou l'offre de puissance de calcul pour entraîner ou utiliser des IA.

Cette dépendance nous empêchera de bénéficier d'une part importante de la valeur ajoutée de l'IA, qui sera captée par des acteurs étrangers. Elle pose plus largement des enjeux de souveraineté : difficultés de protection des données, difficultés d'élaboration des normes appropriées à des enjeux technologiques mal maîtrisés, exposition aux tensions géopolitiques… Il est donc impératif d'y remédier.

Que se passe-t-il dans le reste du monde ? L'exemple des États-Unis[114]

Aux États-Unis, le décret du 30 octobre 2023 impose aux fournisseurs de capacités de calcul en nuage (*cloud*) de déclarer à l'administration américaine la localisation et la capacité de leurs infrastructures de calcul au-delà d'un certain seuil[115].

2.2.2.1 Investir dans une filière de composants semi-conducteurs optimisés pour l'IA

À ce jour, la filière de composants semi-conducteurs à destination des systèmes d'IA générative est extrêmement concentrée. En matière de conception de certains semi-conducteurs clés, l'entreprise américaine NVIDIA est en situation de quasi-monopole. En plus de la conception des composants, cette entreprise propose une couche logicielle (appelée *Cuda*) qui domine également le marché.

La fabrication des composants semi-conducteurs spécifiques à l'IA générative est tout aussi monopolistique, l'entreprise taiwanaise TSMC détenant 85 % du marché des semi-conducteurs de moins de 7 nanomètres. Notons cependant que ce fabricant asiatique s'appuie sur des équipements produits dans le monde entier et rassemblés dans des machines conçues par une entreprise néerlandaise dénommée ASML.

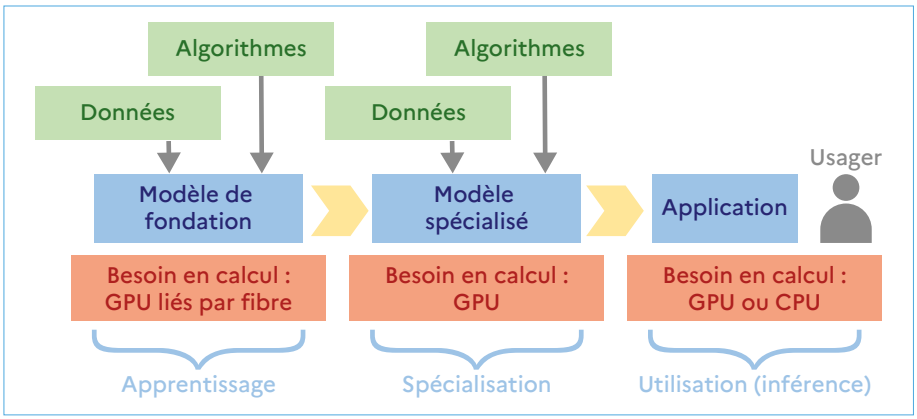
Ce positionnement monopolistique n'est pas une fatalité. D'une part, le marché des semi-conducteurs devrait augmenter rapidement sous l'effet de la demande de puissance de calcul. D'autre part, l'électronique actuelle n'est pas pleinement optimisée à l'entraînement et surtout à l'utilisation des systèmes d'IA générative (voir l'encadré ci-dessous). Ce double contexte, de croissance et de segmentation du marché, est propice à l'entrée de nouveaux acteurs (Google et Amazon conçoivent ainsi déjà leurs propres composants spécialisés dans l'IA).

114. Source : direction générale du Trésor, service économique aux États-Unis.

115. A noter que ces seuils mesurés en opérations de calcul mesurent assez mal la « puissance » d'un modèle ou ses « capacités », d'autant qu'il est possible de couper le modèle pour l'entraîner « par morceaux » ou au contraire de dépasser le seuil simplement parce qu'on a raffiné un modèle pour le rendre plus spécifique.

Quelle électronique pour les modèles d'IA générative ?

Les différentes étapes de création d'une application d'IA générative requièrent différents réseaux de composants semi-conducteurs qui vont réaliser les calculs nécessaires. Les étapes d'apprentissage et de spécialisation nécessitent un calcul massif à réaliser de manière intensive, ceux-ci sont réalisés par des GPU (*graphics processing unit*), qui ont été créés initialement pour le jeu vidéo et non l'IA. L'étape d'inférence, quant à elle, requiert beaucoup moins de puissance de calcul. Elle peut faire appel aux GPU ou à des composants semi-conducteurs plus répandus : les CPU (*central processing unit*).



Pour la plupart d'entre eux, les composants semi-conducteurs actuellement utilisés pour les différentes étapes de conception d'une solution mobilisant l'IA générative ont des applications multiples. Une minorité de ces composants est optimisée pour l'IA, que ce soit sur le plan de la rapidité ou encore de la consommation énergétique. Ces critères sont particulièrement essentiels en prévision d'un déploiement massif de solutions mobilisant l'IA dans les centres de données ou dans les objets du quotidien.

Au-delà de la nature des semi-conducteurs à utiliser, il faut aussi considérer le volume d'acteurs concernés par chacune des phases afin de mesurer la puissance de calcul globale. En effet, un modèle de fondation permet la création de multiples modèles spécialisés dans plusieurs domaines, chacun de ces modèles spécialisés étant ensuite utilisé par de multiples usagers. Au total, l'entraînement et l'inférence créent un besoin très important de puissance de calcul tout au long de la chaîne de valeur.

L'IA représente donc pour l'Europe une opportunité de se positionner sur les prochaines générations de semi-conducteurs qui formeront une brique technologique stratégique. Il ne s'agit pas de viser un objectif autarcique, mais de bénéficier de la croissance de nouvelles activités économiques et de rééquilibrer les rapports de force stratégiques.

Au regard de l'ampleur des investissements nécessaires, des risques associés et des délais inhérents au développement d'une filière industrielle de semi-conducteurs optimisés pour l'IA, l'intervention de la puissance publique apparaît nécessaire. Cette intervention pourrait cibler d'une part les puces spécialisées pour l'inférence et d'autre part les puces adaptées aux systèmes

d'IA embarqués (*edge*). Le cadre d'un tel soutien pourrait s'inscrire dans le règlement européen sur les semi-conducteurs, adopté en 2023, qui vise justement à réduire la dépendance de l'Union européenne dans le domaine.

En revanche, le soutien public ne devrait pas viser une concurrence des acteurs sur les technologies utilisées actuellement. Les montants d'investissements sont en effet tels que le soutien public ne suffirait pas, sans traction du marché ou alliance avec un géant du numérique.

À cet effet, la France doit poursuivre le soutien de la recherche publique et apporter un soutien financier par le biais de subventions et de fonds propres pour faire émerger une telle filière industrielle. Au-delà de la conception et de la fabrication des composants, le soutien des acteurs européens devra aussi porter sur la promotion de solutions ouvertes (*open source*) pour la couche logicielle, pour sortir de la dépendance à la couche logiciel propriétaire, majoritairement utilisée aujourd'hui.

Recommandation n° 13

Accélérer l'émergence d'une filière européenne de composants semi-conducteurs adaptés aux systèmes d'IA.

2.2.2.2 Accroître la présence de centres de données sur notre territoire

Avec la filière de semi-conducteurs, les services associés aux supercalculateurs et aux centres de données constituent l'incontournable deuxième brique de la puissance de calcul pour l'IA. Le monde se livre donc à une course de vitesse sur la croissance de la puissance de calcul : en 2023, la puissance cumulée des dix supercalculateurs les plus puissants du monde a augmenté de 72 % par rapport à 2022[116].

L'Europe est en retard en ce qui concerne la puissance de calcul publique et elle est très largement distancée en ce qui concerne la puissance de calcul installée par des acteurs privés.

La puissance de calcul publique de l'Europe est nettement plus faible qu'outre-Atlantique. Elle héberge aujourd'hui 143 des 500 plus grands supercalculateurs — majoritairement publics — du monde (contre 161 pour les États-Unis) représentant 24 % de la puissance de calcul correspondante (contre 53 % pour les États-Unis). L'Union européenne et ses États membres investissent toutefois dans le cadre d'un programme intitulé EuroHPC (*European high performance computing*). De prochaines installations de supercalculateurs massivement parallèles et de très grande taille, dits « *exascale* » sont prévues, en Allemagne en 2024 (« Jupiter ») et en France en 2026 (« Jules Verne »). Toutefois, en parallèle, les États-Unis et la Chine ont déjà lancé des travaux pour des supercalculateurs dits « *post-exascale* ».

La puissance de calcul des acteurs privés est pour sa part très largement installée aux États-Unis. Cette prédominance est double. D'une part, les centres de données sont en général bien plus présents sur le territoire américain qu'européen : les États-Unis disposent de 2 109 centres de données (soit une part de 37,8 % des centres de données du monde) contre 1 244 centres de données sur l'ensemble des pays de l'Union européenne (soit une part de 22,3 %)[117]. D'autre part, les plus grands fournisseurs de services à partir de ces centres de données (*cloud*) sont

116. Selon les données du Top500 de novembre 2023. La source est identique pour le paragraphe qui suit.
117. D'après les données du site DataCenterMap de janvier 2024.

américains. Trois grands acteurs représentent ainsi 80 % de la hausse des dépenses en infrastructures et applications de services de *cloud* public en France.

Ce déséquilibre a tendance à s'aggraver en ce qui concerne la puissance de calcul dédiée à l'IA. Pour la seule année 2023, les achats cumulés de Meta, Microsoft, Alphabet et Amazon s'élèveraient à 400 000 GPU, contre quelques milliers d'unités pour les acteurs européens. À lui seul, Meta prévoit d'acheter, en 2024, 300 000 cartes graphiques de dernière génération spécialisées dans l'entraînement d'IA (le modèle H100 de NVIDIA ou équivalent), alors que l'entreprise en dispose déjà de 300 000. Certes, une partie sera installée sur des infrastructures en Europe. Cependant, tirer tout le potentiel de ces machines nécessite aujourd'hui de les localiser dans des infrastructures de grande taille, qui sont moins présentes en Europe qu'aux États-Unis.

Il serait tentant de considérer que ce retard de l'Europe n'est pas un problème, qu'il n'est que le prolongement d'un déficit en matière de services d'hébergement. De fait, dans une économie mondialisée, il n'est pas nécessaire que les IA utilisées sur le sol européen ou par des consommateurs européens soient toutes entraînées et déployées sur des infrastructures strictement européennes.

Le retard de l'Europe en matière de puissance de calcul pose toutefois trois difficultés. Premièrement, il existe des usages sensibles de l'IA, sur lesquels il n'est pas envisageable de d'entraîner les modèles ou les déployer sur de la puissance de calcul extra européenne. Ces usages se limitent certes aujourd'hui à quelques acteurs, notamment militaires, qui disposent de leurs propres infrastructures. Ce ne sera demain plus le cas, que l'on songe aux impôts, à la santé ou à la justice. Deuxièmement, dans un monde où l'IA sera intégrée à de nombreux aspects de notre vie sociale et de notre économie, la résilience et l'autonomie stratégique nécessitent de disposer d'une quantité minimale de puissance de calcul sur notre sol. Pour faire un parallèle avec l'électricité : il n'est pas nécessaire que tous les électrons alimentant notre économie soient produits en France, mais il serait dangereux d'être obligé d'importer 80 % de son électricité. La puissance de calcul constitue un point de contrôle dans la chaîne d'approvisionnement de l'IA, car on peut la détecter, interdire son usage, la quantifier et qu'elle est concentrée. Certains y voient donc déjà un point névralgique pour gouverner l'IA[118]. Enfin, la hausse de la demande de puissance de calcul conduira à une dégradation de notre balance commerciale, si la France n'accroît pas simultanément son offre.

118. Sastry, G., L. Heim, H. Belfield et al. (2024) "Computing Power and the Governance of Artificial Intelligence" *Working paper.*

Que se passe-t-il dans le reste du monde ? Les exemples de la Suède, de Taiwan et du Japon[119]

En Suède, le mix énergétique fortement décarboné (les combustibles fossiles représentent 1 % de la production d'électricité) et le plus faible prix de l'électricité sont un atout pour l'installation de supercalculateurs. La *start-up* allemande de service de traduction DeepL a par exemple choisi ce pays pour son nouveau supercalculateur Mercury, qui est aujourd'hui considéré comme le plus puissant ordinateur de Suède.

À Taiwan, les centres de données et la puissance de calcul pour l'IA doivent être localisés sur le sol taiwanais, pour des motifs de résilience. En revanche, il n'existe pas de restriction sur la nationalité des fournisseurs de puissance de calcul. De fait, les trois *hyperscalers* américains (Amazon Web Services, Google Cloud et Microsoft Azure) fournissent la quasi-totalité de l'hébergement et de la puissance de calcul du pays.

Au Japon, l'État subventionne l'utilisation des ressources de calcul pour soutenir le développement des modèles de base de l'IA générative. Il soutient également le développement des centres de données en subventionnant une portion du coût des infrastructures pour les entreprises fournissant ces services. Le financement public est plus élevé en cas de localisation dans les zones rurales.

Notre Commission considère que la France et l'Europe doivent impérativement devenir un pôle majeur de la puissance de calcul installée avec trois objectifs : (i) fournir une puissance de calcul publique pour les cas d'usage sensibles ; (ii) fournir une puissance de calcul accessible et abordable pour stimuler la recherche et le développement des *start-up en IA* ; (iii) être en mesure d'entraîner et d'utiliser sur le sol européen les modèles d'IA les plus avancés.

Pour les cas d'usage sensible, la recherche et le soutien aux *start-up*, il convient de poursuivre les investissements dans les supercalculateurs publics. L'accès aux supercalculateurs financés par l'Europe ou la France est gratuit pour les acteurs publics ou privés qui contribuent à la recherche ouverte. Cette logique permet de bénéficier des infrastructures financées sur les fonds publics en contrepartie d'une contribution permettant de faire avancer la connaissance collective. Cette dynamique doit être poursuivie pour permettre aux acteurs européens de la recherche de rester à la frontière technologique.

Pour les usages privés ou non sensibles, il est impératif de développer des centres de données sur notre territoire. L'installation d'un centre de données sur le territoire français soulève des enjeux de gestion du foncier, de coordination d'acteurs, de procédures administratives et d'accès à une électricité à un prix compétitif. À cet égard, la consommation d'électricité représente un coût important pour les acteurs du marché : il s'agit donc d'un facteur clé de décision d'implantation. Face au soutien financier proposé par certains États américains, qui entendent attirer les centres de données, la France pourrait renforcer sa compétitivité fiscale. Pour assurer à nos acteurs développant et utilisant de l'IA un accès à une puissance de calcul abordable, il faut par ailleurs prendre acte du contexte de tension très forte dans lequel se trouve le marché des processeurs spécialisés pour l'IA. L'accès à la puissance de calcul sera un

119. Source : direction générale du Trésor, service économique en Suède et au Japon.

différenciant important dans les quelques années à venir, et le coût un paramètre crucial pour la compétitivité des acteurs entraînant des IA.

Quels objectifs se fixer pour la puissance de calcul en France ?

Dans un marché en très forte croissance, et aux paramètres technologiques incertains, notre Commission a tenté d'estimer les besoins en puissance de calcul selon plusieurs méthodes. Les besoins en puissance de calcul varient largement, selon que l'on considère une entreprise qui entraîne des modèles de fondation, qui affine des modèles plus spécialisés ou qui utilise des modèles d'IA dans ses opérations.fre

La France pourrait se fixer comme objectif d'ici 2026 de pouvoir offrir sur son sol de la puissance de calcul pour soutenir 5 entreprises entraînant des modèles de fondation de la prochaine génération en un temps raisonnable, 50 entreprises affinant des modèles de fondation spécialisés et 1 000 entreprises utilisant des modèles d'IA dans leurs opérations.

En partant des données publiées pour Llama 2, entraîner trois versions d'un modèle de pointe en 2023 a nécessité l'équivalent de 100 processeurs graphiques de dernière génération en 2024 (« équivalent H100 ») pendant un peu plus de quatre mois. Les modèles de la prochaine génération (dès 2024) nécessiteront probablement environ dix fois plus de puissance de calcul pour l'entraînement, soit 1 000 équivalents H100, et autant pour l'inférence. Les entreprises affinant des modèles auront besoin de 10 fois moins de puissance pour l'entraînement (100 GPU) et pour l'inférence (100 GPU). Les entreprises utilisant seulement les modèles auront besoin de 100 fois moins de puissance de calcul (10 GPU), concentrée sur l'inférence.

La France pourrait donc se fixer comme objectif en 2024 de sécuriser au minimum l'installation de 30 000 processeurs graphiques équivalents H100 par des acteurs privés. Cela représenterait 3 % de la production mondiale prévue par NVIDIA en 2024, ce qui correspond par ailleurs à la part de la France dans le PIB mondial. Si cet objectif était tenu, la France disposerait d'une puissance de calcul privée de 21 MW ou près de 2 *exaflops*[120].

Avec un taux d'utilisation de 60 %, ces 30 000 GPU auraient une consommation électrique de 110 GWh, soit l'équivalent de la consommation annuelle de 50 000 personnes ou 0,22 % de nos exportations d'électricité en 2023. Avec une telle capacité de production électrique bas carbone, la France pourrait doubler ou tripler cet objectif afin d'accueillir de la puissance de calcul pour ses voisins européens. Au niveau européen, le même raisonnement conduirait à fixer un objectif de 150 000 GPU équivalent H100 (ou 105 MW).

Enfin, à la livraison du supercalculateur « Jules Verne », la France disposera d'une puissance de calcul publique d'un peu plus d'un *exaflops*. Étant donné les délais, des travaux sont donc à engager dès maintenant pour augmenter la puissance de calcul publique après « Jules Verne ».

Cet objectif n'inclut pas des besoins de puissance de calcul pour un géant du numérique. Leurs besoins sont tels qu'accueillir simplement un ou deux centres de calcul pour l'IA de l'un d'entre eux viendrait aisément doubler le nombre affiché ici. La France pourrait engager des discussions avec l'un d'entre eux pour accueillir l'un de leurs centres de calcul. De manière générale, ces chiffres sont à prendre comme une projection à court terme, tant les besoins croissent rapidement.

120. Unité de mesure utilisée pour calculer la puissance d'un processeur, qui correspond à un milliard de milliards d'opérations en virgule flottante exécutées en une seconde.

Dans cette perspective, nous proposons d'agir simultanément du côté de l'offre et de la demande de calcul. Du côté de l'offre, nous recommandons à la fois d'accélérer les travaux d'extension des supercalculateurs français et européens de taille « *exascale* », de lancer à court terme une opération d'achat groupé pour l'écosystème et de fixer un objectif pour l'implantation de centres de calcul sur le territoire européen, assorti à la fois d'une garantie publique d'utilisation de la puissance de calcul et d'un accompagnement à l'implantation et au raccordement électrique. Cet appel d'offre vise à la fois l'augmentation de la puissance de calcul – pour répondre aux besoins croissants – et la diversification des solutions d'hébergement – pour faire émerger des solutions européennes. Il n'est donc pas fermé aux solutions de fournisseurs extra européens. Du côté de la demande, un crédit d'impôt IA soutiendrait les projets de recherche et de développement dans la location de la puissance de calcul, sous la condition d'utiliser un centre de calcul établi sur le territoire. Il s'agit ici moins de subventionner généralement le calcul que d'inciter à l'installation de puissance de calcul sur notre territoire.

Recommandation n° 14

Faire de la France un pôle majeur de la puissance de calcul, à court comme à moyen terme.

2.2.3 ACCÉDER À DES DONNÉES DE QUALITÉ

L'IA a besoin de données en quantités considérables, et l'IA générative aussi, même si elle a plus besoin de données culturelles que de données personnelles. L'évolution de ces modèles de plus en plus précis, la volonté d'économiser de la puissance de calcul et l'épuisement des données disponibles (pour le texte les plus grands modèles ont déjà utilisés la plupart des corpus existants) conduiront vraisemblablement à l'avenir à mettre l'accent sur la qualité davantage que la quantité de données. En Europe, en France, l'enjeu n'est pas seulement technique, mais culturel, de présence et de découvrabilité de la langue, des images, des vidéos.

2.2.3.1 Les données personnelles

L'intelligence artificielle permet d'appréhender la masse de données disponibles, que l'intelligence humaine ne peut plus embrasser. Par exemple, plus de 5 millions d'articles scientifiques sont publiés chaque année, dont la moitié dans le seul domaine de la recherche médicale. Il est évidemment impossible qu'un chercheur ou une équipe de chercheurs, même de haute volée, puisse les lire, et encore moins les évaluer et les analyser.

Inversement, les données constituent un ingrédient indispensable aux développements récents de l'intelligence artificielle. Ainsi, la découverte d'un nouvel antibiotique permettant de lutter contre le staphylocoque doré[121] (voir *2.1.3. Mieux soigner grâce à l'IA*) n'a été rendue possible qu'après avoir entraîné un système d'IA avec près de 40 000 résultats déjà connus de structures moléculaires d'antibiotiques existants. Ce chiffre est d'ailleurs très faible au regard des capacités d'apprentissage et de traitement de tels systèmes ; cela tient en particulier à la qualité des

121. Wong, F., *et al.* (2023) « Discovery of a structural class of antibiotics with explainable deep learning », *Nature*

données utilisées pour cet apprentissage. Pour le dire autrement, la qualité des données utilisées pour l'entraînement est au cœur de la fiabilité d'un système d'IA.

Ces données ne sont pas nécessairement personnelles, mais force est de constater que beaucoup de données intéressantes pour entraîner des IA ont un caractère personnel. Dans la santé, bien entendu, mais pas uniquement. Même l'IA générative, a priori plus intéressée par les données culturelles, peut en avoir besoin pour développer une capacité d'interaction spécifique. Dans l'éducation, entraîner un modèle capable d'interagir de façon crédible et pertinente avec un élève nécessitera probablement un entraînement sur des données de dialogue entre élèves et enseignants, qui sont des données personnelles.

Exploiter le potentiel de l'intelligence artificielle et permettre son déploiement au service de l'humain exige par conséquent que les chercheurs, les développeurs et les innovateurs disposent d'un accès à des données massives, fiables, aisément manipulables et dont la représentativité et la qualité peuvent être évaluées. Dans un contexte d'évolution technologique rapide et de concurrence accrue, cet accès doit en outre pouvoir leur être ouvert rapidement et les données être utilisées sans contraintes excessives, au risque de favoriser davantage encore les acteurs en place ou de voir d'autres s'approprier nos recherches et nos innovations, en nous devançant dans leur expérimentation et leur diffusion.

Or en l'état, les difficultés d'accès aux données et les contraintes régulièrement regardées comme excessives liées à leur utilisation sont des constats très largement partagés par les acteurs de l'IA, quels qu'ils soient (entreprises, chercheurs, laboratoires, institutions publiques et privées, associations…). Ces difficultés sont de deux ordres.

En premier lieu, **certaines règles et pratiques françaises sont plus contraignantes que le cadre européen** en matière de traitement de données personnelles. Le cadre actuel est défini par le règlement général sur la protection des données (RGPD), entré en vigueur en 2018.

Le RGPD a renversé complètement la logique du droit qui prévalait en France depuis la loi « informatique et liberté » de 1978. Alors que la possibilité de traiter des données à caractère personnel reposait sur des procédures d'autorisation ou de déclaration préalables auprès de l'administration, le RGPD a posé les principes de liberté et de responsabilité : les acteurs sont libres de créer et de mettre en œuvre des traitements de données à caractère personnel, sous réserve de veiller eux-mêmes à la conformité de ces traitements aux principes et règles prévus par le règlement européen. Ils doivent en particulier analyser les risques spécifiques que peuvent créer les traitements les plus sensibles et prendre les mesures appropriées pour y remédier. En contrepartie de cette liberté, instituée dans le but précis de favoriser l'innovation, les exigences de protection des données personnelles ont été renforcées, de même que les pouvoirs de contrôle et de sanction *a posteriori* des autorités en charge de la protection des données. En France, il s'agit de la Commission nationale de l'informatique et des libertés (Cnil).

En France, cette évolution n'a pas été conduite jusqu'à son terme. Il demeure des procédures d'autorisation préalables non prévues par le droit européen[122]. C'est en particulier le cas pour l'accès aux données de santé pour la recherche. Une procédure simplifiée de déclaration de conformité à des méthodologies de référence existe mais elle est loin d'être généralisée. En pratique, la procédure simplifiée reste l'exception par rapport à la procédure d'autorisation préalable[123] car le moindre écart par rapport à ces méthodologies implique d'en passer par une autorisation préalable qui peut impliquer jusqu'à trois niveaux d'autorisation préalable. On

122. Le législateur national a mobilisé des marges de manœuvre ouvertes par le RGPD aux États membres dans le sens du maintien, voire du renforcement de procédures préalables d'autorisation pour les données de santé.

123. Fédérer les acteurs de l'écosystème pour libérer l'utilisation secondaire des données de santé, rapport de la mission conduite par Jérôme Marchand-Arvier (décembre 2023).

trouve des lourdeurs analogues dans les domaines de l'ordre public, de la sécurité et de la justice[124].

En second lieu, nous observons un **décalage croissant entre la logique centrée sur la protection de *l'individu* et l'évolution des modes d'utilisation *collective* des données**. Alors que le droit de la protection de l'accès et de l'utilisation des données est, depuis ses origines, centré sur la personne (les données personnelles), le développement et l'utilisation des systèmes d'IA se concentrent sur des données massives, agrégées et dynamiques.

Sur le plan technique, plusieurs notions clés du RGPD sont ainsi d'un maniement malaisé face au fonctionnement de l'IA. Tel est notamment le cas du « responsable du traitement », pour laquelle la répartition des responsabilités entre le développeur qui a procédé à l'entraînement d'une IA générative et qui la met à disposition de tiers et l'utilisateur final du système pour ses propres besoins n'apparaît pas forcément aller de soi. La notion de finalité du traitement, qui conditionne la nature des données pouvant légalement être utilisées et sur laquelle porte le consentement des personnes concernées est également plus complexe à appréhender, eu égard aux nombreuses utilisations possibles d'une IA générative une fois celle-ci entraînée.

Sur le plan des principes, la notion même de donnée personnelle, qui constitue la clé d'application du RGPD, suscite des interrogations dans un contexte croissant d'utilisation de données collectives. En l'état, juridiquement, seul un processus d'anonymisation des données personnelles permet de « sortir » du régime de protection des données personnelles du RGPD. Or, la technologie ouvre de plus en plus loin des possibilités de réidentification de données anonymisées. Par ailleurs, une gestion plus collective des données (aujourd'hui embryonnaire dans le RGPD) pourrait améliorer la protection des intérêts et l'exercice des droits, notamment face aux acteurs mondiaux de la donnée. En effet, que ce soit au travers d'associations, de syndicats ou de tout autre collectif organisé, des accords relatifs aux traitements des données et à l'utilisation de systèmes d'IA pourraient permettre d'accroître l'effectivité de la garantie des droits de chacun.

Recommandation n° 15

Transformer notre approche de la donnée personnelle pour protéger tout en facilitant l'innovation au service de nos besoins.

Notre Commission recommande de poursuivre la modernisation de notre approche de la donnée en conjuguant mieux protection et innovation. Des évolutions sont en cours. La Cnil a par exemple créé en 2023 un service de l'IA et de la publication, et publie depuis quelques mois des « fiches pratiques IA » destinées à guider les développeurs de systèmes d'IA. L'autorité consolide ainsi une mission d'accompagnement des acteurs de l'innovation.

Pour aller plus loin, nous recommandons notamment de supprimer des procédures d'autorisation préalable d'accès aux données de santé et de réduire les délais de réponse de la Cnil. Ce mouvement devrait s'accompagner d'une réforme du mandat qui est confié à la Cnil, pour y intégrer un objectif d'innovation. Cette évolution impliquera un ajustement de la composition de son collège, pour qu'une palette plus large de compétences soit représentée (innovation,

124. Une procédure spécifique d'autorisation du traitement par décret en Conseil d'État de tous les traitements mis en œuvre pour le compte de l'État dans les domaines de l'ordre public, de la sécurité et de la justice a été maintenue, rendant particulièrement lourde toute évolution, même minime des modalités de traitement des données personnelles dans ce champ.

recherche…), et un renforcement de ses moyens de fonctionnement. Plus largement il conviendra de renforcer la coordination de l'ensemble des régulateurs intervenant dans le champ de la donnée et du numérique.

Au-delà, il importe de trouver la voie d'une gouvernance collective de la donnée qui pourrait, dès aujourd'hui, utiliser les marges de manœuvre juridiques sous-exploitées du RGPD et, à terme, poser les jalons d'une évolution du cadre juridique qui prendrait mieux en considération l'évolution des modes d'utilisation des données. En appelant à une amélioration de la gouvernance de l'*open data* du secteur public et à la mise en œuvre immédiate des nouveaux règlements européens relatifs à la donnée, c'est à la pleine exploitation d'une formidable ressource d'intérêt général et porteuse de croissance économique qu'appelle notre Commission.

2.2.3.2 Les données patrimoniales

La France dispose d'un ensemble hors du commun de corpus en partie numérisés depuis le début des années 2000. C'est un atout cultivé depuis François Ier, grâce à une politique continue, par un dépôt légal devenu universel (livres, revues, presse, images, sons, cartes, cinéma, jeux vidéo) étendu à la télévision et la radio, mais encore à travers des politiques d'acquisitions et de conservation. Ce corpus représente linguistiquement, artistiquement, culturellement un paysage intellectuel et émotionnel très large de la francophonie.

Que se passe-t-il dans le reste du monde ? Les exemples de l'Inde et de la Corée du Sud[125]

En Inde, l'IA est identifiée comme un levier pour renforcer l'inclusion sociale et économique. Le pays entend être en tête des solutions d'IA appliquées et s'attache à rassembler des données de qualité. Le programme « Bhashini » vise notamment à mettre à la disposition des entreprises des ensembles de données audio et textuelles provenant des 22 langues nationales de l'Inde, afin qu'ils puissent développer des solutions innovantes.

En Corée du Sud, le ministère de la culture et l'institut national de la langue coréenne prévoient de construire un corpus coréen de données de haute qualité d'environ 120 millions de syntagmes pour soutenir le développement de « K-Chat GPT qui parle bien le coréen », puis de l'étendre à 1 milliard de syntagmes d'ici 2027.

La mobilisation adaptée de ce patrimoine unique dans le cadre de l'entraînement des systèmes d'intelligence artificielle représente donc un enjeu de diversité culturelle et de souveraineté. A l'heure actuelle, les modèles d'IA sont moins performants en français qu'en anglais, ayant principalement été entraînés sur des corpus anglais.

Répondre à cet enjeu suppose des évolutions rapides. Si une large part des contenus est déjà disponible, numérisée, accessible par API, une autre partie justifie une accélération des financements de mise à disposition, de libération des droits qui peuvent exister, comme les droits des photographes. La plupart des institutions comme les Archives ont vocation à être

125. Source : direction générale du Trésor, service économique en Inde et en Corée du Sud.

ouvertes et disponibles numériquement. Cependant, des enjeux d'infrastructures de mise à disposition doivent être pris en charge pour accélérer l'accès.

Des freins juridiques demeurent. Ainsi du dépôt légal, alimenté quotidiennement par des publications sous droits issus des éditeurs d'imprimés, de phonogrammes, de producteurs de films ou de programmes audiovisuels. Ce dépôt ne saurait servir de fonds ouvert pour la mise à disposition des IA. D'ailleurs, nos engagements internationaux s'y opposent. En revanche, les métadonnées sont susceptibles d'être ouvertes. Plus largement, des modèles économiques restent à concevoir selon les usages de recherche ou commerciaux.

Pour lever certaines de ces barrières, un vaste plan de mise à disposition est nécessaire avec des financements à l'échelle de cet objectif. Il devra composer avec des particularités tenant aux standards internationaux propres à chaque catégorie de contenus.

Trois pistes doivent être explorées. D'abord, il importe de créer un cadre de confiance pour l'accès aux contenus libres de droits, notamment par la mise à disposition des métadonnées. Ensuite, il est nécessaire de préparer la création d'un registre des contenus à disposition (et incluant toutes les informations utiles, dont les modalités d'accès), pour lequel la Bibliothèque nationale de France, l'Institut national de l'audiovisuel et le ministère de la culture pourraient constituer le tiers de confiance. Une plateforme pourrait alors mettre en relation les grands détenteurs de données culturelles publiques, voire des détenteurs de données culturelles privées qui le souhaiteraient, et des développeurs d'IA. Enfin, il s'agit de construire des outils publics à disposition d'acteurs privés, par exemple pour lutter contre la désinformation en permettant par exemple des vérifications sur les informations audiovisuelles passées.

Recommandation n° 16

Mettre en place une infrastructure technique favorisant la mise en relation ntre les développeurs d'IA et les détenteurs de données culturelles patrimoniales.

Au-delà, **les données publiques demeurent insuffisamment ouvertes**. Au niveau mondial, le mouvement pour l'ouverture des données publiques a commencé à la fin des années 2000 et a eu son apogée au milieu des années 2010, pour ensuite s'estomper. Il nous faut revitaliser ce mouvement, dans un contexte nouveau d'IA générative.

Le secteur public représente effectivement une source de données à la fois nombreuses et de qualité. Leur réutilisation, notamment grâce à l'IA, peut permettre la création de très nombreux services nouveaux, que ce soit pour des activités de nature commerciale ou d'intérêt général. L'ouverture encore trop limitée des données publiques[126] prive la France de nombreux bénéfices.

On peut penser par exemple aux données de santé pour la recherche, aux données relatives au trafic routier qui peuvent permettre d'optimiser l'utilisation des moyens de transport, aux données agricoles pour améliorer les rendements et réduire l'utilisation d'intrants, etc. La circulation des données à l'intérieur même du secteur public est source de simplification des procédures administratives, en évitant de multiplier les dossiers administratifs et la production de justificatifs, et de meilleure effectivité des politiques publiques. À titre d'illustration, elle permet d'identifier les personnes remplissant les conditions pour bénéficier de prestations

126. Voir par exemple le rapport de la mission d'E. Bothorel, *Pour une politique publique de la donnée*, décembre 2020.

sociales et dès lors d'éviter les phénomènes de « non recours » faute de connaissance ou de technicité suffisante des bénéficiaires.

Pour les données publiques, les obstacles rencontrés sont pratiques et non d'ordre juridique. Le principe en vigueur en Europe depuis 2003[127] est effectivement celui de l'ouverture et de la libre réutilisation des données élaborées ou détenues par l'ensemble des organismes publics et appartenant au secteur public. Les très récentes modifications du règlement européen sur la gouvernance des données[128] visent d'ailleurs à étendre le champ de ce principe et à en faciliter l'application, notamment en prévoyant une plateforme européenne unique destinée à la réutilisation des données publiques.

2.2.3.3 Les données protégées par un droit de propriété littéraire ou artistique

Un mouvement global de contestation de l'utilisation des données protégées pour l'entraînement de l'IA est en voie de naître des deux côtés de l'Atlantique.

Aux États-Unis, l'utilisation de données protégées est soumise à l'autorisation du titulaire de droits sauf à ce que l'usage soit considéré comme équitable ou loyal (« *fair use* »). L'application de ce cette exception est contestée par les auteurs et les artistes qui ont intenté de nombreux procès aux éditeurs de modèles. Le résultat de ces actions en justice sera déterminant, mais aussi incertain, car l'appréciation du « *fair use* » est très contextuelle. Dans ces conditions, certains acteurs de l'IA — les mêmes qui contestaient au départ que le droit d'auteur (« *copyright* ») s'applique — cherchent déjà à contractualiser l'utilisation des contenus protégés en passant des accords de licence avec les titulaires de droits.

En Europe, l'utilisation de données protégées pour l'entraînement d'un système d'IA générative est encadrée par l'article 4 d'une directive européenne de *2019* (« droit d'auteur et droits voisins dans le marché unique numérique » ; notons qu'elle est antérieure au foisonnement des modèles d'IA générative). Cette directive impose notamment une exception de fouille de textes et de données (dite *text and Data Mining* - TDM), transposée en droit français en 2021[129]. Le monopole est donc suspendu lors de ces opérations, pour toutes finalités (y compris commerciales) et au bénéfice de tous acteurs.

Des conditions sont cependant exigées pour bénéficier de l'exemption. D'abord, les données doivent avoir été obtenues de manière licite. Or, il est désormais avéré que des modèles d'IA générative ont été entraînés sur des bases comprenant des contenus piratés. Ensuite, en contrepartie de l'atteinte à ses droits exclusifs, le titulaire peut décider de refuser l'utilisation de ses données en exerçant un droit d'opposition (« *opt out* »). Ce retour à l'exclusivité peut lui permettre de recouvrer une capacité de négociation en vue d'obtenir une rémunération.

Confrontés à l'utilisation croissante de leurs données, les ayants droit français multiplient ces stratégies d'*opt out* aux conséquences préoccupantes : affaiblissement de la fiabilité des résultats produits par les IA, absence de contenus français et plus généralement de créations authentiques pouvant entraîner la production de résultats stéréotypés et médiocres par les modèles d'IA générative accessibles au public.

127. Directive 2003/98/CE du Parlement européen et du Conseil du 17 novembre 2003 concernant la réutilisation des informations du secteur public. Le principe figure aujourd'hui dans la directive (UE) 2019/1024 du Parlement européen et du Conseil du 20 juin 2019 concernant les données ouvertes et la réutilisation des informations du secteur public.

128. Règlement (UE) 2022/868 du Parlement européen et du Conseil du 30 mai 2022 portant sur la gouvernance européenne des données et modifiant le règlement (UE) 2018/1724 (règlement sur la gouvernance des données)

129. Ordonnance n° 2021-580 du 12 mai 2021 portant transposition du 6 de l'article 2 et des articles 17à 23 de la directive 2019/790 du Parlement européen et du Conseil du 17 avril 2019 sur le droit d'auteur et les droits voisins dans le marché unique numérique et modifiant les directives 96/9 et 2001/29/CE.

La situation actuelle est donc contentieuse. Les titulaires de droits demandent le respect de leurs droits, mais étaient diminués dans leur action du fait de l'absence de transparence sur les données d'entraînement utilisées. Le règlement européen sur l'intelligence artificielle (*AI Act*[130]) prévoit cette transparence tant en amont, s'agissant des sources d'entraînement, qu'en aval, pour lutter contre la désinformation.

En amont, les titulaires de droits doivent en effet avoir la possibilité d'être informés quant aux données utilisées afin de s'assurer qu'elles ont été obtenues de manière licite. À l'inverse de la tendance actuelle, l'entraînement des modèles d'IA générative ne saurait être un nouveau débouché durable offert aux entreprises de piratage des données protégées par « *scrapping* » ou « *crawling* ». La transparence est également essentielle pour permettre aux titulaires de droits d'exercer leur droit d'*opt out* avec discernement, le cas échéant d'autoriser contre rémunération et contrôler le respect des engagements des parties. Sans ce principe, le droit qui leur est conféré en contrepartie de l'exception de TDM est privé de son sens, faute de pouvoir être rendu opérationnel.

Concrètement cela signifie que l'éditeur d'IA doit rendre public le fait qu'il a utilisé des données protégées par le droit d'auteur et indiquer auprès de quelle entité le titulaire des droits peut obtenir des informations. Ces données peuvent être rendues publiques ou communiquées sur demande à des personnes ayant intérêt à agir s'il est nécessaire de préserver leur confidentialité. Le règlement européen sur l'IA ne prévoit pas que cette obligation s'applique aux éditeurs de modèles spécialisés (mais seulement aux modèles généraux). Il sera donc nécessaire d'évaluer régulièrement, à la lumière des évolutions technologiques et des modalités d'entraînement des modèles, la pertinence de ce champ d'application de l'obligation de transparence.

La plateforme de mise en relation précédemment identifiée pourrait utilement intégrer l'opposition éventuelle exercée par les titulaires de droits et, le cas échéant, les conditions d'une utilisation pour l'entraînement (licence, rémunération…). À terme, l'idée est de pouvoir constituer facilement des corpus de données de qualité pour l'entraînement des modèles d'IA dans le respect des droits. Tous ces corpus n'ont pas la même valeur, et la valeur marginale d'une œuvre ou d'un auteur en particulier peut être assez faible. En revanche, leur valeur collective peut être importante. C'est le cas du dépôt légal de l'Institut National de l'Audiovisuel, qui représente 47,85 Pétaoctets de données multimodales de qualité, soit 100 fois plus que la taille de *Common Crawl*, « sauvegarde » partielle d'internet.

En aval, les produits culturels et l'information issue des IA doivent être clairement labellisés. Les contenus générés par une IA doivent être reconnaissables. Cette exigence qui s'affirme au plan international est essentielle pour s'attaquer au parasitisme, identifier les informations non fiabilisées et lutter contre les *deepfakes*. Elle est également indispensable à l'information des consommateurs qui, comme par exemple pour les aliments, ont un droit à l'information sur la méthode, le lieu et les ingrédients de fabrication des produits. Enfin, la transparence en aval est indispensable dans les relations entre clients et fournisseurs, pour gérer les éventuelles responsabilités. C'est donc ici une transparence plus large que celle évoquée dans le règlement sur l'IA qui est soutenue.

130. Accord politique de la proposition de règlement voté en Coreper en février 2024.

Recommandation n° 17

Mettre en œuvre et évaluer les obligations de transparence prévues par le règlement européen sur l'IA en encourageant le développement de standards et d'une infrastructure adaptée.

La qualité, la diversité et la profondeur des données issues de la culture européenne sont à l'évidence un atout pour les éditeurs d'intelligence artificielle. La mise en œuvre de règles de transparence doit permettre l'émergence d'un marché des données « blanches » dont l'usage par les éditeurs de logiciel est licite et sûr d'un point de vue juridique. Ce peut être un atout face aux États-Unis dont les règles de « *fair use* » feront l'objet de contestations devant les tribunaux pendant encore de nombreuses années. Une forte mobilisation collective en Europe serait essentielle pour encourager la création d'entrepôts de données publiques et privées assorties de cartographies des droits et de modèles de rémunération. Seule cette mobilisation permettra d'éviter les trois risques devant nous : la concentration (si seuls quelques éditeurs d'IA très bien financés peuvent accéder à des données culturelles au prix fort), la submersion culturelle (si les IA sont entraînées sur des données purement américaines) ou le contournement du droit de propriété (si se développe un marché « noir » ou « gris[131] » de la donnée culturelle).

2.2.4 ATTIRER LES TALENTS POUR CONSTRUIRE LES TECHNOLOGIES ET LES USAGES DE DEMAIN

2.2.4.1 Faire venir les plus qualifiés du monde sur notre territoire

Les talents sont un des piliers majeurs pour accélérer le développement de l'IA, créer un écosystème dynamique et contribuer à la souveraineté de la France. La notion de talent est prise au sens large comme toute personne possédant des compétences rares qui peut contribuer au développement de l'IA que ce soit dans sa dimension technique, scientifique ou commerciale. Les talents constituent la principale ressource de la recherche publique comme des entreprises du domaine, dont ils représentent environ un tiers des coûts.

Or, la France est en forte concurrence avec les États-Unis, dans une moindre mesure avec le Royaume-Uni, et les écosystèmes français ne sont pas (encore) aussi réputés. De nombreux jeunes français diplômés, notamment des écoles d'ingénieurs, s'expatrient pour y travailler. Dans les entreprises américaines, les services spécialisés en IA comptent ainsi de nombreux Français, soulignant certes la qualité de la formation française, mais aussi le manque d'attractivité de notre pays.

131. Sans accord entre ayants droit et éditeurs d'IA, certains évoquent déjà la possibilité de produire des bases de données synthétiques reproduisant les caractéristiques de bases de données culturelles obtenues illégalement, sans contenir directement d'œuvres protégées.

Que se passe-t-il dans le reste du monde ? L'exemple des États-Unis[132]

Aux États-Unis, le décret du 30 octobre 2023 met en place un ensemble de mesures visant à attirer les plus qualifiés du monde dans le domaine de l'IA. D'abord, il simplifie les processus d'instruction des demandes de visas et en priorisant l'instruction de ces demandes. Ensuite, il révise les critères d'octroi rapide de visas (J-1). Enfin, il met en œuvre un programme global pour identifier et attirer sur le sol américain les talents étrangers en IA.

La situation s'est améliorée ces dernières années, sous l'effet de l'implantation à Paris de laboratoires de recherche d'entreprises américaines. Celles-ci ont recruté à Paris de jeunes Français, qui ont par exemple ensuite créé une *start-up* en France. Cette dynamique a contribué à faire émerger un écosystème français de l'IA et à renforcer son attractivité. Une entreprise étrangère spécialisée dans la génération de code grâce à l'IA est par exemple en cours de localisation à Paris.

La dynamique est néanmoins insuffisante. Notre Commission considère que parmi les trois à cinq mille profils internationaux très qualifiés susceptibles d'avoir un impact significatif sur la croissance de l'écosystème de l'IA, la France doit en attirer entre 10 et 15 %. Les compétences recherchées sont diverses. Il faut attirer les talents capables de développer les grands modèles de langue ainsi que les futures générations d'IA. Il s'agit également d'attirer les meilleurs chercheurs et ingénieurs du domaine pour créer de nouveaux produits et services, mais aussi des dirigeants d'entreprises innovantes en croissance. Il en est de même pour les ingénieurs-produits qualifiés dans le domaine de la commercialisation.

La diversité des profils favorisera l'innovation, mais aussi limitera les risques de biais, encourus lorsque les équipes affichent des profils trop similaires. Cet enjeu de diversité est particulièrement important en France où la part des ingénieurs étrangers est faible[133] par rapport à la situation dans la *Silicon Valley* ou à Londres. L'attractivité et la diversité de l'écosystème d'innovation reposeront donc également sur la venue d'étudiants étrangers, issus des cycles de licence réputés à l'étranger, afin qu'ils poursuivent leurs études puis débutent leur carrière en France.

Recommandation n° 18

Attirer et retenir des talents internationaux avec des compétences scientifiques, entrepreneuriales et managériales dans le domaine de l'IA.

L'attractivité des talents va de pair avec celle des entreprises. Ces dernières peuvent en effet jouer un rôle de catalyseur dans l'écosystème français, notamment en attirant des profils qualifiés étrangers sur notre territoire. En complément d'un ensemble d'actions visant à soutenir la croissance des *start-up* françaises, les mesures d'attractivité des groupes étrangers sur notre territoire peuvent donc participer à la souveraineté de la France et de l'Europe.

132. Source : direction générale du Trésor, service économique aux États-Unis.
133. La part des ingénieurs étrangers en Île-de-France est estimée à moins de 10 %. Elle est de 68 % dans la Silicon Valley.

Notre Commission recommande la création d'une mission « attractivité des talents en IA », inspirée de celle mise en place pour attirer en France les profils financiers en prévision de la sortie du Royaume-Uni de l'Union européenne (*Brexit*). Cette mission consiste en une démarche ciblée et personnalisée d'identification d'entreprises et de talents à l'étranger. Elle doit être rapidement mise en place pour bénéficier de l'actuel élan de l'écosystème français d'IA et de sa réputation croissante à l'étranger.

En dehors de la recherche publique, recruter des talents individuels relève de la compétence des entreprises. Pour autant, l'État doit mettre en place toutes les conditions pour faciliter l'installation des profils qualifiés notamment *via* une aide pour les démarches administratives : visa, scolarité des enfants, communication sur le régime fiscal des impatriés, etc. Ces conditions seront également offertes aux salariés des entreprises étrangères qui s'installent et aux étudiants qui auront été démarchés.

2.2.4.2 Permettre aux chercheurs de chercher

Il est essentiel de disposer d'une recherche publique dans l'IA de haut niveau en France. La recherche publique permet d'explorer des pistes de plus long terme, qui ne répondent pas nécessairement à des objectifs économiques immédiats. Elle garantit également la qualité de formation des futurs ingénieurs et chercheurs qui rejoindront le secteur privé. Elle constitue un facteur d'attractivité.

À la suite du rapport de C. Villani[134] (2018), qui soulignait cet enjeu, plusieurs mesures ont été prises. Plusieurs lois[135] ont facilité le recours aux contrats à durée déterminée ou indéterminée (CDD ou CDI), assoupli le régime de primes, simplifié la déclaration de cumul d'activités[136], renforcé les relations avec les entreprises, etc. En particulier, la possibilité de proposer dans la recherche publique des CDD ou des CDI avec des conditions de rémunération supérieures à celles des grilles de l'État a facilité le recrutement de jeunes chercheurs ou de chercheurs étrangers renommés.

Cependant, les obstacles pratiques demeurent nombreux. En matière de rémunération, par exemple, les chercheurs rencontrés par notre Commission estiment qu'ils finissent presque toujours par trouver des solutions de compléments de rémunération, mais au prix de montages complexes, fastidieux et longs à mettre en place. Alors que la recherche publique ne peut rivaliser avec les entreprises pour les conditions salariales (les salaires sont communément dix fois plus élevés), la simplicité du recours aux compléments de rémunération est indispensable : les lourdeurs administratives découragent les chercheurs.

Renforcer l'attractivité de la recherche publique française est non seulement impératif mais aussi accessible. Celle-ci offre toujours plusieurs conditions qui restent appréciées des chercheurs, comme l'accès à des étudiants de qualité ou la liberté académique des chercheurs (ce point distinguant notamment la France des États-Unis). Dans le domaine de l'intelligence artificielle (pris largement), une rémunération supérieure à 80 000 € par an, plus ou moins élevée en fonction de l'âge et de la renommée, doit être atteinte sans aucune pesanteur administrative. Au-delà des chercheurs en poste, il faut également prévoir des compléments de rémunération aux doctorants et post-doctorants, qui contribuent largement à la recherche publique et dont la qualité constitue un facteur d'attractivité pour des chercheurs plus matures.

134. Cédric Villani, mars 2018, Donner du sens à l'intelligence artificielle : pour une stratégie nationale et européenne
135. En particulier la loi de programmation de la recherche et la loi pour la croissance et la transformation des entreprises.
136. En dépit de la simplification, de nombreux établissements continuent à exiger des autorisations qui doivent passer par de nombreuses strates de signatures.

D'innombrables autres complexités administratives, chronophages et disproportionnées, doivent être combattues. La validation des embauches, la signature de contrats avec des entreprises, le montage des dossiers administratifs pour les réponses aux appels à projets ou encore la création d'une formation occupent au moins un tiers de l'emploi du temps d'un chercheur dans le public. Ce sont autant d'activités longues, laborieuses, décourageantes et à faible valeur ajoutée par rapport aux compétences des chercheurs. Elles ralentissent donc la production de nouvelles connaissances. Dans un monde de l'innovation en IA qui progresse chaque jour, ces complexités sont impensables.

Enfin, les conditions de travail ne sont pas à la hauteur des enjeux et des standards internationaux. À titre d'illustration, les étrangers recrutés dans la recherche publique française ne bénéficient pas d'un appui global pour faciliter leur installation (déménagement, accompagnement pour les démarches administratives, aides à l'embauche pour les conjoints, inscriptions dans les écoles pour les enfants…). Le dispositif *Choose Paris Region* a permis quelques avancés, mais il est nécessaire d'aller plus loin.

Si la France a mis en place une réflexion sur la simplification des mesures administratives dans la recherche, le calendrier ne s'inscrit pas dans celui du développement de l'IA au niveau mondial. De ce fait, il est proposé d'aller plus vite pour l'intelligence artificielle en mettant en place une « Exception IA ». Sous la forme d'une expérimentation, cette exception vise un objectif de « zéro entrave pour les chercheurs », notamment par un engagement sur les délais de réponse aux sollicitations des chercheurs. L'exception IA doit également permettre de rehausser la rémunération et de faciliter les temps partiels avec des entreprises ou d'autres acteurs socio-économiques de l'IA. Au-delà, il est indispensable de doubler au moins les financements de la recherche publique en IA pour rejoindre le *top* 5 mondial en termes de production et de publication en IA.

Recommandation n° 19

Assumer le principe d'une « Exception IA » sous la forme d'une expérimentation dans la recherche publique pour en renforcer l'attractivité.

2.2.5 DÉPLOYER MASSIVEMENT L'INTELLIGENCE ARTIFICIELLE DANS NOTRE ÉCONOMIE

2.2.5.1 Recourir à l'IA pour gagner en productivité et en compétitivité

L'écosystème français et européen spécialisé dans l'intelligence artificielle doit se développer parallèlement à l'adoption massive des systèmes d'IA par les entreprises qui composent actuellement notre tissu productif. Il faut donc miser sur l'économie de l'IA, mais aussi sur l'économie avec l'IA. Le potentiel de gains de productivité est effectivement immense[137]. Les particularités de l'IA générative (réalisme, simplicité, rapidité, aptitudes) ouvrent la voie à d'importants gains de productivité rapides, presque immédiats. Cependant, les principaux bénéfices nécessiteront des actions de transformation (des infrastructures de données, de l'organisation, des ressources humaines, etc.).

L'impact des précédentes innovations majeures sur la productivité française est souvent considéré comme modéré. Le retard à l'adoption de ces technologies par rapport aux autres pays, notamment les États-Unis, constitue une des raisons principales. Afin de tirer les bénéfices de l'IA, les entreprises françaises doivent donc adopter rapidement les systèmes d'IA. Si elles sont plus lentes dans cette adoption, elles perdront des parts de marché vis-à-vis des entreprises qui auront recours plus vite à l'IA. L'inverse est tout aussi vrai : si la France se mobilise plus vite, elle pourra gagner des parts de marché national ou à l'export.

Que se passe-t-il dans le reste du monde ?
La diffusion de l'IA en Europe et à l'étranger

L'IA est une technologie à usage général, sans mesure unique de son usage (comme pour l'électricité). Il est donc difficile de comparer la vitesse de diffusion de l'IA dans l'économie européenne, a fortiori de l'IA générative. Les publications de 2023 à ce sujet ne permettent par ailleurs pas toujours de distinguer ce qui relève de l'expérimentation du projet en production. Quoi qu'il en soit, les auditions menées par la Commission sont convergentes.

Premièrement, le dynamisme de l'écosystème IA est plus important aux États-Unis qu'en Europe (l'activité sur des plateformes de contribution y est 50 % plus importante). Mais cet écart s'explique principalement par la taille du secteur technologique aux États-Unis.

Secondement, l'adoption de l'IA générative est à peu près similaire aux États-Unis et en Europe lorsque l'on compare secteur par secteur. Elle est très largement liée à l'adoption du cloud dans les entreprises. Les grands fournisseurs de cloud proposent en effet tous des services d'IA et d'IA générative, facilitant leur expérimentation. Or la France est l'un des pays d'Europe où les entreprises utilisent le moins le *cloud* (30 % des entreprises, contre plus de 70 % dans les pays nordique, 60 % en Irlande et 45 % en Allemagne).

137. Voir précédemment « *L'IA nous rendra-t-elle plus prospère ?* »

Il revient donc à chaque entreprise française et européenne d'identifier les cas d'application les plus pertinents de l'IA. Les cas les plus simples se situeront probablement là où les erreurs de l'IA ont peu de conséquences, peuvent être facilement détectées et corrigées. Étant donné que les coûts d'inférence restent encore substantiels, il s'agira aussi de cas où la valeur du service rendu par l'IA est assez élevée. L'État n'a pas vocation à identifier ces cas d'usage à la place des acteurs économiques mais il peut favoriser la création d'une dynamique collective permettant un partage d'expérience et d'information entre ces acteurs. Cela pourrait prendre la forme d'une « Convention des entreprises à l'ère de IA », à l'image de la « Convention des entreprises pour le climat », permettant aux entreprises d'un même territoire ou d'une même filière de s'informer, d'expérimenter et de partager de bonnes pratiques et d'établir des feuilles de route de transformation avec l'IA.

Une impulsion pour créer une dynamique collective est nécessaire, mais ne sera pas suffisante pour favoriser une adoption généralisée de l'IA par tous les acteurs économiques. Celle-ci dépendra aussi largement de la rentabilité des systèmes d'IA. Au-delà de la baisse générale du coût de ces technologies, faciliter la collecte de données ou l'utilisation à grande échelle du même système d'IA sera crucial pour qu'il devienne rentable. Pour faciliter l'adoption de l'IA tout en évitant qu'elle ne soit conduite que par quelques acteurs à très grande échelle, la dynamique collective doit être complétée par une approche plus sectorielle (aéronautique, luxe, automobile, agriculture…) ou plus fonctionnelle (ressources humaines, finances). La puissance publique, État et Conseils régionaux en tête, pourrait ainsi pourrait ainsi soutenir davantage les projets communs de consortiums d'acteurs privés, réunissant des start-up de l'IA et des entreprises du tissu économique traditionnel.

Que se passe-t-il dans le reste du monde ? L'exemple du Royaume-Uni[138]

Au Royaume-Uni, le programme *BridgeAI* a été lancé en avril 2023 pour stimuler l'adoption de l'IA dans les secteurs de l'économie à faibles gains de productivité et ainsi rehausser leur productivité. Ce programme, doté de 100 M£, cible particulièrement les secteurs de l'agriculture, de la construction, des transports et des industries créatives.

Enfin, l'État doit davantage jouer un rôle de primo-adoptant de la technologie, c'est-à-dire contractualiser avec des entreprises, notamment petites, qui développent des solutions innovantes, capables d'avoir un impact majeur, mais qui ne sont pas encore stabilisées, faute de clients. Une telle démarche contribue également au développement de l'écosystème et permet à l'État d'avoir un rôle d'exemple dans l'adoption de l'IA, vis-à-vis des entreprises.

Recommandation n° 20

Inciter, faciliter et amplifier le recours aux outils d'IA dans l'économie française en favorisant l'usage de solutions européennes.

138. Source : direction générale du Trésor, service économique au Royaume-Uni.

2.2.5.2 Accélérer le déploiement de l'intelligence artificielle dans la culture

Les chaînes de valeur des secteurs culturels et des médias nationaux devront s'engager dans une transition rapide, parfois importante. La transition a débuté dans la presse (sélection, documentation, vérification, labellisation des contenus) ; elle est en cours dans les grandes agences d'information (*Associated Press*, Reuters, Agence France Presse) et de photographies (Getty Images) ; l'appropriation est rapide dans les grands groupes de communication (Publicis). Elle se déploie encore dans des secteurs culturels à forte intensité technologique, comme l'animation et bien sûr les jeux vidéo (Ubisoft ou Mac Guff). L'utilisation d'IA s'est portée principalement sur des tâches répétitives, engendrant des gains pouvant être investis vers plus de qualité d'information, de créativité et de différenciation.

Les usages de l'IA dans les autres secteurs sont variés. Certains sont moins sensibles – que l'on songe au spectacle vivant hors des aspects techniques (son, lumière). La plupart peuvent y recourir dans les processus créatifs, comme dans l'architecture et déjà dans le jeu vidéo avec de nouvelles interactions avec les joueurs. Les IA participent déjà à la production et la post-production pour les effets visuels et sonores dont l'emploi devrait progresser dans l'audiovisuel, le cinéma ou la musique (mixage, arrangement…). Cependant, des activités entières et des emplois sont évidemment affectés par l'essor de l'IA : traduction, doublage, figuration, photographie…

La recherche de gains de productivité se répand. Dans l'édition de livres ou de musique, elle tend à se développer (sélection de manuscrits ou musiques, détection de talents). Elle peut comme dans toute l'économie accroître l'efficacité de la distribution physique (livres, disques). Couplés avec la diffusion en ligne, ce sont les segments du marketing relationnel et individualisé qui semblent les plus prometteurs pour les IA à la recherche des goûts et préférences.

Cet emploi de l'IA n'est pas sans risque en termes de diversité culturelle, informationnelle et de découvrabilité. Car, c'est sur le segment de la commercialisation et du marketing que se joue la mutation la plus importante. L'IA peut conduire à raréfier et homogénéiser la création. Elle s'accompagne d'une forte polarisation des acteurs : d'un côté, un petit nombre d'acteurs puissants pour une offre de *blockbusters* pour de vastes publics ; de l'autre, un tissu plus fragile d'acteurs petits et moyens souvent au cœur du renouvellement de la créativité.

L'enjeu consiste donc à permettre la réussite du virage de l'IA grâce à l'accélération et la structuration d'un écosystème. Or, l'actuel écosystème de conception d'IA spécialisées existe, mais est naissant : génération et simulation de foules en 3D (Golaem), Face Engine (MacGuff) pour le traitement d'images de visages, colorisation d'archives de films (Composite Films), ChatBox de médiation culturelle (Ask Mona), transformation d'images (Photoroom), usages de voix DeepVoice de l'Ircam. Ces nouveaux outils ont pour champ de croissance les secteurs culturels et tout secteur qui fait appel à la création (publicité, luxe, industrie textile, *design*). Il convient de soutenir cet écosystème d'acteurs appuyés sur la recherche, mais aussi de prendre en compte le tissu industriel des ETI et PME culturelles en évitant une polarisation de l'économie de la culture et des médias.

Recommandation n° 21

Faciliter l'appropriation et l'accélération des usages de l'IA dans la culture et les médias pour limiter la polarisation entre grands groupes et petits acteurs et lutter contre la désinformation.

2.3. RESPONSABILITÉ : MAÎTRISER, AUDITER, PROTÉGER

2.3.1 BÂTIR UNE GOUVERNANCE INTERNATIONALE QUI FAIT AUJOURD'HUI DÉFAUT

La prise de conscience généralisée des progrès fulgurants réalisés dans le domaine de l'IA a suscité un foisonnement d'initiatives de portée internationale sur le thème de l'IA. Au total, on dénombre pas moins d'une cinquantaine d'initiatives auxquelles la France est le plus souvent partie !

L'année 2023 a été particulièrement riche en débats. Dans le cadre de l'ONU, un comité consultatif de haut niveau sur l'IA auprès du secrétaire général a été annoncé. Dans le cadre de l'Unesco, un nouveau sommet mondial sur l'éthique de l'intelligence artificielle a été préparé et s'est tenu début février dernier. Dans le cadre du G7, le processus dit de Hiroshima a conduit à plusieurs déclarations communes sur le développement et l'encadrement des systèmes d'IA. Par ailleurs, le Royaume-Uni a organisé un sommet sur les risques liés à l'IA en novembre 2023.

Ces événements prolongent un premier ensemble de travaux sur le développement de cette technologie. Au sein de l'Union européenne, le groupe d'experts de haut niveau sur l'IA mis en place en 2017 a posé les bases du règlement sur l'intelligence artificielle. Au sein de l'OCDE, un réseau d'experts et un observatoire des politiques relatives à l'IA ont été installés en 2018, conduisant à l'établissement de principes directeurs sur l'IA (2019), d'un cadre de classification des systèmes d'IA (2022) et d'une plateforme de suivi des incidents liés à l'IA (2023).

Les États ont aussi porté des projets communs. En 2020, le partenariat mondial pour l'intelligence artificielle (PMIA) a été lancé sous l'impulsion de la France et du Canada, afin de développer une expertise scientifique robuste sur l'IA et de formuler des recommandations concrètes pour le développement de systèmes d'IA responsables et respectueux des droits humains. Le PMIA rassemble une solide communauté de parties prenantes, avec 29 États membres de tous les continents et de tous les niveaux de développement, ainsi que des experts de l'industrie, du monde universitaire, de la société civile. Le PMIA a par exemple travaillé sur l'intégration des systèmes d'IA dans les organisations et ses effets sur l'emploi, et sur le recours à l'IA par les petites et moyennes entreprises.

Faut-il aller plus loin que ces initiatives ? Oui car elles ne constituent pas une véritable gouvernance internationale de l'IA. D'une part, les travaux menés jusqu'à présent relèvent de l'analyse, de la déclaration ou de la recommandation et n'ont pas de portée contraignante. D'autre part, les initiatives sont dispersées et aucune d'entre elles ne dispose isolément ni des moyens ni de la légitimité suffisante pour assurer une gouvernance efficace de l'IA au niveau international.

Notre Commission considère qu'une gouvernance mondiale doit aujourd'hui émerger de ce concert d'initiatives et recommande la fondation, lors du sommet sur l'IA accueilli par la France, d'une Organisation mondiale de l'IA (*World AI Organization*). Intéressons-nous donc à quatre questions clés liées à cette organisation : sa composition, ses missions, son fonctionnement et sa temporalité.

Recommandation n° 22

Structurer une initiative diplomatique cohérente et concrète visant la fondation d'une gouvernance mondiale de l'IA.

Recommandation n° 23

Structurer dès maintenant un écosystème national ouvert de gouvernance de l'IA.

Une gouvernance, composée de qui ? La participation des États est bien sûr indispensable, car ils ont la capacité de conclure des traités et de les rendre applicables sur leur territoire. Cependant, les progrès de la recherche en IA et la diffusion mondiale de la technologie sont très fortement le fait de très grandes entreprises privées à dimension mondiale. Qu'on le regrette ou non, la légitimité de la dimension interétatique ne suffit donc pas, à elle seule, à concevoir une gouvernance de l'IA à l'échelle internationale qui soit pleinement effective. Par ailleurs, à l'ère numérique, la société civile est incontournable. Très largement utilisatrice, elle est aussi contributrice au développement et à la diffusion de l'IA, avec un ensemble de principes et de référentiels culturels qui fondent par exemple l'*open source*.

Nous écartons ainsi plusieurs schémas existants de gouvernance. Le modèle de gouvernance fondé exclusivement sur des relations interétatiques, tel que celui hérité des organisations internationales du XXe siècle, serait incomplet. De même, une gouvernance de l'IA constituée autour des seules composantes de la société civile internationale, même largement entendue (associations, organisations non gouvernementales, entreprises, structures de recherche, etc.) serait en grande partie dépourvue d'effectivité. En particulier, le poids de quelques grandes entreprises de l'IA ne permet pas de s'inspirer du modèle du secteur aérien qui verrait se confronter seuls face à face, d'un côté, une organisation dotée de compétences de normalisation réunissant ces seuls acteurs privés commerciaux — l'association internationale de l'aéronautique commerciale (Iata) — et, de l'autre côté, une organisation interétatique, sous l'égide de l'ONU, dotée d'une compétence de réglementation et de régulation — l'organisation de l'aviation civile internationale (OACI).

La gouvernance de l'IA doit donc être d'un genre nouveau. Pour assurer pleinement la légitimité et l'effectivité de son action, la gouvernance devrait être constituée, à parts égales, d'une part,

de représentants des États et des organisations interétatiques et, d'autre part, de représentants individuels répartis en quatre collèges correspondant aux ensembles d'acteurs clés de l'IA (recherche, structures privées d'intérêt général, entreprises, territoires).

Une organisation, pour quelles missions ? Le traité constitutif de l'Organisation mondiale de l'IA lui confierait un ensemble de compétences. Celles-ci seraient de trois ordres. Premièrement, l'organisation serait chargée d'établir des normes contraignantes sur les systèmes d'IA, notamment en matière de standardisation des processus d'audit de l'IA. Deuxièmement, à la manière du GIEC, elle serait chargée de faire un état des connaissances sur l'évolution des systèmes d'IA et de ses impacts. Troisièmement, elle déciderait des orientations stratégiques pour des projets d'intérêt général mondial, par exemple en matière d'outils d'IA au profit de la transition environnementale. Le financement des projets en tant que tel serait confié à un fonds international spécifique (voir ci-dessous).

Une organisation, fonctionnant comment ? L'assemblée générale de l'Organisation rassemblerait l'ensemble des représentants des États et des représentants individuels. Un organe exécutif collégial restreint (au maximum 10 personnes) serait chargé d'assurer, non seulement le fonctionnement et la gestion, mais aussi la conduite des projets de l'organisation.

Le financement de cette organisation pourrait être assuré de manière pérenne à parts égales par les États et l'ensemble des composantes non étatiques en étant suffisant pour garantir une participation effective des acteurs non étatiques et pour assurer une expertise forte et reconnue.

Afin de garantir la confiance de l'ensemble des acteurs, les missions de standardisation et de normalisation, notamment en matière d'audit, pourraient être confiées à une instance spécialisée, dotée d'une grande autonomie et de garanties fortes d'impartialité. La gouvernance de cette structure spécialisée pourrait être composée de personnes de confiance élues ou désignées par les différentes composantes et collèges de l'organisation principale, à la manière d'un « *trust* » anglo-saxon, et des professionnels techniques nationaux ou régionaux. Au-delà de son expertise propre, cette instance pourrait s'appuyer notamment sur des représentants des professions d'experts techniques nationales ou régionales (CEN-CENELEC, etc.).

Une organisation, à quel moment ? Le sommet sur l'IA qu'accueillera la France fin 2024 ou début 2025 constitue l'occasion de fonder cette gouvernance mondiale de l'IA. Dans cette perspective, il apparaît nécessaire de construire dès à présent ce projet, avec les partenaires diplomatiques de la France ainsi que les acteurs non étatiques appelés à participer à cette future organisation. En amont et en aval du sommet, un ensemble d'actions pourraient être organisées afin de mobiliser la société civile en France et à l'étranger.

Le sens d'une candidature française pour accueillir l'Organisation mondiale de l'IA

Sans être en tête de la compétition internationale dans le domaine de l'IA, la France contribue au développement de la technologie et de ses modèles d'affaires. Lieu de formation de grands chercheurs contemporains en IA, la France accentue l'effort de recherche et de formation depuis le rapport de 2018 de C. Villani (lauréat de la médaille Fields). Dans le même temps, son tissu d'entreprises de l'IA prend de l'ampleur, avec par exemple deux entreprises figurant au classement des modèles de fondation les plus précis (sur douze dans le monde[139]).

Par ailleurs, la France s'est engagée ces dernières années à poser les premières briques d'une gouvernance mondiale. Co-fondateur du PMIA avec le Canada, elle accueille aussi à Paris le siège de l'OCDE, qui mène d'importants travaux dans le domaine de l'IA.

En outre, au sein de l'Union européenne, la France a exprimé à l'occasion de la négociation du règlement européen sur l'IA une position visant à concilier les deux impératifs de protection et d'innovation. Lors du sommet de Londres de novembre 2023, qui faisait prévaloir des discours fondés sur les risques existentiels, la France a également promu une vision reconnaissant les risques de l'IA, mais aussi soulignant les bénéfices de cette technologie. Cette ligne pourrait contribuer à construire un espace de convergence internationale.

Enfin, dans le contexte actuel de compétition économique accrue entre les États-Unis et la Chine, la France pourrait apparaître comme un point d'équilibre, favorisant l'émergence rapide d'une gouvernance incluant des éléments de régulation, de normalisation et d'audit.

Aux côtés de cette Organisation mondiale de l'IA, la France pourrait promouvoir la création d'un **Fonds international pour l'IA d'intérêt public** (*International Fund for Public Interest AI – IFPAI*). Ce fonds contribuerait à faire émerger un ensemble de projets bénéfiques pour l'humanité : des services d'IA libres et gratuits (*open source*), des projets de recherche indépendante, des innovations (dans l'environnement, la science, la santé…). Le budget du fonds pourrait être de l'ordre de 500 M€ par an.

Nous suggérons un Fonds international, distinct l'Organisation mondiale présentée précédemment pour quatre raisons principales. D'abord, il apparaît nécessaire de séparer — afin de prévenir d'éventuels conflits d'intérêts — la mission d'établir les normes et les standards, qui emporte des enjeux politiques et relèverait donc de la compétence de l'Organisation mondiale, et la mission de financer des initiatives d'intérêt général mondial (recherche indépendante, construction d'écosystèmes…), qui relèverait de la compétence du Fonds international. Ensuite, un Fonds indépendant permet de réunir l'expertise en matière d'évaluation financière, d'accompagnement et de suivi des projets innovants. Par ailleurs, le Fonds s'appuiera sur les contributions des donateurs, tandis que l'Organisation mondiale s'appuiera sur les contributions statutaires des États membres et de tiers. Enfin, la constitution d'un Fonds permet d'accentuer l'engagement des acteurs non étatiques, avec un conseil d'administration multipartite.

139. Centre de recherche sur les modèles de fondation de Stanford (janvier 2024).

Au-delà de la cristallisation de la gouvernance mondiale, le sommet sur l'IA qu'accueillera la France pourrait être l'occasion de faire progresser la coopération internationale dans quatre directions.

Mettre l'IA au service du bien commun. La France pourrait proposer la création d'un mécanisme d'accès à la puissance de calcul pour les pays en voie de développement, par exemple dans le cadre d'un « **1 % IA** » : tous les acteurs investissant significativement dans la puissance de calcul pourraient s'engager à allouer 1 % à ces pays.

Favoriser le dynamisme de l'innovation et l'ouverture de l'écosystème. Face au risque de concentration de l'écosystème de l'IA, la France pourrait promouvoir le financement international de communs de l'IA, notamment en matière de données. De bonnes pratiques en matière de politique de la concurrence (voir *2.3.3. Éviter les positions concurrentielles dominantes*) pourraient également être identifiées et mises en œuvre à l'échelle mondiale.

Préparer l'avenir du travail avec l'IA. Dans le prolongement des travaux de l'Organisation internationale du travail et de notre Commission, la France pourrait porter un projet international d'évaluation de l'impact de l'IA sur l'emploi et le travail, afin de faciliter le dialogue social et l'orientation de la technologie vers la qualité de l'emploi. Elle pourrait également initier une réflexion sur certains métiers se trouvant aujourd'hui au cœur de la chaîne de valeur de l'IA, comme ceux de l'apprentissage par renforcement humain.

Promouvoir des systèmes d'IA sûrs et sécurisés. Le sommet pourrait aboutir au principe de convergence internationale des standards et méthodes d'évaluation des systèmes d'IA afin d'éviter le morcellement des règles de sécurité et de sûreté dans le monde. Ces règles sont destinées à protéger les utilisateurs, mais aussi à créer un cadre de confiance favorable au développement des modèles d'affaires de l'IA.

2.3.2 DISPOSER EN FRANCE D'UNE CAPACITÉ D'ÉVALUATION DES SYSTÈMES D'IA

Les performances des systèmes d'IA nous étonnent souvent mais leurs limites sont presque aussi évidentes que leurs qualités : quelques échanges avec un robot conversationnel piloté par une IA générative suffisent pour comprendre qu'un système d'IA peut répondre de façon incomplète et trompeuse. Les spécialistes emploient le mot « hallucinations » (ou *confabulations*) pour désigner les résultats erronés ou les faits imaginaires énoncés par un système d'IA générative. Elles sont loin d'être le seul défaut de ces systèmes : des essais ont montré que, mal contrôlés, ils pouvaient produire des résultats discriminatoires ou sexistes, révéler des informations confidentielles présentes dans leurs données d'entraînement, créer des contenus pédopornographiques ou menacer leur utilisateur…

Pour prévenir les risques que peut occasionner l'utilisation des systèmes d'IA générative, les entreprises qui développent ces systèmes les soumettent à des tests poussés et ont recours à diverses techniques pour améliorer la qualité des résultats. Il peut s'agir, par exemple, de fournir à la machine le retour d'un humain pendant la phase d'entraînement, ou encore d'intervenir directement sur l'algorithme pour contrôler les résultats produits. Ces entreprises doivent être en mesure de vérifier et de démontrer que les précautions prises permettent réellement de réduire les risques à un niveau acceptable, mais aussi, plus simplement, que leurs systèmes

remplissent efficacement l'usage pour lequel ils sont prévus. Notons à ce propos qu'il n'est pas évident d'améliorer la qualité en amont d'un modèle général, car il est ensuite amené à être peaufiné pour des usages particuliers.

De façon symétrique, les utilisateurs des systèmes d'IA et les pouvoirs publics veulent avoir la garantie que ces systèmes présentent un niveau suffisant de fiabilité et de sécurité. Cette garantie est particulièrement importante lorsque le système est destiné à être utilisé dans un domaine sensible, comme la santé humaine, le maintien de l'ordre ou le recrutement. Les pouvoirs publics doivent, en outre, être en mesure de suivre dans la durée l'évolution des performances et des biais des systèmes d'IA afin d'anticiper de nouveaux risques. Ce besoin d'évaluation selon des standards partagés dépasse la simple performance, puisqu'il concerne aussi bien les biais sociaux (comme la discrimination) ou l'impact environnemental. Tout ceci implique de disposer de capacités d'évaluation holistique des systèmes d'IA, qui soit fiable et dont la fiabilité soit reconnue par tous.

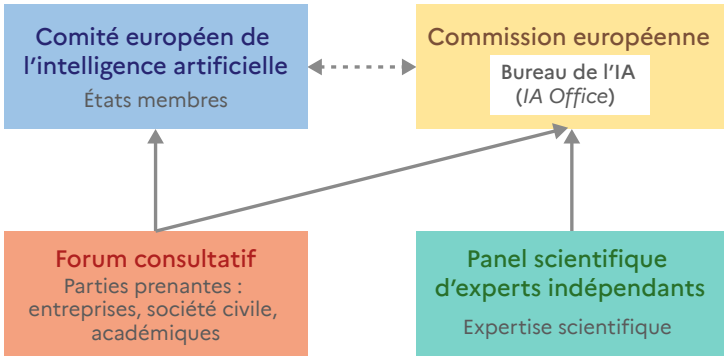
La demande d'évaluation viendra d'une part de la capacité des vendeurs de systèmes d'IA à se différencier grâce à ces évaluations et d'autre part d'obligations réglementaires. L'offre pour sa part sera facilitée par une clarification des standards d'évaluation et des normes.

Du point de vue des obligations réglementaires, un cadre pour l'évaluation des systèmes d'IA est en cours de mise en place. Le règlement européen sur l'intelligence artificielle, en cours d'adoption, impose la réalisation d'une évaluation de conformité des systèmes d'IA dits « à haut risque » préalablement à leur mise sur le marché. Cette évaluation vient s'ajouter à celles qui sont prévues par de nombreuses réglementations sectorielles déjà applicables : un appareil de radiographie intégrant une IA est avant tout un équipement médical ; il doit être évalué dans les mêmes conditions et présenter les mêmes garanties de fiabilité et de sécurité qu'un appareil de radiographie « classique ». Des initiatives de même nature sont en cours dans d'autres régions du monde.

Qui mettra en œuvre le règlement sur l'IA ?

L'objectif du règlement européen sur l'IA est de fournir un cadre favorable au développement de l'IA dans l'Union européenne, en tirant notamment avantage de la taille du marché unique, tout en apportant de solides garanties quant à la protection des droits fondamentaux. Le règlement fixe un ensemble de règles communes pour la mise sur le marché et l'utilisation des systèmes d'IA, mais différenciées selon le degré de risque associé à ces systèmes. Il prévoit une gouvernance complexe.

Au niveau européen, quatre instances sont créées. Le « bureau européen de l'IA » est la principale instance d'application du règlement. Un « panel scientifique » d'experts indépendants appuie le bureau dans l'exercice de ses missions. Un « comité européen de l'IA » rassemble les États à des fins de coordination et de partage de bonnes pratiques. Enfin, un « forum consultatif » réunit des acteurs économiques et de la recherche pour apporter une expertise technique à l'Union et aux États membres.



Au niveau national, la mise en œuvre du règlement sur l'IA s'appuie sur deux ensembles d'acteurs. D'une part, les autorités dites notifiantes, qui sont chargées de désigner (« notifier ») et de contrôler les organismes qui évaluent la conformité des systèmes d'IA mis sur le marché européen. D'autre part, les autorités de surveillance du marché, qui sont chargées de surveiller des enjeux propres à l'IA (supervision des tests par exemple) ou des enjeux sectoriels au croisement avec l'IA (marché financier, protection des données, répression des fraudes…).

Dans ce contexte, notre Commission souligne l'intérêt de désigner des autorités de surveillance du marché en privilégiant la compétence et la connaissance du marché en cause, sans chercher à désigner une autorité unique. Il sera donc nécessaire de soutenir la montée en compétence de ces organismes, notamment *via* la mise à disposition de moyens mutualisés.

S'il existe des standards, ils continuent d'évoluer dans cet écosystème encore naissant. Ces standards doivent traduire en solutions techniques les exigences imposées aux systèmes d'IA, par exemple en termes de robustesse ou de fiabilité. Il s'agit de répondre de façon concrète à des questions telles que : qu'est-ce qu'une IA fiable dans le domaine médical ? qu'est-ce qu'une IA non biaisée dans le domaine du recrutement ?

La définition rapide de standards permettant l'évaluation des systèmes d'IA est dans l'intérêt de tous, industrie comprise, car elle en a besoin tant pour remplir ses obligations réglementaires

et que pour créer la confiance et la différenciation. En l'absence de standards permettant d'évaluer effectivement les performances des systèmes d'IA, de nombreux marchés pourraient rester fermés aux solutions d'IA. La définition de ces standards n'est cependant pas uniquement un enjeu économique et juridique, c'est aussi un enjeu politique, car il s'agit de donner un contenu concret aux valeurs que doivent respecter les systèmes d'IA. Cette question ne peut pas être laissée entièrement entre les mains des entreprises.

En Europe, les normes harmonisées sur lesquelles repose l'application du règlement sur l'IA sont appelées à jouer un rôle central. La Commission européenne a donné mandat au Comité européen de normalisation appelé CEN/CENELEC pour les établir, avec un objectif ambitieux : être prêt dès l'entrée en vigueur du règlement, en 2026. Au plan mondial, de nombreuses enceintes mènent des travaux afin de définir les standards et méthodes nécessaires à l'évaluation des systèmes d'IA, qu'il s'agisse d'organismes de normalisation (l'ISO, au niveau mondial, le NIST, aux États-Unis, le BSI, au Royaume-Uni…) ou de forums « privés ».

Malgré leur caractère technique, la Commission considère que plusieurs enjeux doivent être portés à haut niveau.

Premièrement, les standards devraient être harmonisés au niveau mondial, malgré le foisonnement d'initiatives. Ce sujet devrait donc être pris en compte dans la gouvernance internationale à mettre en place (voir *2.3.1*). Au niveau européen, l'application de l'IA Act devrait se faire de façon harmonisée en Europe pour éviter toute course au moins disant entre Etats-membres.

Deuxièmement, les standards à définir devront être évolutifs : l'IA investit de nouveaux domaines, les besoins en matière d'évaluation des systèmes d'IA évolueront parallèlement. Surtout, il est peu probable qu'une réponse définitive aux questions posées par l'évaluation des systèmes d'IA puisse être trouvée du premier coup. Il s'agira plus vraisemblablement d'un processus itératif, comme on l'a vu dans le domaine de la cybersécurité. Des travaux restent d'ailleurs nécessaires sur de nombreux sujets. Ainsi, le recours à des équipes indépendantes pour mener des tests adversaires (*red teaming*) est considéré comme une méthode d'évaluation très prometteuse, déjà utilisée d'ailleurs par certaines entreprises, mais il n'y a pas encore de consensus sur ce que constitue une approche efficace et exhaustive en la matière. D'autres sujets relèvent encore de travaux de recherche, comme le développement de techniques de désapprentissage machine, permettant de faire « oublier » à un système d'IA générative une donnée ou une information ingurgitée lors de son entraînement.

Troisièmement, les normes et notamment la définition des obligations de pratiques de documentation et d'évaluation devront être définis en tenant compte de la réalité des pratiques de développement de systèmes d'IA. Les développeurs utilisent aujourd'hui de nombreux outils de documentation automatique et d'évaluation de la performance et des biais de leurs systèmes. Les grilles de mesures doivent être définies et harmonisées, mais elles gagneraient à s'inscrire facilement dans ces outils. Pour le dire clairement, la conformité à l'IA Act peut devenir un poids en forçant à une documentation manuelle très lourde pour de petites entreprises, ou être un exemple de régulation technologique.

Pour assurer une mise en œuvre au plus proche des usages, évolutives et technologique, la Commission propose d'anticiper la mise en œuvre de l'IA Act en France, et notamment de désigner des autorités de surveillance au plus proches des marchés, soutenues par des capacités mutualisées d'évaluation renforcées. Concentrer l'ensemble de la mise en œuvre du règlement IA en France dans une seule autorité (comme un « *AI Safety Office* » français) réduirait la résilience de l'écosystème français et limiterait la montée en compétence des autorités sectorielles. Si l'IA

est bien une technologie à usage générale, toutes les autorités sectorielles doivent s'en préoccuper.

Du point de vue de l'offre, il est nécessaire de faire émerger une capacité d'audit des systèmes d'IA qui soit indépendante des fournisseurs. Pour cela, à l'appui du règlement européen sur l'IA, il convient d'accompagner la montée en compétence des organismes intervenant dans l'évaluation de conformité des systèmes d'IA (appelés « organismes notifiés ») dont le règlement prévoit la désignation. Il semble également opportun d'encourager la certification des systèmes d'IA, qui peut constituer un avantage concurrentiel pour les entreprises qui y ont recours et pour laquelle un certain nombre de programmes sont d'ores et déjà disponibles[140].

Enfin, la capacité d'évaluation et d'audit devra être complétée par le développement de solutions de sécurisation des systèmes d'IA. Il s'agit par exemple de permettre de détecter des cyberattaques (des modèles d'IA ou avec des modèles d'IA) ou de mener des actions de protection ou de remédiation.

Recommandation n° 24

Doter la France et l'Europe d'un écosystème d'évaluation public et privé des systèmes d'IA au plus proche des usages et des derniers développements technologiques.

2.3.3 ÉVITER LES POSITIONS CONCURRENTIELLES DOMINANTES

Le début des années 2000 a marqué l'émergence de géants du numérique, en particulier Alphabet, Apple, Meta, Amazon et Microsoft, qui dominent aujourd'hui la majorité des segments de la chaîne de valeur des technologies numériques. Le manque de réactivité de la politique de la concurrence, aux États-Unis comme en Europe, a permis cet état de fait. Or, le développement de l'IA confirme et renforce le risque de concentration de la chaîne de valeur du numérique.

Limiter les positions concurrentielles dominantes favorise la croissance et une juste répartition des gains économiques. Il est donc crucial d'étudier, mais plus encore d'anticiper l'évolution concurrentielle de la chaîne de valeur de l'IA pour permettre le bon exercice de la concurrence. Les institutions européennes devront donc faire des enjeux concurrentiels liés à l'IA une des priorités de leurs actions, afin que les marchés puissent rester contestables[141], c'est-à-dire que les barrières à l'entrée de nouvelles entreprises et à la sortie d'entreprises existantes soient limitées. Les enjeux concurrentiels apparaissent à différents échelons de la chaîne de valeur IA.

En amont, le poids largement prépondérant d'un seul acteur sur le marché de la conception des processeurs graphiques (*GPU*), qui sont une brique essentielle pour l'entraînement et l'optimisation des grands modèles de langue, génère des risques de comportement anticoncurrentiel. Par ailleurs, comme l'indique l'autorité de la concurrence dans son avis de

140. Afnor Certification et le LNE, notamment, ont mis en place des certifications portant sur l'IA.
141. Un marché est contestable s'il est possible pour de nouveaux producteurs d'entrer et pour des producteurs actuels de sortir dans des conditions supportables, i.e. si les barrières à l'entrée et à la sortie sont acceptables.

juin 2023[142], le secteur du *cloud*, également en amont de la chaîne de valeur de l'IA, est dominé par trois grands acteurs, les « *hyperscalers* » (Amazon Web Services, Google Cloud Platform et Microsoft Azure). En 2021, ces trois acteurs représentaient 80 % de la croissance des dépenses en infrastructures et applications de services cloud public en France. Ces entreprises, aussi bien en raison de leur puissance financière que de leur écosystème de services numériques très riche, sont en mesure d'entraver la bonne concurrence sur la chaîne de valeur.

En aval, la vente liée (*« bundling »* en anglais) est une perspective préoccupante. Elle consiste pour des acteurs présents sur un segment de la chaîne de valeur, à proposer une offre commerciale composée, en plus du produit principal, de produits (logiciels par exemple) supplémentaires sur d'autres segments de la chaîne de valeur. Cet enjeu n'est pas propre à la chaîne de valeur de l'IA, puisqu'il existe depuis le début de la révolution numérique (on peut penser à la commercialisation de la suite *Microsoft Office* par exemple). Si cette pratique commerciale n'est pas anticoncurrentielle par nature, elle peut induire des comportements répréhensibles, et il est important de les identifier.

Au cœur de la chaîne de valeur, la vigilance est de mise en matière de concentration. En effet, les entreprises américaines les plus prometteuses sur le développement des plus grands modèles, dits de fondation, sont déjà liées à des acteurs historiques de la révolution numérique (bien que selon des manières différentes) : OpenAI à Microsoft et Anthropic à Amazon et Google. Même si ces entreprises ne sont pas formellement contrôlées par ces acteurs historiques, les financements des géants du numérique pourraient constituer une forme de contrôle et soulèverait donc des enjeux concurrentiels.

Les pouvoirs publics ne sont pas démunis face à ces enjeux. L'entrée en vigueur du règlement sur les marchés numériques (*digital markets act* - DMA) au premier semestre 2023 a pour objectif de permettre aux marchés dans le secteur numérique de rester contestables et équitables. Plus précisément, cette réglementation vise à empêcher qu'une grande plateforme jouissant d'une position de « contrôleur d'accès » (*gatekeeper*) vis-à-vis d'un grand nombre d'utilisateurs n'abuse pas de cette position en empêchant des entrants potentiels d'accéder à ces utilisateurs, et donc à ce marché.

Ainsi, le DMA définit un cadre clair, qui permettra de répondre à la majorité des enjeux concurrentiels du secteur numérique « historique ». Il devrait permettre de s'attaquer aux comportements anticoncurrentiels en aval de la chaîne de valeur : on peut penser à celui de la vente liée, ou à la réutilisation de données personnelles collectées grâce à l'utilisation d'un premier produit (par exemple la collecte de données personnelles via un système de messagerie pour faire de la publicité ciblée sur un réseau social appartenant à la même plateforme). Le DMA devrait également permettre de s'intéresser aux problèmes de concentration sur certains marchés numériques, situation dans laquelle un faible nombre d'acteurs voire un seul acteur domine le marché, grâce à l'obligation faite aux « contrôleurs d'accès » de déclarer l'ensemble de leurs acquisitions, y compris celles de petites entreprises dont l'acquisition passait jusqu'ici sous les radars de la Commission européenne.

Toutefois, le DMA répond à ces enjeux au travers d'un prisme bien défini, celui des plateformes. Or, la chaîne de valeur de l'IA, encore émergente et sujette à de nombreuses évolutions, ne convergera pas nécessairement vers un mode de fonctionnement exclusif de type « plateforme en ligne ». C'est pourquoi le DMA tel qu'il existe actuellement pourrait ne pas être opérant pour traiter l'ensemble des enjeux concurrentiels de la chaîne de valeur de l'IA, mais plutôt certains

142. Avis 23-A-08 du 29 juin 2023 portant sur le fonctionnement concurrentiel de l'informatique en nuage ("cloud").

points (référencement sur les moteurs de recherche, portabilité des données, interopérabilité des agents personnels par exemple).

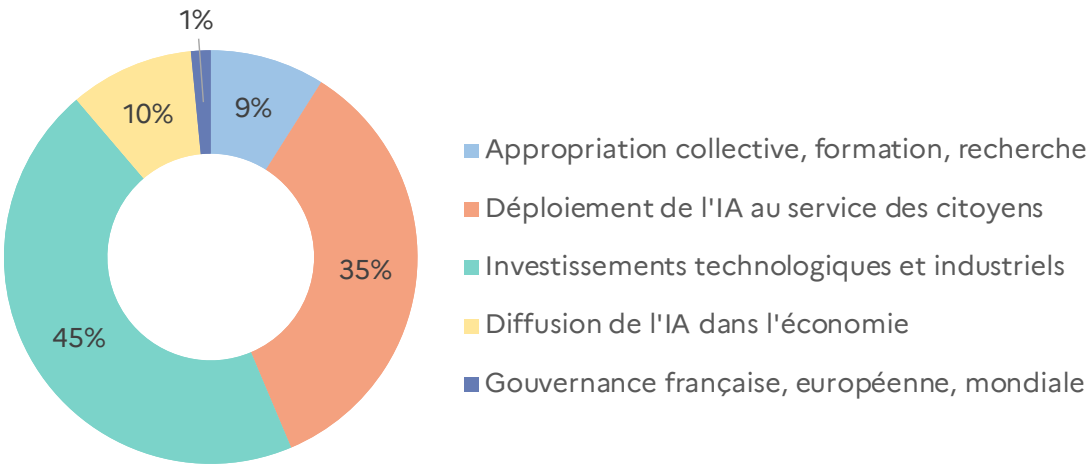
Ainsi, le DMA est en mesure d'apporter des solutions aux enjeux concurrentiels présents sur la chaîne de valeur de l'IA tant que ces enjeux sont reliés à une problématique de plateforme. Il ne sera probablement pas suffisant pour couvrir tous les enjeux concurrentiels. Ce règlement mériterait donc d'être complété pour tenir compte des spécificités de l'IA. Enfin, à moyen terme, il convient d'envisager un changement de doctrine de la politique de concurrence, en passant d'un système *statique* (quelles parts de marché détient *aujourd'hui* cette entreprise ?) à une vision *dynamique* (quelles parts de marché pourraient *demain* détenir cette entreprise et quelles entreprises pourraient *demain* entrer sur ce marché ?), permettant d'anticiper les concentrations plutôt que d'attendre de pouvoir les constater.

Recommandation n° 25

Assurer un suivi de l'évolution des concentrations de marché et mettre en place, rapidement, la réglementation nécessaire pour éviter les abus de position dominante.

NOS RECOMMANDATIONS

Le plan proposé par notre Commission inclut 25 recommandations, qui représentent un engagement annuel d'environ 5 Md€ au cours des cinq prochaines années. Les moyens peuvent être répartis selon cinq grandes catégories : l'appropriation collective, la formation et la recherche ; le déploiement de l'IA au service des citoyens ; les investissements technologiques et industriels ; la diffusion de l'IA dans l'économie ; la gouvernance française, européenne et mondiale.



	Recommandations	Impact financier en 5 ans (M€)
1	Créer les conditions d'une appropriation collective de l'IA et de ses enjeux afin de définir collectivement les conditions dans lesquelles elle s'insère dans notre société et nos vies quotidiennes. *Pilotes : ministère de la cohésion des territoires*	10
2	Investir dans l'observation, les études et la recherche sur les impacts des systèmes d'IA sur la quantité et la qualité de l'emploi. *Pilotes : ministère du travail et ministère de la fonction publique*	5

	Recommandations	Impact financier en 5 ans (M€)
3	Faire du dialogue social et professionnel un outil de co-construction des usages et de régulation des risques des systèmes d'IA. *Pilotes : ministère du travail*	0
4	Porter une stratégie de soutien à l'écosystème d'IA ouverte au niveau international en soutenant l'utilisation et le développement de systèmes d'IA ouverts et les capacités d'inspection et d'évaluation par des tiers. *Pilotes : ministère de l'enseignement supérieur et de la recherche ; ministère de l'économie*	0
5	Faire de la France un pionnier de l'IA pour la planète en renforçant la transparence environnementale, la recherche dans des modèles à faible impact, et l'utilisation de l'IA au service des transitions énergétique et environnementales. *Pilotes : ministère de l'Europe ; ministère de la recherche*	100
6	Généraliser le déploiement de l'IA dans toutes les formations d'enseignement supérieur et acculturer les élèves dans l'enseignement secondaire pour rendre accessibles et attractives les formations spécialisées. *Pilotes : ministère de l'enseignement supérieur et de la recherche, ministère de l'éducation nationale*	1 200
7	Investir dans la formation professionnelle continue des travailleurs et dans les dispositifs de formation autour de l'IA. *Pilotes : ministère du travail et ministère de la fonction publique*	200
8	Former les professions créatives à l'IA, dès les premières années de l'enseignement supérieur et en continu. *Pilotes : ministère de l'enseignement supérieur et de la recherche ; ministère de la culture*	20
9	Renforcer la capacité technique et l'infrastructure du numérique public afin de définir et de passer à l'échelle une réelle transformation des services publics grâce au numérique et à l'IA, pour les agents et au service des usagers. *Pilote : ministère de la transformation et de la fonction publique*	5 500
10	Faciliter la circulation des données et le partage de pratiques pour tirer les bénéfices de l'IA dans les soins, améliorer l'offre et le quotidien des soignants. *Pilote : ministère de la santé*	3 000
11	Encourager l'utilisation individuelle, l'expérimentation à grande échelle et l'évaluation des outils d'IA pour renforcer le service public de l'éducation et améliorer le quotidien des équipes pédagogiques. *Pilote : ministère de l'éducation*	1 000

	Recommandations	Impact financier en 5 ans (M€)
12	Investir massivement dans les entreprises du numérique et la transformation des entreprises pour soutenir l'écosystème français de l'IA et en faire l'un des premiers mondiaux. *Pilotes : services du Premier ministre ; ministère de l'économie et des finances*	3 600
13	Accélérer l'émergence d'une filière européenne de composants semi-conducteurs adaptés aux systèmes d'IA. *Pilotes : services du Premier ministre ; ministère de l'économie ; ministère de l'enseignement supérieur et de la recherche*	7 700
14	Faire de la France et de l'Europe un pôle majeur de la puissance de calcul installée. *Pilotes : services du Premier ministre ; ministère de l'économie et des finances*	1 000
15	Transformer notre approche de la donnée personnelle pour mieux innover. *Pilote : services du Premier ministre ; ministère de la justice*	16
16	Mettre en place une infrastructure technique favorisant la mise en relation entre les développeurs d'IA et les détenteurs de données culturelles patrimoniales. *Pilote : ministère de la culture et ses opérateurs*	35
17	Mettre en œuvre et évaluer les obligations de transparence prévues par le règlement européen sur l'IA en encourageant le développement de standards et d'une infrastructure adaptée. *Pilote : ministère de la culture*	0
18	Attirer et retenir des talents de stature internationale avec des compétences scientifiques ou entrepreneuriales et managériales dans le domaine de l'IA. *Pilote : ministère de l'économie ; ministère des affaires étrangères*	10
19	Assumer le principe d'une « Exception IA » sous la forme d'une expérimentation dans la recherche publique pour en renforcer l'attractivité. *Pilote : ministère de l'enseignement supérieur et de la recherche*	1 025
20	Inciter, faciliter et amplifier le recours aux outils d'IA dans l'économie française en favorisant l'usage de solutions européennes. *Pilote : services du Premier ministre ; ministère de l'économie*	2 600
21	Faciliter l'appropriation et l'accélération des usages de l'IA dans la culture et les médias pour limiter la polarisation entre grands groupes et petits acteurs et lutter contre la désinformation. *Pilote : ministère de la culture*	60
22	Structurer une initiative diplomatique cohérente et concrète visant la fondation d'une gouvernance mondiale de l'IA. *Pilote : ministère des affaires étrangères ; ministère de la culture ; ministère de l'économie*	300

	Recommandations	Impact financier en 5 ans (M€)
23	Structurer dès maintenant un puissant écosystème national de gouvernance de l'IA. *Pilotes : services du Premier ministre, ministère de l'économie, ministère de la recherche*	5
24	Doter la France et l'Europe d'un écosystème d'évaluation public et privé des systèmes d'IA au plus proche des usages et des derniers développements technologiques. *Pilotes : ministère de l'économie et des finances*	15
25	Anticiper les concentrations de marché sur l'ensemble de la chaîne de valeur de l'intelligence artificielle. *Pilote : ministère de l'économie*	0
	Ensemble du plan	27 Md€

LETTRE DE MISSION

Le Premier Ministre

Paris, le 28 FEV. 2024

Madame, Monsieur,

L'arrivée progressive de l'intelligence artificielle (IA) au cœur de notre quotidien révèle chaque jour un peu plus son potentiel et soulève de très nombreuses questions, notamment dans les domaines de l'éthique, de l'économie, de la productivité, du travail, de l'organisation des entreprises ou encore de la souverainetés industrielle et numérique des États. Cette innovation constitue un point de rupture dans nos sociétés modernes, venant bouleverser nos modes de pensée, nos modes de production, nos modes de consommation, en bref; nos modes de vie.

Les bouleversements technologiques et sociaux qui en découlent prouvent que la bataille pour la souveraineté passe par la maîtrise de ces technologies. Dès 2018, la première phase de la stratégie nationale pour l'IA, lancée par le Président de la République, a permis de développer une filière de rang mondial. Avec le lancement du plan France 2030 en octobre 2021, notre pays a rehaussé son ambition en la matière.

Aujourd'hui, face à ces nouveaux défis, nous devons aller plus loin, notamment en renforçant nos formations pour développer davantage de talents en France, en nous assurant que notre tissu économique ait les moyens de tirer le maximum de cette technologie, en investissant pour favoriser l'innovation française sur la scène internationale et en définissant une régulation adaptée des différents secteurs pour protéger des dérives.

C'est la raison pour laquelle ma prédécesseure, Mme Elisabeth Borne, a souhaité installer la Commission de l'intelligence artificielle, le 19 septembre 2023 à Matignon.

Sous votre présidence, cette Commission réunit des acteurs de différents secteurs (culturel, économique, technologique, de recherche), pour contribuer à éclairer les décisions du Gouvernement et faire de la France un pays à la pointe de la révolution de l'intelligence artificielle. Elle doit présenter au Gouvernement des propositions concrètes permettant d'adapter notre stratégie nationale.

Vous animez cette Commission composée de treize experts de l'intelligence artificielle : Gilles Babinet, Joëlle Barral, Alexandra Bensamoun, Nozha Boujemaa, Bernard Charlès, Luc Julia, Yann Le Cun, Arthur Mensch, Cédric O, Isabel Ryl, Franca Salis-Madinier, Martin Tisné, Gaël Varoquaux.

Monsieur Philippe AGHION
Professeur au Collège de France
Madame Anne BOUVEROT
Présidente du conseil d'administration de l'École normale supérieure

Chacun des experts a été désigné *intuitu personae* par le Gouvernement et ne peut donc en aucun cas se faire représenter. Les présidents, les membres et les experts de la Commission sont par ailleurs tenus à la confidentialité des débats et travaux auxquels ils participent ou assistent. Vous veillez à ce que cette obligation de confidentialité et de discrétion soit respectée.

Vous êtes responsables de l'avancée des travaux, dont je souhaite qu'ils s'articulent autour de cinq thématiques clés : impacts économiques, souveraineté industrielle et numérique, éthique et impacts sociétaux, enjeux culturels et service public. Pour chacun de ces thèmes, vous veillez à la diversité des personnes auditionnées qui vous permettent d'étayer les constats et les propositions établis par la Commission.

Pour mener ces travaux thématiques, la Commission de de l'intelligence artificielle bénéficie de l'appui de treize rapporteurs issus de l'administration : Marc Auberger, inspecteur général des finances, Simon Bunel, administrateur de l'Insee, Philippe Chantepie, inspecteur général des affaires culturelles, Eloy Dorado, administrateur général, Emilie-Pauline Gallié, inspectrice générale de l'éducation, du sport et de la recherche, Paul Jolie, ingénieur général des mines, Arnaud Mazier, ingénieur général des mines, Vincent Montreuil, inspecteur général de l'éducation, du sport et de la recherche, Timothée Paris, maître des requêtes au Conseil d'État, Christophe Ravier, ingénieur général des mines, Erwan Paitel, inspecteur général de l'éducation, du sport et de la recherche, Ulrich Tan, ingénieur des mines, et Louis--Charles Viossat, inspecteur général des affaires sociales.

Afin d'assurer la coordination d'ensemble des travaux de la Commission, vous êtes également accompagnés de deux rapporteurs généraux mis à votre disposition, Cyprien Canivenc, conseiller référendaire à la Cour des comptes, et Arno Amabile, ingénieur des mines. Vous pouvez également vous appuyer sur l'expertise de Guillaume Avrin, coordonnateur national pour l'intelligence artificielle.

Je souhaite que vous me rendiez régulièrement compte de l'avancée de vos travaux, jusqu'à la remise de vos conclusions définitives début mars 2024. Les propositions devront être opérationnelles, réalistes et ambitieuses soutenues par une vision long terme, globale et objectivée.

Je vous remercie et vous prie de croire, Madame, Monsieur, à l'assurance de notre considération distinguée.

Gabriel ATTAL

Maquette : DSAF / DPSG/BAEC / Section Édition / EP.
Illustration : Céline JOUDET - Mars 2024.