

Où va l'IA ? Quelles innovations pour quels usages ?



Poursuivant la révolution de l'intelligence artificielle générative amenée par les grands modèles de langage (LLM), les technologies d'IA se développent à un rythme soutenu. Les LLM ont ainsi vu leurs performances démultipliées depuis 2024. L'utilisation de modèles de raisonnement reposant sur les « chaînes de pensée » leur offre des capacités logiques et leur permet d'auto-corriger leurs erreurs. Au-delà des LLM, des innovations

comme les « modèles de monde » visent à s'affranchir des limites actuelles des algorithmes d'IA.

L'Office a organisé une audition publique pour clarifier les enjeux et les perspectives de ce bouillonnement scientifique et comprendre comment les innovations qu'il porte se diffusent dans la société. En conclusion de cette audition ⁽¹⁾, l'Office formule sept recommandations qui visent à encourager le développement d'une IA souveraine, frugale et innovante tant dans ses fondements techniques que dans les cas d'usage qu'elle ouvre à ses utilisateurs.

Alexandre Sabatou, député

Corinne Narassiguin, sénatrice

(1) Rapport Assemblée nationale n° 2744 (17^e législature) – Sénat n° 591 (2025-2026)

Les fondements technologiques de l'intelligence artificielle : questions actuelles

L'intelligence artificielle occupe une place à part au sein des technologies de pointe. Jusqu'à récemment, son étude se limitait à la communauté des chercheurs en informatique ou, dans un autre registre, à la science-fiction, très prolifique sur le sujet. Depuis 2017, la révolution technologique qu'a été l'architecture *Transformer* et son application aux LLM a fait sortir l'IA de ce cadre restreint en ouvrant son utilisation au grand public sous la forme d'agents conversationnels ou *chatbots*. L'Office s'est penché sur cette innovation et ses conséquences dans un rapport publié en 2024¹. Dans le domaine de l'IA, le délai entre la mise au point d'une innovation et sa généralisation est remarquablement court. L'IA est ainsi un rare exemple de diffusion quasiment en temps réel d'une technologie en train de s'inventer.

Depuis 2024, les performances des outils fondés sur l'intelligence artificielle ont augmenté exponentiel-

lement, comme le montrent par exemple les résultats des benchmarks ARC-AGI². Ces progrès reposent à la fois sur l'accroissement de la quantité de données utilisées pour entraîner les modèles et sur la mise au point de méthodes nouvelles d'entraînement et de fonctionnement.

Les grandes tendances de l'IA

Ces dernières années, deux innovations majeures ont permis un saut qualitatif des grands modèles de langage : les modèles de raisonnement et l'IA agentique.

Les modèles de raisonnement ont fait leur apparition en 2024, d'abord avec le modèle o1 créé par OpenAI, puis avec Gemini créé par Google. Ils utilisent les « chaînes de pensée », ou *Chains of Thought*, une manière pour les LLM de se parler à eux-mêmes pour déployer une plus longue réponse cachée à l'utilisateur. Le terme de « raisonnement » est trompeur, car il ne décrit pas un raisonnement mathématique utilisant des règles logiques. Il correspond plutôt à une suite parfois très longue de réflexions écrites par un LLM, étapes intermédiaires vers l'obtention de la réponse à la demande formulée par l'utilisateur. Le processus suivi par

¹ Alexandre Sabatou, Patrick Chaize, Corinne Narassiguin. *ChatGPT, et après ? Bilan et perspectives de l'intelligence artificielle*. Assemblée Nationale n° 642 (29 novembre 2024) – Sénat n° 170 (28 novembre 2024)

² Les benchmarks ARC-AGI sont des tests de performance qui visent à mesurer la capacité des IA à résoudre des problèmes inédits sans entraînement préalable. Ils permettent de comparer des modèles entre eux et de suivre leur évolution. La spécificité de ces tests et de tenter d'évaluer l'adaptabilité face à la nouveauté, ce qui peut être considéré comme une forme d'intelligence.

les modèles de raisonnement pour rédiger leurs « pensées latentes » améliore significativement leur capacité à se corriger et à donner une réponse valide. Les tests d'intelligence ARC-AGI ont permis de mettre en évidence l'augmentation des performances suite à la mise en œuvre des chaînes de pensée. Les modèles antérieurs avaient de grandes difficultés à résoudre les problèmes inédits proposés par ce benchmark ; à partir de l'arrivée de o1 puis de son successeur o3, les scores sur ARC-AGI se sont améliorés de plusieurs ordres de grandeur.

L'IA agentique est une IA capable de répondre à un problème posé en planifiant des stratégies qu'elle applique en interagissant avec des agents, qui produisent sous ses ordres les éléments qu'elle ne peut réaliser elle-même. L'IA devient ainsi semi-autonome ; elle est en mesure de décomposer un objectif complexe en tâches plus simples et de trouver les bons outils pour les résoudre. Si cette technologie n'est pas fondamentalement liée aux LLM, elle s'appuie très largement sur eux en pratique.

Avec l'amélioration des capacités des LLM et leur diffusion toujours plus large, une quantité croissante de texte généré par IA est présente sur internet. On pourrait craindre que cela crée un risque de contamination des données qui servent à entraîner les modèles. Au contraire, certains proposent de profiter de la possibilité de générer des **données synthétiques** par IA pour les présenter volontairement en entrée des LLM. Ce contrôle plus fin des données d'apprentissage procurerait une meilleure maîtrise de l'étape d'entraînement.

Au-delà des LLM, d'autres modalités d'interaction que le texte connaissent des progrès spectaculaires : c'est le cas de la génération de sons, de musiques et de vidéos. Le développement de modèles produisant ou intégrant ces autres modalités, appelés de ce fait **modèles multimodaux**, pourrait bientôt révolutionner le secteur de la robotique, dont l'avenir semble très prometteur et étroitement lié à l'IA.

Les **modèles de monde** s'inscrivent dans une optique similaire. Développés en France sous l'impulsion de Yann Le Cun, qui en est un promoteur ambitieux, ils visent à permettre aux modèles de comprendre leur environnement. Les modèles d'IA se construiraient une vision du monde à partir de laquelle ils seraient capables de saisir les événements se produisant autour d'eux. Ils disposeraient ainsi de connaissances physiques effectives. La robotique est à nouveau un domaine susceptible de bénéficier grandement de cette piste de recherche.

IA frugale vs. gros modèles : la question énergétique

Le numérique est depuis longtemps un secteur intensif en énergie et en capacités de calcul et de stockage. Des investissements importants y sont un prérequis pour la conception d'outils utilisables par le plus grand nombre. Ceci favorise la concentration des solutions techniques chez un petit nombre d'acteurs d'envergure mondiale. Les principales entreprises, toutes américaines, sont d'ailleurs regroupées sous l'acronyme

Gafam (Google, Apple, Facebook, Amazon et Microsoft), ce qui met en lumière la nature oligopolistique du domaine.

Cette caractéristique de l'économie numérique se retrouve naturellement dans le secteur spécifique de l'intelligence artificielle, en particulier pour les LLM. Les Gafam se sont d'ailleurs positionnés sur cette technologie, généralement par l'acquisition ou la participation au capital de start-up spécialisées. Microsoft détient par exemple 27 % des parts d'OpenAI, créateur de ChatGPT, qui est valorisé 500 milliards de dollars environ. Nvidia, qui fournit en processeurs graphiques les centres de données nécessaires à l'entraînement des modèles, a été la première entreprise à atteindre une valorisation de 5 000 milliards de dollars. En matière d'IA, les montants financiers mobilisés par les grands acteurs américains sont extrêmement élevés. Face à eux, peu de pays peuvent suivre la course au gigantisme ; seule la Chine se positionne comme un concurrent sérieux.

Les financements se matérialisent dans l'achat de toujours plus de capacités de calcul et de traitement de données, afin d'entraîner des modèles d'IA générative toujours plus importants. La taille d'un modèle se mesure en nombre de paramètres, ou « poids ». Ils sont caractéristiques du modèle en question et sont issus de son entraînement. Ils doivent être distingués des données utilisées en entrée, qui occupent bien plus de mémoire. Les modèles actuels comprennent souvent plusieurs centaines de milliards de paramètres, et la course à l'augmentation des performances s'appuie sur une hausse comparable de ce nombre. Dans ce paradigme, taille du modèle, investissements financiers et résultats obtenus aux évaluations peuvent sembler liés.

Ce lien pose cependant de nombreux problèmes, dont la dépendance des usagers à un tout petit nombre de fournisseurs, l'inadaptation des outils à certains cas d'usage, et l'augmentation de l'intensité énergétique de cette industrie. C'est pourquoi les experts s'exprimant dans le cadre de la première table ronde de l'audition publique ont mis en évidence l'intérêt de l'IA dite « frugale », non comme une voie par défaut, mais comme une alternative à la course aux gros modèles.

La frugalité consisterait à travailler plus finement sur la technologie afin de construire des modèles plus petits sans perte de performance, ce qui peut passer par une plus grande spécialisation du modèle, au détriment de son caractère généraliste. Il faut toutefois noter que même ces modèles relativement petits comptent des milliards de paramètres et demandent aussi de grandes quantités de données et une puissance de calcul significative ; ils restent néanmoins beaucoup moins gourmands en énergie que les « grands » modèles d'OpenAI ou de Google par exemple.

L'IA frugale offre un premier avantage, d'ordre pratique : elle ouvre le secteur aux pays n'ayant pas des moyens comparables à ceux des États-Unis. De l'avis commun des intervenants, la France est bien positionnée sur ces technologies et dispose d'un écosystème

dynamique et performant. Si cela n'évacue pas totalement la question de la souveraineté, qui s'apprécie sur toute la chaîne de valeur, du semi-conducteur jusqu'aux paramètres du modèle, en passant par l'origine des données d'entraînement, la possibilité de développer des modèles d'IA français ou européens durablement compétitifs est réelle.

Par ailleurs, certaines applications nécessitent ce genre de modèles plus légers. C'est le cas de l'IA embarquée, notamment pour la robotique, qui semble être un usage prometteur de l'IA. Un modèle destiné à fonctionner sur un smartphone ou un drone, par exemple, voit sa demande de mémoire et de processeurs contrainte par les capacités matérielles de son vecteur, ce qui justifie de sacrifier la généralité du modèle contre sa frugalité. À ce stade, un modèle plus léger reste capable d'atteindre les performances des modèles généralistes dans son domaine de spécialisation.

Il existe aussi un intérêt théorique à la frugalité : elle est un excellent stimulant pour les scientifiques et ingénieurs qui travaillent à développer ces modèles, car elle constitue un défi à surmonter. Plutôt que de s'enfermer dans une course au gigantisme et aux financements, ils doivent réfléchir à des solutions alternatives plus subtiles pour arriver aux mêmes résultats. Les contraintes peuvent être un moteur d'innovation, comme souvent en science. La vitalité de la recherche académique française et sa forte interaction avec les entreprises d'innovation doit à ce propos être saluée.

Le concept de frugalité s'inscrit aussi en réaction à la consommation énergétique galopante liée à la diffusion des usages de l'IA. L'alimentation électrique de nouveaux centres de données est devenue un élément important des politiques énergétiques et devient une préoccupation majeure vis-à-vis du respect des trajectoires climatiques. Sans pour autant parler de sobriété, puisque les modèles qui s'en revendiquent utilisent aussi des capacités de calcul significatives, la frugalité serait un moyen de s'extraire du paradigme dominant qui voudrait consacrer à l'IA des moyens infinis en faisant fi des limites planétaires.

L'IA peut aussi contribuer à la lutte contre le réchauffement climatique. Associée à la collecte de grandes quantités de données, elle pourrait optimiser la consommation ou la production d'énergie. L'IA offre sans doute de nombreuses possibilités dans ce domaine ; encore faut-il s'assurer que les gains qu'elle induit font plus que compenser son coût énergétique propre. Le travail d'estimation du bilan des outils d'IA, et plus généralement du numérique, a commencé ; il doit se poursuivre. Il est cependant extrêmement complexe dans un domaine par nature global dont les chaînes de valeur suivent de multiples ramifications dans des zones géographiques souvent éloignées. Il importe pourtant d'être en mesure de quantifier les bénéfices apportés par les outils d'IA afin de pouvoir les comparer. Comme on le fait pour d'autres domaines – les transports, la production d'énergie, etc., on doit évaluer les conséquences du numérique sur les émissions de gaz à

effet de serre et sur les trajectoires climatiques et les faire connaître aux usagers.

Il est intéressant de remarquer que si des efforts sont menés en France pour quantifier ces impacts, avec des travaux faisant notamment intervenir l'INRIA, l'Institut national de recherche pour l'agriculture, l'alimentation et l'environnement (Inrae) et l'Agence de la transition écologique (Ademe), ce sujet ne semble pas être une source de préoccupation dans tous les pays du monde. Les grandes entreprises de l'IA, en particulier aux États-Unis, semblent assez indifférentes à leur empreinte climatique et poursuivent l'inflation de leurs modèles. La frugalité est donc un argument non négligeable pour faire advenir une forme de souveraineté numérique. Pour être en mesure d'estimer correctement les conséquences environnementales d'un modèle, il faut connaître le processus de conception à un niveau de détail inaccessible à ce stade, notamment pour les modèles américains. Des IA frugales développées en France pourraient donc s'intégrer dans une optique de développement durable de deux manières : par leur moindre consommation énergétique, et par la transparence à construire sur leur impact climatique. La France et l'Europe ont un rôle à jouer pour défendre un numérique vivable et soutenable, à la condition de disposer de solutions d'IA souveraines efficaces.

IA symbolique et IA connexionniste : vers une réconciliation ?

Comme l'a montré le rapport de l'Office de 2024, deux visions divergentes de l'IA se sont historiquement opposées : l'IA symbolique, plus ancienne et reposant sur la mise en œuvre de règles logiques, et l'IA connexionniste fondée sur des méthodes statistiques. L'aspect probabiliste de l'IA connexionniste a toujours suscité une profonde méfiance, y compris chez les spécialistes. C'est pourtant ce type d'IA qui est à l'origine des LLM et plus généralement des progrès massifs de l'IA depuis le développement de l'architecture *Transformer* en 2017. On pourrait donc considérer que l'IA connexionniste a gagné la bataille de l'IA. Les deux innovations majeures de la période actuelle que sont l'IA agentique et les modèles de raisonnement concernent d'ailleurs cette technologie.

Pourtant, la recherche essaie toujours d'hybrider les approches symbolique et connexionniste, dans l'espoir de profiter du meilleur des deux mondes. Les efforts en ce sens sont déjà anciens et les échecs d'ores et déjà constatés dans la longue histoire de l'IA incitent à la prudence. Plus généralement, les prédictions sur l'avenir de l'IA doivent être considérées avec précaution au regard des incertitudes majeures qui entourent ces technologies en constante évolution. Il semble toutefois que cet axe de recherche soit prometteur ; il est en tout cas sérieusement étudié en France.

L'approche conceptuelle de l'hybridation entre logique et stochastique a de quoi intéresser les scientifiques. En effet, les informaticiens pourraient alors remettre du contrôle humain sur une intelligence artificielle dont le fonctionnement est encore très mal compris. Réussir ce défi serait un grand pas vers une IA

« de confiance », qui serait suffisamment vérifiable pour être utilisée avec le même niveau de validité que les algorithmes symboliques actuels. On pourrait par exemple imaginer un outil qui utiliserait un LLM pour écrire en langage formel un problème posé en langage naturel, puis qui se servirait de modèles physiques et de règles logiques pour y répondre.

De nombreux cas d'usage bénéficieraient de telles méthodes. Dans l'exemple déjà très répandu d'écriture de code informatique à l'aide de l'IA, des algorithmes hybrides pourraient démontrer l'absence d'erreurs dans le code produit par un LLM, ou dans le cas contraire lui demander de le corriger. Cette tâche est difficile à l'heure actuelle car le code généré par IA peut être peu compréhensible pour l'être humain.

Ces recherches n'annoncent donc pas la fin des LLM, qui sont là pour durer. Ils continuent d'ailleurs à produire des résultats surprenants, voire inquiétants. Claude Mythos, la nouvelle IA d'Anthropic, semble capable de détecter des failles dans quasiment tous les logiciels informatiques. Mais, au même titre que l'IA agentique ou les modèles de raisonnement, l'hybridation pourrait à terme être à l'origine d'une nouvelle évolution démultipliant les capacités de l'IA.

La diffusion des innovations dans les usages de l'IA

Il existe aujourd'hui un consensus assez fort parmi les décideurs politiques et les acteurs du secteur numérique quant à l'intérêt de développer des outils d'IA avancés et de soutenir la filière. Cela se traduit en France par un riche écosystème de start-up plus ou moins installées, de nombreux emplois et une communauté de spécialistes de l'IA qui s'intéresse de près au développement de nouvelles capacités. Le risque peut être alors d'aborder cette technologie par le seul prisme des experts, dont le point de vue est largement favorable à sa diffusion la plus large dans la société. D'où la multitude des appels à réfléchir aux manières de transformer la société ou l'entreprise par l'IA, comme si l'IA était une fin en soi. Faire primer la technologie sur la pertinence de son utilisation est un biais auxquelles les prouesses impressionnantes des outils d'IA peuvent facilement mener.

À l'inverse, toute discussion sur les usages de l'IA devrait d'abord se fonder sur leur utilité sociale. L'intelligence artificielle doit permettre d'« augmenter » les humains, individuellement et collectivement, cette « augmentation » ne pouvant avoir un sens uniquement matériel. Certes, chacun semble s'accorder sur l'importance de mettre l'humain au cœur de ces questions, mais ce principe est parfois perdu de vue au moment où sont prises les décisions concrètes. La notion d'intelligence artificielle centrée sur l'humain, qui a notamment suscité la création de l'*Institute for Human-Centered Artificial Intelligence* par Fei-Fei Li à Stanford, ne doit pas rester une idée théorique. L'IA a la capacité de bouleverser la vie sociale et économique, des pans

majeurs de la vie humaine. Ses usages doivent être interrogés à l'aune de leurs impacts éventuels.

La diffusion de l'IA dans les entreprises est en cours. Soutenue par les politiques publiques nationales et européennes, elle est parfois ralentie par un manque d'acculturation. Les cas d'usage sont pourtant nombreux et les innovations récentes en démultiplient les possibilités. L'adoption relativement plus rapide de ces outils par la société fait peser le risque d'une utilisation cachée de l'IA au sein des entreprises. Articulée avec l'application des réglementations et la formation initiale et continue, la production de modèles d'IA souverains sera nécessaire à une diffusion sûre de cette technologie.

La diffusion de l'IA en entreprise est en cours mais recouvre des réalités contrastées

Alors que les outils d'IA se sont généralisés auprès du grand public, la situation est plus contrastée au sein des entreprises. Le baromètre France Num réalisé en 2025 par la DGE estime que 26 % des TPE-PME utilisent l'IA. Cela correspond certes à une augmentation de 13 points par rapport à l'année précédente, mais l'usage de l'IA reste encore minoritaire. La taille des entreprises est un facteur important dans l'adoption des outils d'IA, les grandes structures étant plus avancées.

Les entreprises dont la taille leur permettait s'étaient précédemment positionnées sur le traitement des données massives et avaient à cette fin intégré des équipes de spécialistes en science des données. Des cas d'usage avancés, comme la maintenance prédictive de systèmes industriels ou la détection automatique de défauts dans une chaîne de production, nécessitaient de disposer d'une expertise interne peu accessible, à l'origine d'une asymétrie entre petites et grandes structures.

Les innovations dans l'IA, et en particulier l'arrivée de modèles généralistes, ont pu réduire ces barrières à l'entrée. Ces modèles sont plus faciles d'accès et nécessitent un entraînement interne moins important. Cependant, les grandes entreprises gardent l'avantage car elles disposent d'équipes susceptibles d'être facilement réorientées vers ces nouveaux outils.

Les entreprises voient principalement deux raisons différentes d'adopter des outils d'IA : l'optimisation des coûts par l'automatisation plus poussée de certaines tâches, qui génère des gains de productivité ; la capacité à créer de nouveaux produits.

Les agents conversationnels servant d'intermédiaire pour la relation client peuvent assurer une tâche précédemment effectuée par des humains ; on peut aussi considérer qu'ils apportent un service nouveau grâce à une plus grande facilité d'utilisation par leurs usagers. Les outils d'IA chargés de missions d'assistance juridique comme la veille ou la rédaction assistée de clauses contractuelles sont aussi très répandus ; un gain de productivité allant jusqu'à 20 % est parfois évoqué. Cet exemple illustre la nouveauté de la révolution de l'IA pour le monde du travail, par contraste avec les révolutions industrielles précédentes : pour la première

fois, ce sont les tâches intellectuelles qui sont automatisables.

L'IA peut aussi être utilisée dans les processus industriels, comme les inspections d'infrastructures et de matériels, ou pour fiabiliser les chaînes de production. Enfin, l'impact de l'IA sur la recherche ne doit pas être négligé. Elle ouvre la voie à des transformations majeures dans certains secteurs, comme la découverte de nouveaux matériaux pour l'industrie des semi-conducteurs ou le développement de médicaments. Les gains de productivité apportés par l'IA sont difficiles à quantifier ; Philippe Aghion et Simon Bunel ont proposé une estimation non exhaustive de 0,7 point de croissance chaque année pendant dix ans.

Si l'utilisation de l'IA en entreprise semble prometteuse, elle comporte aussi des risques qui ralentissent son acceptation. De nombreux chefs d'entreprise, en particulier dans les PME et ETI, qui bénéficient d'un moindre accompagnement pour s'acclimater à ces outils, sont réticents à les intégrer. Par manque d'information, ils peuvent avoir des difficultés à identifier les cas d'usage adaptés à leur entreprise. Des organisations comme le Hub France IA ou des programmes tels que Osez l'IA visent à améliorer la compréhension de cette technologie par les acteurs concernés.

Un deuxième frein important est la réticence des équipes face à l'incorporation non concertée de l'IA. La crainte de se voir remplacé par un algorithme aboutit souvent à son rejet par les personnels. La difficulté des employés à mettre efficacement ces outils en pratique ralentit aussi la diffusion de l'IA. Des chercheurs du Massachusetts Institute of Technology (MIT) estiment que 90 % des projets d'IA échouent, probablement parce que la méthode d'intégration de l'IA a été mal conçue. De ce fait, des projets mieux réfléchis, préparés avec les salariés et articulés avec une évolution de l'organisation du travail pourraient avoir une plus grande chance de succès. Ceci n'est toutefois pas toujours à la portée de toutes les entreprises. L'IA a en tout cas à terme un potentiel important de transformation des méthodes de travail, et son impact à venir sur les emplois reste difficile à prédire.

Les entreprises doivent pouvoir disposer d'outils souverains en mesure de répondre à leurs cas d'usage

L'intégration d'outils d'IA dans les entreprises fait courir le risque d'une augmentation de la dépendance aux acteurs extérieurs, notamment étrangers, qui est déjà très élevée dans le domaine du numérique en général. C'est un facteur que les entreprises souhaitant intégrer l'IA dans leurs processus doivent prendre en compte dans leurs analyses de risque. La dépendance peut être un danger majeur, voire rédhibitoire. Elle pose en effet plusieurs problèmes. Les enjeux de sécurité liés à la possibilité de fuites de données viennent au premier rang. Avec un nombre restreint d'acteurs, les entreprises manquent de choix dans leurs outils et sont à la merci de leurs fournisseurs. La disparition de Sora, logiciel de génération de vidéo par IA créé par OpenAI, est un

exemple susceptible de se reproduire. Or les entreprises ne peuvent prendre le risque de se positionner sur des outils dont la pérennité n'est pas assurée. Par ailleurs, la taille et la position dominante des sociétés qui proposent les grands modèles d'IA générative limitent toute discussion avec les entreprises qui souhaitent utiliser leurs produits, ce qui rend difficile l'acquisition d'une solution personnalisée.

Un outil, l'indice de résilience numérique, peut aider à apprécier le niveau de dépendance et aider à identifier des leviers pour s'en défaire. Encore faut-il disposer de solutions alternatives.

Il ne suffit pas d'avoir accès à des outils en source ouverte, qui sont le plus souvent américains ou chinois, même s'ils permettent d'éviter les solutions propriétaires. Leur gouvernance échappe de toute façon aux acteurs européens alors que c'est justement le sujet clef. La diffusion de l'IA au sein de l'économie doit donc s'appuyer sur une production souveraine. C'est un objectif de la stratégie nationale pour l'IA.

L'écosystème français est dynamique, fort de plusieurs centaines de start-up. Elles sont soutenues par différents mécanismes, notamment France 2030. La filière dispose de moyens pour regrouper ses forces. L'AI Factory France a été lancé en 2025 dans le cadre du programme européen EuroHPC. Il vise à assurer la cohérence d'ensemble du secteur. Les start-up y trouvent un accès unifié aux moyens nécessaires à l'innovation que sont les capacités de calcul, les données et les modèles, ainsi qu'un support technique. Le Grand équipement national de calcul intensif (Genci), l'Inria et d'autres acteurs académiques participent à cet effort.

Les liens entre la recherche académique et les start-up sont solides et fructueux. Dans un domaine en permanente évolution, il est essentiel de continuer à favoriser le développement de nouvelles technologies. L'IA a souvent connu des cycles, les réseaux de neurones étant un exemple d'une technologie ancienne qui n'est devenu dominante que très récemment. Il est difficile de prédire dans quelle voie les prochains bouleversements de l'IA emmèneront le secteur. Étudier une large gamme de techniques est un gage de résilience pour faire face à cette incertitude. La diversité des recherches menées actuellement en France promet une adaptabilité qui peut s'avérer précieuse dans la compétition internationale.

L'importance prêtée à l'innovation de rupture s'étend aussi aux applications et aux usages. Le dispositif Pionniers de l'IA soutient ainsi les projets de recherche et développement dans des domaines stratégiques tels que la santé, l'industrie, la transition écologique ou la sécurité.

Enfin, les mécanismes comme l'AI Factory France aident à faire émerger des communs d'usage, réfléchis avec les entreprises clientes qui ont souvent des besoins proches. On s'assure ainsi que les outils produits par les start-up trouvent leurs utilisateurs. C'est l'un des intérêts d'un écosystème local que les demandes des utilisateurs soient bien comprises par les fournisseurs de technologies. En parallèle, une politique industrielle du logiciel doit être mise en place, afin de disposer de

communs logiciels pour répondre à ces demandes. Ceci doit permettre d'éviter la prolifération désordonnée des briques technologiques nécessaires à leur réalisation et les risques d'arrêt de maintenance.

La diffusion de l'IA fait peser des risques sur la société et les entreprises qui ne seront pas évités par son rejet

L'adoption de l'IA par la société est encore plus rapide que celle du numérique à ses débuts. Les outils d'IA grand public sont très accessibles, et peuvent avoir des applications directes dans la vie quotidienne. L'utilisation grandissante d'agents conversationnels par les entreprises pour la communication avec leurs clients est à la fois un signe et un facteur de cette diffusion. Tout porte à croire que les étudiants actuels se servent massivement d'outils d'IA dans le cadre de leurs études. L'arrivée dans l'emploi de générations complètement acclimatées à ces usages devrait interroger les entreprises, qui seront obligées de s'adapter à ces publics et ces employés.

Alors que certaines entreprises sont encore réticentes à adopter l'IA par peur des risques associés, le danger que font peser les usages cachés de l'IA, ou *shadow AI*, est encore plus grand. Contrairement à une utilisation autorisée par la hiérarchie et prévue par l'entreprise, l'adoption sans contrôle d'outils d'IA grand public par les salariés pour effectuer des tâches à caractère professionnel leur fait prendre des risques juridiques, au-delà même du risque de fuite de données. Les secteurs de l'informatique et de la santé y sont particulièrement exposés, mais tous sont concernés.

Imposer des règles strictes restreignant l'utilisation de l'IA au sein de l'entreprise ne réglera pas seul ce

problème. La formation continue peut être utile pour acculturer les personnels, mais seuls 21 % des salariés ont suivi une telle formation, selon une étude réalisée en décembre 2025 par Ipsos. Ce n'est d'ailleurs pas qu'une question de compréhension des enjeux : il sera difficile d'endiguer la diffusion de l'IA dans la société. Empêcher le *shadow AI* passera par l'adoption volontaire et organisée par l'entreprise d'outils d'IA de confiance.

L'utilisation d'IA dans la vie quotidienne est aussi porteuse de risques. Des contenus générés artificiellement inondent internet, ce qui favorise la mésinformation, voire la diffusion de *fake news* et la désinformation. Les LLM facilitent les cyberattaques sur les plans qualitatif et quantitatif. En réponse, certains demandent la mise en place d'une formation à l'IA et à ses risques dès le plus jeune âge. Le développement d'une forme de littératie en IA semble faire consensus. C'est pourtant une demande particulièrement exigeante car on parle là d'un domaine extrêmement complexe dans lequel les bouleversements technologiques sont fréquents. Les utilisateurs ne sauraient être tenus pour seuls responsables des dangers que font peser les outils d'IA sur la société. Les entreprises d'IA ne peuvent échapper à leur propre responsabilité, entendue ici sur un plan moral.

L'Union européenne s'est dotée d'un cadre réglementaire parmi les plus avancés au monde, reposant sur le *Digital Services Act* (DSA) et le règlement européen sur l'intelligence artificielle (AI Act). La protection de la société – qui est la raison d'être de ces réglementations – invite à les appliquer de façon rigoureuse et à sanctionner les entreprises qui ne s'y conformeraient pas. Les institutions ont, elles aussi, une responsabilité morale envers la société.

Recommandations

- Renforcer le soutien à l'écosystème de recherche public-privé en IA, notamment en s'assurant de la disponibilité de moyens de calcul et de données ; mettre en relation les entreprises utilisatrices de solutions d'IA avec les start-up innovantes pour favoriser le passage aux applications ;
- Encourager la diversité des recherches fondamentales en IA pour susciter l'apparition d'innovations de rupture, favoriser l'émergence d'une IA de confiance et garantir l'adaptabilité du secteur ;
- Encourager le développement d'IA frugales efficaces dans le cadre des efforts pour mettre au point une IA souveraine ;
- Approfondir les études visant à évaluer l'impact climatique de l'IA et du numérique ; favoriser les cas

d'usage de l'IA qui bénéficieraient à la transition énergétique et à la réduction des émissions de CO₂, et en faire un avantage de différenciation face à la concurrence ;

- Accompagner la diffusion de l'IA dans les entreprises et la société par la professionnalisation de la formation continue et initiale ;
- Mobiliser les entreprises sur la lutte contre l'usage caché de l'IA ;
- Assurer un strict respect de la réglementation française et européenne en matière d'IA pour protéger les Français face aux nouvelles menaces amplifiées par l'IA.

Pour consulter le rapport :

<https://www.assemblee-nationale.fr/dyn/17/organes/delegations-comites-offices/opecst>

<https://www.senat.fr/opecst>

Compte X : @OPECST_

Assemblée nationale - 126 rue de l'Université - 75355 Paris 07 SP – Tél : 01 40 63 26 81 – Mél : secretariat-opecst@assemblee-nationale.fr
Sénat - 15 rue de Vaugirard - 75291 Paris Cedex 06 – Tél : 01 42 34 27 20 – Mél : opecst-secretariat@senat.fr