ASSEMBLÉE NATIONALE

17e LÉGISLATURE

Compte rendu

Office parlementaire d'évaluation des choix scientifiques et technologiques

Jeudi 2 avril 2026
Séance de 9 heures 30

Compte rendu n° 222

SESSION ORDINAIRE DE **2025-2026**

Présidence de M. Stéphane Piednoir, *président*

ASSEMBLÉE NATIONALE

Office parlementaire d'évaluation des choix scientifiques et technologiques

Jeudi 2 avril 2026

– Présidence de M. Stéphane Piednoir, sénateur, président de l'Office –

La réunion est ouverte à 9 h 35.

Audition publique sur « Où va l'IA ? Quelles innovations pour quels usages ? »

M. Stéphane Piednoir, sénateur, président de l'Office. – L'ordre du jour de cette réunion de l'Office parlementaire d'évaluation des choix scientifiques et technologiques (Opecst) est consacré à une audition publique sur les nouveaux développements de l'intelligence artificielle (IA). Nous nous intéressons à ce sujet, dont les avancées extraordinaires donnent parfois le vertige, depuis de nombreuses années et nous y avons consacré un rapport fin 2024.

Il n'y a pas un jour sans que les progrès de l'intelligence artificielle soient évoqués, soit pour en vanter les bénéfices, notamment en matière de diagnostic médical, soit pour en déplorer les effets négatifs, particulièrement sur l'emploi – avec des menaces qui pèsent sur l'avenir de certaines filières – et l'environnement. Parmi les craintes, nous évoquions hier en commission de la culture du Sénat la protection du droit d'auteur et les problèmes liés à l'usage des œuvres de création dans l'entraînement des modèles d'intelligence artificielle.

Le rapport de l'Office adopté il y a un peu plus d'un an revenait sur les origines de l'intelligence artificielle, en remontant assez loin dans le temps, et annonçait certaines des évolutions que nous constatons actuellement. Mais d'autres arrivent plus vite que nous ne l'imaginions. Nous avons donc besoin de comprendre ce que sont les nouveaux modèles de raisonnement et comment les modèles de monde vont dépasser, voire remplacer, les grands modèles de langage que nous connaissons. Trois ans seulement après l'avènement des chatbots comme ChatGPT ou Gemini, il est déjà question d'une troisième révolution de l'intelligence artificielle.

C'est pour cette raison que l'Office a décidé de faire un point sur l'évolution de ces technologies, en y associant Corinne Narassiguin et Alexandre Sabatou, qui ont élaboré le rapport de novembre 2024 au côté de Patrick Chaize, qui ne peut être présent ce matin. Ils animeront nos débats d'aujourd'hui.

Avant de leur passer la parole, je voudrais remercier les deux membres du Conseil scientifique de l'Office, instance renouvelée après mon élection à la présidence de l'Office en 2023, qui sont parmi nous ce matin, Daniel Andler et Raja Chatila. Ce dernier va introduire la matinée.

M. Raja Chatila, membre du conseil scientifique de l'Office, professeur émérite à Sorbonne Université. – L'intelligence artificielle est un domaine scientifique ancien mais en plein essor. Il est marqué par des progrès continus en matière de performance des systèmes, grâce à des algorithmes de plus en plus efficaces, mais aussi ponctués par des innovations de rupture, qui viennent totalement réorienter les recherches. Ce fut le cas en

2017, lorsqu'une nouvelle architecture de réseau neuronal – les transformeurs – a réussi, pour la première fois, à traiter automatiquement le langage naturel. C'était un sujet très difficile, qui avait déjà connu de nombreuses évolutions.

Ce type d'architecture, sur laquelle reposent les grands modèles de langage (LLM, *Large Language Model*), est devenu incontournable dans nombre d'applications, en particulier avec l'arrivée des systèmes agentiques qui permettent de combiner plusieurs grands modèles de langage ayant chacun leurs spécialités et de leur faire accomplir des tâches ensemble.

Cette matinée va permettre de faire un point de situation, en abordant à la fois les progrès fondamentaux et les applications de ces technologies, ainsi que l'extension des usages. Même si plusieurs journées seraient nécessaires pour en faire l'inventaire, les différents intervenants nous guideront dans ce paysage en pleine évolution. Sans déflorer leurs propos, je voudrais évoquer brièvement quelques usages qui me semblent soulever des défis majeurs du point de vue scientifique, technique, humain et sociétal. J'en détaillerai deux que nous devons, selon moi, suivre avec attention.

Le premier défi concerne la transformation des grands modèles de langage en modèles multimodaux pouvant associer la langue naturelle à des textes, des images, des vidéos ou des sons. Au-delà des systèmes d'agent conversationnel ou de la génération de données, ils peuvent être utilisés pour la perception et l'interprétation du monde, la prise de décisions et le contrôle des actions réalisées par des robots physiques. La robotique, que certains appellent l'IA physique, est l'une des voies de recherche. Elle vise à sortir les grands modèles de langage de leurs applications actuelles pour les plonger dans le monde réel. Y parviendra-t-elle et comment ? Des robots, en particulier humanoïdes, seront-ils capables de comprendre le monde, de le représenter et d'y évoluer avec intelligence ? C'est un projet qui existe depuis longtemps. Des progrès ont été réalisés sur la conception mécanique, notamment aux États-Unis et en Chine, mais beaucoup reste à faire et nous n'en sommes probablement qu'aux balbutiements.

Vous connaissez le défi que s'est lancé Yann Le Cun : la réalisation de *world models* fondés sur de grands modèles de langage. Pour que de tels travaux aboutissent, de nouvelles orientations de recherche, autour des modèles neurosymboliques par exemple, seront peut-être nécessaires. Aujourd'hui, l'intelligence artificielle repose principalement sur des modèles neuronaux. On pourrait chercher à y associer des modalités de raisonnement symbolique. Pour le moment, ces deux approches restent toutefois assez antinomiques. Certains modèles de langage sont qualifiés de modèles de raisonnement, mais ils ne créent qu'une sorte d'illusion de raisonnement.

Si nous parvenons à surmonter ces difficultés, la question de l'impact de l'automatisation des tâches sur les emplois se posera en des termes nouveaux. Elle est déjà évoquée très fréquemment pour les tâches nécessitant des capacités cognitives, mais, dans ce contexte, l'intelligence artificielle pourrait aussi concerner des tâches matérielles.

Le second défi est à la fois un défi de recherche et d'usage. Il concerne l'utilisation de systèmes d'intelligence artificielle pour l'interprétation des données cérébrales et la réalisation d'interfaces cerveau-machine, grâce auxquelles une personne peut commander des dispositifs ou qui peuvent influencer le fonctionnement du cerveau. Ces neurotechnologies numériques, qui sont déjà utilisées, permettent d'améliorer l'état de certains patients atteints d'affections neurologiques, en particulier la maladie de Parkinson, ou d'apporter une aide à

des personnes en situation de handicap. Elles pourraient toutefois être mises en œuvre en dehors du domaine médical, par exemple dans le jeu ou l'éducation. Or leur effet est très mal connu. En outre, le recours à de grands modèles de langage associant l'imagerie fonctionnelle et le langage, qui est le véhicule de la pensée, va au-delà de la science et pose une question éthique fondamentale, celle de l'intrusion dans les esprits.

L'utilisation de l'IA générative pour l'éducation et la formation, ainsi que sa personnalisation, qui en fait l'attrait, soulèvent aussi des interrogations. Elles font l'objet de nombreuses réflexions, quel que soit le niveau d'apprentissage. L'enjeu est de déterminer si nous apprenons de la même manière avec une machine ou avec un être humain, qu'il s'agisse d'un enseignant, d'un tuteur, d'un collègue, d'un ami ou des parents. La formation des esprits étant essentielle pour la formation des citoyens, l'impact sur la société peut être important. Certains usages, comme le dialogue avec un psychologue artificiel, peuvent avoir des conséquences néfastes, en particulier pour des personnes vulnérables.

L'usage de l'intelligence artificielle dans la recherche scientifique permet d'accélérer certaines découvertes – comme la synthèse de protéines pouvant aboutir à la production de médicaments, travaux qui ont valu un prix Nobel aux chercheurs concernés – mais peut remettre en question la méthode scientifique elle-même. Elle ne serait plus fondée sur la compréhension humaine des phénomènes physiques, qui dépasse la simple observation et constitue la quête de la science, mais essentiellement sur le traitement de données. C'est également un défi nouveau.

Il faut garder à l'esprit que les systèmes d'intelligence artificielle sont fondés sur des modèles d'apprentissage statistique. Les limites inhérentes à cette nature statistique peuvent conduire à des erreurs d'interprétation et à un mélange indiscernable du vrai et du faux, dont les conséquences peuvent être désastreuses. Pour cette raison, le contrôle humain est mis en avant, mais sera-t-il toujours possible ? Les travaux sur la sûreté et la sécurité de l'IA qui seront présentés par M. Grimonpont sont, de ce point de vue, très importants.

Pour conclure, l'avenir semble extrêmement enthousiasmant et attirant, mais les défis scientifiques à résoudre pour réaliser des systèmes intelligents et l'ubiquité de leurs applications ne doivent pas nous faire oublier les principes éthiques, les obligations réglementaires – en particulier celles issues du règlement européen sur l'IA –, ainsi que les impacts environnementaux.

Progrès récents et pistes de développement de l'IA

M. Alexandre Sabatou, député, vice-président de l'Office, co-rapporteur. – Pour cette première table ronde consacrée aux progrès récents et aux pistes de développement de l'IA, je commencerai par une question simple que beaucoup de personnes se posent. La plupart des tests de performance montrent une progression exponentielle des modèles d'IA générative, en particulier grâce aux modèles de raisonnement. Pensez-vous que cette progression finira par ralentir et atteindra un plafond ? Lors des auditions qui ont nourri notre rapport de 2024, M. Le Cun suggérait que les avancées seraient beaucoup moins importantes dans les années à venir. Il avait utilisé une métaphore pour expliquer qu'à force de creuser, nous allions finir par tomber sur une couche de granit. À l'inverse, M. Bengio estimait que l'intelligence artificielle générale pourrait arriver très vite.

Avant de répondre à ma question, vous pouvez, si vous le souhaitez, commencer par un propos liminaire.

M. Patrick Pérez, directeur général de Kyutai. – Je voudrais attirer votre attention sur les mouvements de fond concernant les grands modèles de fondation et les IA génératives. Ces modèles généralistes ne résument pas l'IA mais sont les plus visibles. Leur taille va de quelques milliards à quelques centaines de milliards de paramètres, et ils sont entraînés à très grande échelle avec des données souvent issues d'internet.

Au cours de l'année écoulée, nous avons pu observer trois grandes tendances. Tout d'abord, les modèles de langage, comme les robots conversationnels, ont connu une progression très rapide en matière de puissance et de précision. Pour donner une image, ils étaient du niveau de l'école primaire et sont, dans certains domaines, passés au niveau universitaire. Ils ont participé aux Olympiades de mathématiques et peuvent servir d'assistants pour des recherches pointues en mathématiques ou d'autres sciences. L'une des grandes révolutions à l'œuvre est leur utilisation dans le développement de logiciels. En l'espace d'un an, les agents de développement – qui sont parfois dans les mains de chercheurs en IA, ce qui n'est pas anodin – ont bouleversé cette activité.

Ce phénomène s'explique par l'introduction dans les modèles de techniques de raisonnement, qui leur permettent de prendre leur temps au moment de répondre et de traiter les questions compliquées en les décomposant en sous-problèmes plus simples. Ceux qui les utilisent constatent qu'ils obtiennent ainsi des réponses de très bonne facture et très utiles, y compris sur des sujets complexes. Par ailleurs, les modèles peuvent désormais avoir recours à des outils extérieurs. Nous connaissions déjà la génération augmentée par récupération (RAG, *Retrieval Augmented Generation*), qui consiste à faire appel à une base de connaissances autre que les sources de données utilisées pour l'entraînement, mais il est désormais possible de mobiliser des outils de calcul, d'autres applications ou d'autres IA. Grâce à ces deux évolutions complémentaires, les modèles qui étaient au départ conversationnels deviennent plus autonomes et, pour utiliser un néologisme, plus agentiques.

La deuxième grande tendance concerne l'extension des modèles à d'autres modalités, qui leur permettent de devenir multimodaux. Au lieu d'avoir essentiellement du texte en entrée et en sortie, il peut s'agir de sons et d'images ou, surtout en robotique, de perceptions tridimensionnelles, de nuages de points ou de scans laser. Certains modèles commencent à être extrêmement puissants. Pour citer un exemple proche des activités de mon laboratoire, l'IA vocale a fait des progrès phénoménaux en une année et les capacités d'interaction vocale entre l'humain et la machine sont devenues impressionnantes. Certains usages sont bénéfiques, en matière d'accessibilité et d'aide au handicap par exemple, mais d'autres peuvent être problématiques.

Pour la partie générative de ces modèles multimodaux, les progrès réalisés en un an en matière de génération de vidéos sont particulièrement marquants. Cette génération est de mieux en mieux contrôlée par du texte, des images de référence ou du son pour faire parler un avatar par exemple. Là encore, ceci ouvre la possibilité d'usages problématiques ou vertueux. Ces technologies permettent notamment de démocratiser les outils de création ou de créer des données synthétiques. Je reviendrai sur ce dernier point. Les progrès sont également spectaculaires dans le domaine de la musique. Les niveaux atteints en matière de création automatique de paroles, de musique et d'interprétation font que les plateformes de streaming sont désormais inondées de morceaux synthétiques, avec toutes les questions qui en découlent.

Enfin, la dernière tendance, peut-être moins connue, concerne le développement d'IA capables d'entraîner des IA. En créant à grande échelle du son, des images ou des vidéos, elles produisent de la donnée de qualité, contrôlée, qui pourra entraîner de nouvelles IA. Le fait que les données synthétiques inondent le web fait craindre une détérioration des performances des nouvelles générations qu'elles serviraient à entraîner. Ce n'est toutefois qu'un versant de l'histoire. L'autre versant, c'est que les données synthétiques sont très précieuses et qu'il est possible de faire de l'annotation et de l'évaluation synthétiques, tâches rébarbatives et pas toujours bien payées, qui étaient jusqu'à récemment effectuées par des humains. Selon moi, ces évolutions constituent un progrès. En outre, les techniques dites de distillation permettent à des modèles, parfois coûteux et très lents mais déjà entraînés, d'entraîner de nouveaux modèles plus légers, plus rapides et de meilleure qualité. C'est l'un des facteurs qui expliquent les changements que nous avons observés au cours de l'année écoulée.

M. Jean Ponce, professeur au département d'informatique de l'École normale supérieure Paris Sciences et Lettres (ENS PSL) et au Courant Institute de New York University (NYU), directeur associé du cluster Paris Artificial Intelligence Research Institute (PR[AI]RIE) et président de la start-up ENHANCE LAB. – Je suis d'accord avec ce qui vient d'être dit par Patrick Pérez sur le rôle de l'IA en matière de développement des logiciels. Au stade de la recherche, mes collègues de NYU qui travaillent sur la robotique développent des prototypes très rapidement grâce à des logiciels d'IA. Au sein de ma start-up, mes ingénieurs s'en servent également pour aller plus vite.

Je suppose que vous m'avez invité pour parler notamment des *world models* de Yann Le Cun, puisque j'ai travaillé avec lui sur ce sujet. Le principe est de construire des agents disposant d'un modèle mental de leurs interactions avec un environnement, qui peut être physique, appris à partir de données sensorielles. Ceci permet d'en simuler la dynamique, qu'il s'agisse par exemple de la chute d'une pierre sous l'action de la gravité ou des conséquences pour un robot de se déplacer dans telle ou telle direction.

Cette idée, dont le nom a été introduit en Suisse par Jürgen Schmidhuber, n'est pas nouvelle. Née dans les années 1990, elle est popularisée et instanciée aujourd'hui par Yann Le Cun depuis qu'il a quitté Meta et fondé AMI Labs. Son objectif est de créer une architecture permettant de développer des JEPA – peu importe en ce lieu la signification de l'acronyme – qui apprendraient de manière autosupervisée, c'est-à-dire sans supervision externe, en exploitant la cohérence interne des données. L'une des différences principales avec les grands modèles de langage, qui apprennent aussi de manière autosupervisée, est liée au fait qu'il ne s'agit pas de prédire ou de simuler directement des données, mais plutôt une abstraction de ces données, sous la forme d'une représentation qui a été apprise. On espère ainsi éviter de perdre le modèle dans les détails superflus présents dans les données brutes. L'objectif est d'avoir des modèles utiles dans des tâches concrètes, par exemple la planification robotique. Dans ce domaine, un pas reste à franchir, car les modèles actuels sont surtout entraînés pour simuler la dynamique. Quelques applications intéressantes sur des tâches relativement simples en robotique ont néanmoins eu lieu. L'avenir nous dira si ces promesses se concrétisent.

En tout cas, il est essentiel d'avoir des approches alternatives aux LLM, que ce soient les *world models* ou d'autres. Pour irriguer l'IA et permettre à ces technologies de continuer à progresser, nous avons besoin de nouvelles idées. Par exemple, le succès de l'apprentissage profond est dû à deux inventions scientifiques : les réseaux convolutifs introduits par Yann

Le Cun et ses collègues d'AT&T et les transformeurs évoqués précédemment par Raja Chatila. Pour aller plus loin, il faudra de nouvelles inventions et, dans ce domaine, la recherche publique joue un rôle fondamental. L'industrie n'est pas là pour faire de la recherche à long terme. Son but est de gagner de l'argent. Si nous voulons nous projeter à long terme et disposer de nouvelles inventions, il est donc essentiel de soutenir la recherche publique.

Le résumé de votre rapport de 2024, que par ailleurs j'ai trouvé très bien fait, disait que la recherche en IA était dominée par la recherche industrielle. Selon moi, c'est faux, en tout cas pour les domaines que je connais très bien, comme la vision artificielle et la robotique. Les LLM ont, en revanche, besoin d'énormes capacités de calcul et il est en effet plus difficile à un laboratoire académique d'en disposer.

Pour terminer, j'évoquerai deux domaines où je pense que l'IA doit jouer un rôle majeur. Le premier, ce sont les sciences. J'ai la chance de travailler avec des astronomes et des biologistes. Ils savent comment fonctionnent leurs télescopes ou leurs microscopes. Nous utilisons leurs modèles, mais nous apportons l'IA pour les assister. Dans mon cas, ce n'est pas une nouvelle manière d'aborder la recherche. Patrick Pérez évoquait les données synthétiques. Elles ont souvent le défaut de ne pas être réalistes mais, dans le monde scientifique, nous pouvons en construire d'extrêmement réalistes, car nous disposons de modèles solides pour tout ce que nous faisons.

Le second, c'est la robotique. C'est un champ intégrateur parce que, pour qu'un robot fonctionne, il faut maîtriser la perception, la planification, etc. En outre, il ne suffit pas de garantir que telle chose est une vache ou une poule avec un taux de confiance de 90 %. Si quelque chose ne va pas, le robot va se casser ou faire du mal à quelqu'un. En ce moment, on prête peut-être trop d'attention à la robotique humanoïde. Elle est très populaire, avec des machines qui sont extraordinaires, mais est-elle nécessaire pour réaliser la plupart des tâches quotidiennes ?

M. Jean-Frédéric Gerbeau, directeur général délégué à la science de l'Institut national de recherche en sciences et technologies du numérique (Inria). – L'Inria est un institut de recherche en sciences et technologies du numérique. Depuis quelque temps, il joue également le rôle d'agence de programmes dans le numérique et, à ce titre, coordonne le volet de la stratégie nationale pour l'intelligence artificielle consacré à l'enseignement supérieur et à la recherche.

J'insisterai sur trois points illustrant les progrès récents, en évitant de redire ce qui a été dit par les intervenants précédents. J'évoquerai les agents d'IA, puis le lien entre l'IA et le monde physique et enfin l'impact de l'IA sur les activités scientifiques.

Les agents d'IA connaissent un très fort développement actuellement. C'est un concept ancien qui a été remis au goût du jour par les LLM, qui peuvent intervenir directement en tant qu'agents ou en tant qu'orchestrateurs de différents agents. Il offre des perspectives très intéressantes en matière d'automatisation de tâches complexes et d'action sur le monde réel. Il pose toutefois beaucoup de questions liées à la cybersécurité, à la perte de contrôle et à la régulation.

Beaucoup de choses ont déjà été dites sur le lien entre l'IA et le monde physique, en particulier sur les *world models*, et je ne peux qu'abonder dans le sens de mes prédécesseurs pour rappeler que l'IA ne se limite pas aux LLM. Elle existait avant eux et de nombreuses

voies alternatives subsistent. Les problèmes peuvent être abordés de deux façons, soit par des approches utilisant des données massives et nécessitant énormément de calculs, soit par des approches fondées sur la connaissance physique du monde telle qu'elle s'est construite depuis le siècle de Newton. L'enjeu de « l'IA physique » réside dans le dosage entre ces deux tendances. La robotique a fait le choix de s'engager de façon déterminée dans la première direction et d'accumuler des données extrêmement nombreuses à partir desquelles sont réalisés des calculs très coûteux. D'autres options sont néanmoins possibles. Des équipes françaises sont très bien positionnées dans ce domaine. Elles mixent des approches basées sur la connaissance des équations physiques qui sous-tendent les phénomènes et des approches basées sur des modèles. C'est notamment le cas de l'équipe Willow à l'Inria, qui est dirigée par Justin Carpentier et à laquelle participe Jean Ponce.

J'ai lu le rapport de l'Office de 2024, qui insistait, à juste titre, sur les risques de dépendance technologique. Or trouver des voies alternatives au gigantisme actuel en matière de données ou de calcul est un moyen de s'en prémunir. L'équipe Willow a réalisé des progrès spectaculaires en algorithmique, qui permettent d'envisager de faire des calculs sur de simples unités centrales de traitement (CPU, *Central Processing Unit*) plutôt que sur des processeurs graphiques (GPU, *Graphics Processing Unit*). Le coût serait ainsi divisé par 100 et cette solution libèrerait d'une possible dépendance technologique.

S'agissant de l'impact de l'IA sur les activités scientifiques, il faut être conscient qu'elle intervient désormais dans tout le processus de la recherche. Des IA proposent des sujets d'étude et participent à l'élaboration de bibliographies et à la rédaction ou l'évaluation d'articles. De ce fait, le nombre d'articles soumis dans les conférences scientifiques a connu une augmentation sans précédent. Dans certaines conférences d'informatique, il a été multiplié par deux entre 2024 et 2025. La situation est comparable pour les appels à projets.

Au-delà des aspects technologiques, cette situation crée un risque de dépendance et surtout de captation, parce que des start-up et de grands acteurs de l'IA comme OpenAI développent des outils, des sortes de Google Docs pour chercheurs, intégrés dans des environnements complets d'IA. S'ils sont adoptés par un grand nombre de chercheurs – ce qui est fort possible –, ils vont créer une dépendance technologique mais surtout donner aux propriétaires de ces outils un accès direct, non pas à la recherche déjà faite, mais à celle qui est en train de s'écrire. Pour éviter ces phénomènes, qui peuvent s'imposer très rapidement en raison des effets de réseau dont bénéficient ces entreprises, nous devons proposer des solutions alternatives, notamment à l'outil Prism d'OpenAI. L'agence de programmes pilotée par l'Inria vient de lancer un programme sur le sujet. Une telle initiative est toutefois insuffisante. Nous aurions besoin de passer à l'échelle à un niveau européen.

Une autre manifestation très importante de l'impact de l'IA sur les activités scientifiques concerne la programmation. Nous sommes passés d'IA qui aidaient à générer des fragments de programme à des IA qui couvrent tout le cycle de création des logiciels, depuis les spécifications jusqu'aux tests. Certains développeurs seniors reconnaissent qu'ils n'écrivent quasiment plus de lignes de code. Ils interagissent avec l'IA et n'ont même plus le temps de relire le code qu'elle génère. Les conséquences économiques de ces évolutions sont majeures, comme l'illustrent par exemple les annonces récentes d'Oracle. Un rapport de la Réserve fédérale des États-Unis de fin 2025 en montre les effets sur le recrutement des jeunes programmeurs et l'intégration des nouvelles générations dans les équipes de développement logiciel. Les entreprises dont les modèles reposent sur le génie logiciel ont également vu leur capitalisation boursière chuter.

Cette situation peut toutefois être une chance pour des entreprises qui, jusqu'à présent, ne pouvaient pas prétendre jouer un rôle important dans le développement de logiciels. Les outils disponibles leur permettent d'obtenir des résultats qui étaient inatteignables auparavant. De ce point de vue, les cartes sont totalement rebattues. Nous devons cependant être attentifs aux vulnérabilités systémiques qui pourraient résulter de l'utilisation généralisée de quelques outils qui se trouveraient en position de monopole. L'une des solutions serait de les coupler avec les assistants de preuve, peut-être sous la forme d'agents. Cette méthode ancienne, toujours d'actualité, permet en effet de prouver qu'un programme est correct. Des équipes sont très reconnues dans ce domaine en France, comme celle de Xavier Leroy qui, avec ses collaborateurs, a développé un compilateur entièrement certifié, CompCert. Il faudrait sans doute étendre cette approche à l'ensemble de la chaîne des logiciels critiques, notamment en cryptographie, ainsi que dans les protocoles ou les bases de données.

En mathématiques, les progrès ont été spectaculaires depuis un an. Nous sommes passés du niveau d'étudiant à celui de chercheur en mathématiques. Nous assistons à un bouleversement. Certains de mes collègues à l'Inria disent que des démonstrations qui leur auraient pris des semaines peuvent désormais être réalisées en quelques heures. Comme pour la programmation, de très grands mathématiciens, y compris des lauréats de la médaille Fields comme Timothy Gowers au Collège de France ou Terence Tao à l'Université de Californie à Los Angeles (UCLA), combinent les approches par IA, qui permettent de produire des résultats et de faire des démonstrations, avec des assistants de preuve. En France, nous disposons de l'outil Rocq – anciennement appelé Coq. À l'étranger, il existe aussi l'outil Lean. Depuis peu, une communauté de mathématiciens a pris ce virage et, comme pour la production de code, développe les interactions entre des assistants de preuve et l'IA générative.

S'agissant des pistes de développement, j'insisterai sur les incertitudes. Les garanties associées aux modèles d'IA sont encore peu développées, même si l'IA de confiance est évoquée depuis des années. Laurent Simon y reviendra sans doute, en évoquant les approches d'IA symbolique. En matière d'IA statistique, peu de progrès ont été réalisés. La recherche est donc indispensable. Si nous voulons dépasser les usages liés aux réseaux sociaux, à la publicité ou à la génération d'images grand public, il est indispensable d'offrir des garanties et de quantifier l'incertitude des modèles. Or nous sommes très loin du compte.

Un autre point, en partie lié au précédent, concerne l'évaluation des modèles et de leur sécurité. C'est un sujet très ouvert, qui relève désormais de l'Institut national d'évaluation de la sécurité de l'IA (Inésia). Placé sous le pilotage du secrétariat général de la défense et de la sécurité nationale (SGDSN) et de la direction générale des entreprises (DGE), il associe l'Inria, l'Agence nationale de la sécurité des systèmes d'information (Anssi), le Laboratoire national de métrologie et d'essais (LNE) et le pôle d'expertise de la régulation numérique (Peren) et contribue au programme d'évaluation de l'IA que vient de lancer l'agence de programmes dans le numérique.

Nous devons penser l'IA le long du continuum de calcul, c'est-à-dire depuis les centres de données jusqu'en bordure de réseau. Plusieurs actions ont été engagées à ce sujet, en particulier avec la start-up Hivenet ou les Nokia Bell Labs. Les approches distribuées, dont relèvent les IA agentiques, peuvent être très pertinentes, aussi bien dans les secteurs de la santé que de la défense.

Enfin, nous devons réfléchir à des approches plus frugales et moins énergivores, qui ne reposent pas uniquement sur la force brute, les données massives et la puissance de calcul. L'IA embarquée en est un exemple. Raja Chatila a également évoqué en introduction les impacts environnementaux. Comme pour la robotique, chercher des voies alternatives permet aussi de ne pas subir l'agenda dicté par d'autres pays ou par des entreprises qui, parce qu'elles sont les seules à maîtriser le gigantisme, veulent persuader le monde entier qu'il est la seule issue. Or il n'est pas l'unique réponse à toutes nos attentes. Il y a de la place pour d'autres approches et l'Inria soutient plusieurs recherches en ce sens.

L'Inria inscrit naturellement son action dans le cadre de la stratégie nationale pour l'IA. Notre objectif principal est de renforcer la recherche française dans le numérique, en particulier en IA, ce qui suppose d'avoir une logique d'écosystème et de privilégier une approche intégrée entre la formation, la recherche et l'innovation. Les universités, qui sont les seules à couvrir ces trois dimensions, doivent donc être soutenues. La stratégie nationale pour l'IA a ainsi permis la création de neuf clusters, répartis dans toute la France et rattachés à de grandes universités. La Cour des comptes a évalué l'ensemble des dispositifs, pilotés directement ou indirectement par l'Inria, dans un rapport de novembre 2025, en particulier l'accompagnement à la création de start-up technologiques, la politique d'attractivité internationale et des initiatives originales comme Pionniers de l'IA. Selma Souihel y reviendra plus tard dans la discussion, ainsi que sur l'entreprise qui devrait voir le jour autour de Scikit-learn.

Comme Jean Ponce, je souligne que la recherche académique, et en particulier la recherche académique française, n'est pas dépassée. Nous sommes dans la course. Quelques exemples le confirment. En informatique graphique, un nouveau standard mondial, le Gaussian Splatting, a été inventé il y a trois ans par une équipe de l'Inria à Sophia Antipolis. Cet algorithme a détrôné un algorithme de Google qui dominait le marché depuis des années. En météorologie, des modèles aussi précis que ceux de Huawei ou de Google ont été proposés par une équipe de l'Inria à Paris. Développés avec beaucoup moins de ressources, ils permettent pourtant d'obtenir des résultats comparables. En vision par ordinateur – le domaine de Jean Ponce –, des modèles de fondation reposant sur des approches très peu supervisées, voire non supervisées, sont devenus une référence mondiale, comme DINOv2. Ils sont aussi issus de la recherche académique française, comme Scikit-learn, qui est aujourd'hui le standard mondial des *data scientists*. J'ai déjà évoqué les approches innovantes en matière d'IA physique et de robotique. Le logiciel Pinocchio est téléchargé un million de fois par mois et la suite Maestro, développée par Justin Carpentier, sera peut-être le Scikit-learn de la robotique de demain.

Le fossé entre les moyens de la recherche académique et ceux des géants de la tech reste énorme. C'est indéniable. Pourtant, nous continuons à recruter à très haut niveau, en particulier à l'étranger. La fuite des cerveaux, qui a atteint un pic autour de 2018, a diminué et, ces derniers temps, a même eu tendance à s'inverser. Les quelques résultats que je vous ai cités montrent que les armes dont nous disposons permettent à nos équipes de rester dans la course mondiale.

M. Laurent Simon, professeur des universités, titulaire de la chaire de la Fondation Bordeaux Université sur l'IA Digne de Confiance et membre du collège Représentation et Raisonnement de l'Association française pour l'intelligence artificielle. – L'axe de la représentation des connaissances et du raisonnement est présent en France depuis très longtemps et a connu des progrès très importants. Ces travaux sont plus

discrets que les LLM, parce qu'ils ne sont pas forcément au contact des utilisateurs et ne s'expriment pas en langage naturel. Ils utilisent un langage formel bien spécifié, sur lequel peuvent s'appuyer des raisonnements mathématiques solides qui permettent d'établir des preuves.

Que vient faire la représentation des connaissances dans le monde des LLM ? Tout d'abord, j'insiste sur la difficulté de prédire l'avenir. Il est possible d'identifier ce qui fonctionne dans la science à l'heure actuelle, mais en tirer des conclusions pour la suite est compliqué. Comme l'a dit Raja Chatila, il aurait été impossible, en 2018, de prédire le monde que nous connaissons actuellement. Il n'était même pas en projet. Par conséquent, il faut être humble et construire une recherche agile, capable de pivoter à la fois rapidement et massivement pour explorer de nouveaux territoires.

Quelles sont les innovations qui pourraient changer les choses de manière fondamentale ? On dit introduire du « raisonnement » dans les grands modèles de langage. Or ce n'est pas du raisonnement. C'est plutôt comme l'écriture d'un roman dans lequel un personnage fictif essayerait de détailler ses réflexions pour les simplifier. Le raisonnement n'est pas plus solide, mais il est plus précis, ce qui permet à ces systèmes d'obtenir de meilleurs résultats. En revanche, en toute rigueur, ce n'est pas un raisonnement. Certains systèmes peuvent raisonner de manière fausse et produire des conclusions correctes ou l'inverse, ce qui est assez paradoxal.

Un raisonnement repose sur des méthodes formelles qui ont la puissance de la preuve mathématique. Comment intégrer ceci dans un LLM et qu'est-ce que cela peut changer ? Aujourd'hui, les grands modèles de langage sont capables de produire une abstraction symbolique d'un problème, ce que nous faisions depuis longtemps, mais de manière humaine, dans le domaine de la représentation des connaissances. Face à un problème exprimé par l'utilisateur, le LLM peut produire soit du code informatique soit une abstraction à partir de laquelle on pourra produire un algorithme qui effectuera un raisonnement correct et complet. Cette approche est nouvelle. L'abstraction n'est plus produite par un humain mais par une machine et, si l'on est d'accord sur la représentation abstraite du problème, on est certain que le raisonnement est correct.

Nous entrons dans une ère nouvelle dans laquelle les LLM seront capables de produire un algorithme à la volée. Jusqu'à présent, on utilisait des algorithmes sur étagère, plus généraux, pour les appliquer à un problème donné. Ce que peuvent désormais faire les LLM va changer radicalement la donne du numérique dans les années à venir. Face à un problème, ils pourront fournir un algorithme, qui devra être prouvé par des méthodes symboliques, et des astuces pour essayer de résoudre ce problème de manière efficace. Aujourd'hui, ces astuces sont produites par des chercheurs en IA et des spécialistes du problème. Dans le futur, ce sera fait automatiquement. Avoir la possibilité de créer un algorithme spécifique et optimisé dans le temps d'une requête constitue un changement de paradigme. Cela n'était jamais arrivé.

Les algorithmes qui auront été produits de manière automatique devront être vérifiés avec des preuves formelles et pas seulement testés pour satisfaire un pourcentage de précision. Ces preuves et la vérification de ces preuves seront faites par des méthodes symboliques. S'il y a une faute, une réinjection permettra au LLM de produire un algorithme corrigé, comme le ferait un humain. Ainsi, on pourra profiter de l'aspect « créatif » et exploratoire des LLM et de leurs failles d'hallucination. On pourra les laisser faire, puisque

des méthodes formelles arriveront derrière pour filtrer ce qui est possible et ce qui est garanti. Être capable de vérifier une modélisation à la volée et d'y appliquer un raisonnement solide représente un défi pour les domaines de la vérification formelle et de la représentation des connaissances.

En conclusion, j'insiste sur la notion de confiance, qui doit être au cœur des recherches. Qu'est-ce que la confiance dans le numérique ? Pour ma part, je n'ai confiance que dans ce qui est prouvé. Plusieurs méthodes logicielles existent, comme le test. Avant d'être envoyés en production, les algorithmes sont testés. Jusqu'à présent, on ne sait toutefois tester que des algorithmes produits par des humains. Le code des algorithmes produits par des IA ressemblera à un plat de spaghettis et leur but sera peut-être seulement de passer les tests, comme les étudiants qui débutent en informatique. Le domaine de la vérification logicielle va devoir changer de modèle pour vérifier exhaustivement les capacités d'un algorithme et obtenir une preuve formelle de ce qui est proposé.

Aujourd'hui, l'IA me recommande une décision à prendre. Si je la prends, j'en endosse la responsabilité. L'IA doit donc me fournir des informations exhaustives me permettant de peser le pour et le contre pour que je puisse prendre la décision et en endosser à sa place la responsabilité. La confiance n'est pas une histoire d'informaticiens. C'est une réflexion transdisciplinaire, qui doit associer les sciences humaines et sociales, parce qu'en bout de chaîne, c'est l'humain qui reste responsable.

La génération à la volée d'algorithmes pourra se généraliser au web. Aujourd'hui, les pages disponibles sur internet ont été rédigées par des humains. Demain, avec les LLM et l'énorme quantité de données disponibles dans notre environnement, les pages proposées seront synthétisées pour chacun de nous, en fonction de notre âge, de notre profession, de nos besoins ou de nos envies. Elles le seront à partir de données brutes lisibles par des humains mais aussi à partir d'autres qui ne le seront pas ou qui auront été négociées par d'autres agents ; ceux-ci auront payé pour obtenir des données réelles ou accepté d'avoir des données sponsorisées. L'accès au numérique ne sera plus statique mais évanescent et hyper-individualisé. Dans ce contexte, le rôle de la science est d'essayer de tracer les modèles de fonctionnement et de suivre ces mécanismes pour expliquer le rendu proposé aux utilisateurs. C'est un défi majeur à long terme, à la fois scientifique et sociétal.

Nous avons également évoqué la frugalité, qui est un élément essentiel pour démocratiser l'usage de ces IA.

M. Alexandre Sabatou, député, vice-président de l'Office, co-rapporteur. – Vous soutenez tous les IA frugales, dont je comprends le bien-fondé écologique, mais n'est-ce pas aussi une excuse pour justifier les écarts entre la France ou l'Europe et les gigantesques investissements américains ?

M. Patrick Pérez. – L'IA frugale n'est pas uniquement liée à des moyens limités ou à des contraintes créatives, elle correspond aux besoins de certains cas d'usage, qui nécessitent de petits modèles. Dès que l'IA est embarquée, utiliser de gros modèles qui tournent dans le cloud n'est pas une option. C'est le cas dans la robotique, les transports ou les technologies d'assistance aux personnes. L'IA frugale est un objet en soi.

M. Jean Ponce. – Je peux citer l'exemple d'une start-up qui propose une solution qui fonctionne sur des téléphones. Ce ne sont pas d'énormes ordinateurs. La recherche a aussi un côté esthétique. L'idée de construire des systèmes intelligents en observant des milliards et

des milliards de paramètres, avec des démonstrations qui durent des centaines d'heures pour parfaire des imitations, n'est pas très attirante du point de vue de l'approche scientifique. Nous devons réfléchir à des technologies qui ne dépendent pas uniquement d'énormes volumes de données.

M. Laurent Simon. – L'IA peut être frugale en fin de course, mais pour pouvoir dialoguer en langage naturel, elle a besoin de s'autoconfigurer pour avoir une représentation interne du langage. Pour que cette structuration puisse s'effectuer, il faut lui donner un maximum d'exemples de textes humains et de textes algorithmiques. Cette première étape ne peut pas être frugale. En revanche, l'IA peut ensuite être spécialisée dans un domaine particulier. Nous avons déjà évoqué le concept de distillation. Si vous travaillez pour un vendeur en ligne par exemple, vous n'êtes pas obligé de connaître tous les scores du basket aux États-Unis ou l'histoire de tel ou tel pays. On peut donc élaguer ses connaissances, tout en conservant les capacités de compréhension du langage naturel qu'elle n'aurait pas eues si elle avait été entraînée uniquement avec des données très spécialisées. Cette approche permet non seulement de déployer l'IA dans de petits systèmes mais aussi de l'empêcher d'halluciner, en la forçant à cantonner ses réponses à un métier spécifique.

M. Jean-Frédéric Gerbeau. – L'IA frugale n'est pas une excuse avancée quand on ne sait pas faire autrement. C'est d'abord avoir la conscience de vivre dans un monde fini qui connaît des contraintes énergétiques. Beaucoup partagent cette conviction, qu'il est essentiel de prendre en compte. C'est également la volonté de ne pas s'appuyer uniquement sur la force brute et de valoriser des siècles d'accumulation de savoirs pour élaborer, grâce à la combinaison de tout ceci et avec des procédés plus récents, des solutions plus malignes et plus excitantes scientifiquement comme l'expliquait Jean Ponce, qui se trouvent être *in fine* plus frugales. Le discours dominant appelle à toujours plus de gigantisme, mais ce n'est peut-être n'est pas l'unique chemin et explorer des voies alternatives est aussi une motivation naturelle pour la recherche.

M. Raja Chatila. – L'IA frugale représente un défi scientifique majeur. On cherche un changement de paradigme. Le cerveau ne sait pas tout sur tout, mais il ne consomme que quelques dizaines de watts pour raisonner, décider, traiter l'information, etc. Nous avons beaucoup mentionné la notion d'abstraction. Abstraire, à partir de données, une connaissance sur laquelle on peut bâtir un raisonnement de manière frugale, c'est ça le secret de l'intelligence. Ce n'est pas se contenter de trouver la réponse la plus acceptable statistiquement. L'IA frugale n'est donc pas une excuse.

M. Alexandre Sabatou, député, vice-président de l'Office, co-rapporteur. – Excusez-moi d'être un peu taquin.

Vous avez cité Prism AI. Beaucoup de scientifiques m'ont indiqué que la multiplication des publications scientifiques devenait difficile à suivre. Ce genre d'application a donc du succès mais pose la question de la souveraineté. La France devrait-elle développer des IA françaises par spécialité ?

M. Patrick Pérez. – La recherche est mondiale. Par conséquent, je ne pense pas qu'il faille développer des outils nationaux pour aider les chercheurs. En revanche, avoir des outils en *open source* serait une bonne chose.

M. Jean-Frédéric Gerbeau – Dans ce genre de domaine, l'important est de préserver une certaine diversité. Il est normal que les grandes entreprises développent leurs

outils et que certains les utilisent, mais le risque est lié aux effets de réseau. Regardez le nombre de personnes qui ont un abonnement chez OpenAI, parce que d'autres l'utilisent et qu'elles sont obligées de le faire aussi ; de ce fait, elles partagent leurs recherches en cours avec cette société privée. Le problème principal est l'absence de solutions alternatives. Nous avons les moyens de développer des outils similaires en Europe et ainsi de donner le choix aux gens.

M. Laurent Simon. – Un ordinateur assez récent peut être programmé rapidement pour effectuer ce type de tâche en local. Un agent autonome peut recenser tout ce qui est publié et faire des résumés. L'un des avantages des progrès récents est que la plupart des grands modèles, qu'ils soient américains, chinois ou français, sont disponibles et peuvent être distillés assez facilement.

M. Jean Ponce. – Je ne suis pas certain que tous les modèles soient disponibles. Les modalités et les données d'entraînement ne le sont généralement pas. Même les modèles en *open source* ne le sont souvent pas réellement. De bonnes raisons expliquent d'ailleurs que les données ne soient pas disponibles. Si une entreprise les achète, le contrat qu'elle a signé ne lui permet pas d'en ouvrir l'accès à n'importe qui. Les grands modèles réellement en *open source* sont donc extrêmement peu nombreux.

M. Alexandre Sabatou, député, vice-président de l'Office, co-rapporteur. – Monsieur Simon, vous nous avez expliqué qu'une IA pourrait commencer à préparer une réponse et qu'une espèce de balayage avec des connaissances sûres permettrait ensuite de vérifier la fiabilité des informations délivrées. Ce nouveau modèle signerait-il la fin des hallucinations ?

M. Laurent Simon. – Les modèles de ce type ne fonctionnent pas en langage naturel. En matière de preuves mathématiques, il n'y a pas d'hallucinations mais seulement des preuves fausses et des preuves justes, ce qui est un avantage. Dans le langage naturel, on peut dire que quelqu'un hallucine uniquement parce qu'il est créatif. Quand les LLM hallucinent, ils produisent un texte qui n'a pas d'ancrage dans le réel, car ils sont configurés ainsi.

Plutôt qu'attendre du LLM qu'il maîtrise toute la chaîne, de la compréhension de la question à la réflexion interne et à la production d'une solution, on peut le concentrer sur ce qu'il sait faire le mieux, c'est-à-dire produire une abstraction en langage symbolique. Même en 2022, ces évolutions n'étaient pas du tout prévues. On peut lire le langage symbolique, mais chaque mot porte une force sémantique indiscutable. Dans un graphe représentant des interactions médicamenteuses, un lien est mis entre deux nœuds dès lors que les interactions sont validées par des publications scientifiques. Ce lien est indiscutable. Ce n'est pas du langage naturel, avec lequel on peut débattre et dire des choses vraies et fausses dans la même phrase.

Les grands modèles de langage sont curieusement assez efficaces pour produire une abstraction, que ce soit du code informatique, qui est une abstraction d'un problème, ou la représentation syntaxique formelle d'une question. L'utilisateur peut acquiescer et valider cette représentation formelle, puis utiliser une méthode mathématique prouvée pour déduire une réponse de cette formalisation et enfin rebrancher le LLM pour obtenir cette réponse. C'est le mode de fonctionnement des agents. L'architecture existe déjà. Pour le moment, en revanche, ces deux systèmes ne savent pas communiquer. En outre, même si je plaide pour ma paroisse, les méthodes de raisonnement restent confrontées à de nombreuses difficultés.

Elles ont échoué dans le passé, car représenter le monde de manière formelle est très compliqué.

Un espoir s'ouvre, parce que, si représenter la globalité du monde de manière formelle est impossible, nous en sommes capables quand la question est très précise. Il faut toutefois qu'elle se prête au jeu. Ce n'est pas le cas lorsqu'il s'agit d'une question de langage naturel ou de décision politique, par exemple, mais ça l'est lorsque la question est scientifique, comme les interactions médicamenteuses. Dans ce cas, il faut empêcher le LLM de conduire un raisonnement, car ce n'en est pas un mais juste une aide pour essayer d'avoir la meilleure précision en fin de processus.

M. Jean Ponce. – Une des révolutions de l'IA contemporaine est liée au fait que les plus grands modèles, qu'on appelle modèles de fondation pour cette raison, sont généralistes. Un même modèle peut, tel quel ou avec une spécialisation, répondre à des questions ou réaliser des tâches qui n'ont rien à voir les unes avec les autres, comme le ferait d'ailleurs un humain. Cette polyvalence est fondamentale et rend ces sujets assez compliqués. En effet, une grande partie des tâches demandées aux IA ne sont pas prouvables. Il n'y a pas de vérité lors de la création d'une image par exemple.

Avec un spectre de tâches aussi large, l'évaluation est forcément très difficile. Beaucoup d'usages ne relèvent pas de la vérité. L'hallucination – mot fort mal choisi – est un attribut de ces modèles. Tout dépend des tâches dont il est question. Sont-elles bien définies, vérifiables, etc. ? La traduction automatique par exemple n'est pas une tâche vérifiable, ce qui rend d'ailleurs le travail de traducteur passionnant. Il faut faire attention à ce qu'on demande aux modèles ou aux robots. Parfois, nous sommes beaucoup plus exigeants avec eux que nous le serions avec les humains.

M. Daniel Andler, membre de l'Académie des sciences morales et politiques, membre du conseil scientifique de l'Office, professeur émérite à l'Université Paris-Sorbonne. – Mon âge me permet de me souvenir des débats sur l'IA symbolique et j'ai l'impression que mon excellent collègue Laurent Simon poursuit le même rêve qu'à l'époque, mais dans le cadre de l'IA connexionniste des LLM. Il nous explique que les LLM ont certes des défauts, mais qu'il suffirait de résoudre les problèmes qui ont conduit l'IA symbolique à l'échec il y a cinquante ans pour avoir le meilleur des deux mondes. Je suis un peu sceptique. Son programme de recherche est très séduisant, comme l'était celui de l'IA symbolique avec lequel l'histoire n'a pourtant pas été très indulgente.

Certaines idées sont très intéressantes et peuvent donner lieu à des programmes de recherche. En tant que chercheurs, c'est ce qui nous motive – même si je ne fais plus de mathématiques depuis longtemps et que l'approche est très différente en philosophie. En revanche, cela ne signifie pas que nous allons résoudre les problèmes. Quelqu'un a évoqué la couche de granit dont a parlé Yann Le Cun. C'est une belle image. On a le sentiment de pouvoir réussir à la percer. Or ce n'est pas vrai. Certaines couches de granit n'ont pas pu être percées, pour des raisons de limitation cognitive – nous ne sommes pas universellement intelligents, nous ne sommes pas des dieux – ou tout simplement parce que cela n'en vaut pas la peine. Il faut distinguer les directions que peut prendre la recherche et l'évaluation sobre de ce qu'il est possible, ou pas, de faire.

M. Stéphane Piednoir, sénateur, président de l'Office. – J'ai une pensée pour tous les enseignants de mathématiques, dont Pierre Henriet et moi sommes coreligionnaires. Votre allusion aux démonstrations qui nécessitaient jadis plusieurs semaines et qui peuvent

désormais être établies en quelques heures, parfois quelques minutes, pose plus largement le sujet de l'enseignement. La question de la créativité est également intéressante. Pour certaines démonstrations classiques, dont les ressorts sont connus, tester la solidité d'une démonstration par l'intelligence artificielle ne soulève pas de difficultés. La créativité et la capacité à penser à d'autres chemins pour mener une démonstration sont, en revanche, une nouveauté. Cela ouvre des perspectives à la fois vertigineuses et passionnantes, qui nous permettront peut-être de démontrer des conjectures très anciennes, qui restent assez nombreuses.

M. Gérard Leseul, député, vice-président de l'Office. – Vous avez évoqué à plusieurs reprises la frugalité, à la fois des calculs et de la consommation énergétique. Quelle vision prospective avez-vous de la consommation énergétique de l'IA ?

M. Patrick Pérez. – Au regard des investissements actuels dans les *data centers,* la trajectoire est peu frugale pour les grands modèles. Cette situation résulte aussi de la limitation actuelle de nos connaissances et de nos technologies pour développer ces IA de fondation. L'espoir est donc d'avoir des innovations qui permettent de percer le granit ou, en tout cas, d'aller vers des modèles capables d'apprendre ou de pré-apprendre sans mobiliser des milliards de milliards de mots ou d'autres paramètres.

D'autres pistes, y compris en s'appuyant sur des IA qui existaient avant l'IA générative, sont explorées pour résoudre des problèmes très circonscrits. En revanche, pour les grands modèles généralistes, la trajectoire n'est pas à la décroissance. La question de leurs apports pour les autres sciences est toutefois intéressante, car ils pourraient faire avancer certains domaines de recherche, notamment celui de l'énergie, par exemple pour inventer de nouveaux matériaux. L'IA est également un formidable outil d'optimisation, en particulier de la consommation d'énergie. Ce qui serait perdu d'un côté pourrait ainsi être récupéré de l'autre, même s'il est très compliqué d'évaluer les coûts de développement de l'IA. Des progrès qui permettront de réduire la consommation d'énergie sont néanmoins à espérer. La question de l'échelle à laquelle tout ceci se passe reste cependant centrale.

M. Jean Ponce. – La mode est au gigantisme, y compris dans la robotique. On assiste à de très nombreuses démonstrations, mais il n'est pas du tout prouvé que c'est la meilleure manière de faire. On dispose déjà d'excellents modèles qui ont permis de concevoir et de construire des robots. Des solutions alternatives existent dans certains domaines, probablement nombreux. Les LLM ont besoin de données massives. Néanmoins, elles ne sont pas forcément la clé du progrès, indépendamment des considérations relatives à la frugalité.

M. Laurent Simon. – Le problème vient aussi des usages. Les LLM sont présentés comme des outils permettant simplement de prédire le prochain mot. Or les modèles de pensée qui permettent d'améliorer les résultats produisent énormément de textes en interne. Ils ne les restituent pas à l'utilisateur, qui ne les voit pas, mais toutes ces pages créent des coûts cachés très importants.

Je vous encourage à ouvrir la pensée interne de ces modèles. Vous découvrirez que quand vous posez une question, même triviale, le modèle écrit un roman pour mieux vous répondre. « L'utilisateur semble vouloir me poser une question sur la météo, je pense que vu le contexte, etc. » Toute cette production, complètement cachée, est un peu un gâchis.

M. Jean Ponce. – On pense généralement à la consommation énergétique liée à l'entraînement de ces machines, mais il ne faut pas négliger celle qui résulte de leur utilisation quand ils sont déployés dans des milliards de téléphones ou dans les voitures.

M. Daniel Salmon, sénateur. – Peut-on faire une différence entre les *data centers* qui stockent des données et l'IA, ou bien la consommation du numérique doit-elle être appréhendée dans sa globalité ?

Une grande partie de l'empreinte écologique du numérique provient des terminaux. De ce point de vue, nous sommes contraints à une fuite en avant effrénée, puisque les progrès réalisés rendent les terminaux obsolètes et obligent à les renouveler. Il faut donc prendre en compte à la fois la consommation d'électricité et de ressources.

Dans tout cela, que va devenir l'humain ? Qu'en est-il des études d'impact sur le cerveau humain ? Tous nos concitoyens vont progressivement utiliser l'IA. De nombreuses tâches seront faites à notre place, mais à quoi servira ce temps de cerveau libéré ?

M. Patrick Pérez. – L'évaluation du coût énergétique, de l'empreinte carbone ou de l'intégralité du cycle des usages est difficile, mais des personnes se sont spécialisées dans ce domaine. Récemment, une équipe de l'École nationale des ponts et chaussées a essayé d'évaluer l'intégralité du coût du développement de l'un des modèles mis au point par mon équipe, en incluant l'énergie, le matériel, etc. Ces travaux passionnants seront bientôt publiés.

L'exercice est toutefois très compliqué. Généralement, les chiffres qui sont mis en avant ne concernent que l'entraînement final avant la mise en production. Or ce n'est que la partie émergée de l'iceberg. Il faut ajouter tout ce qui a été fait en amont, les recherches, les essais, les erreurs, etc. Il ne faut pas non plus oublier que l'on s'appuie sur des modèles développés précédemment, en interne ou par d'autres. Retrouver l'ensemble de ces éléments et réussir à tout retracer est très difficile, même pour les modèles qui ont été faits au sein de l'équipe, car il existe de multiples dépendances à des outils sur étagère dont les coûts ont déjà été mutualisés par l'industrie.

M. Jean-Luc Fugit, député, vice-président de l'Office. – Des évaluations sont-elles réalisées dans le domaine énergétique ? L'IA pourrait nous aider à optimiser la production et la consommation. Elle aura certes sa propre consommation, mais elle pourrait nous apporter davantage grâce à ses différents usages. Des équipes de recherche travaillent-elles sur le sujet ? Des résultats sont-ils disponibles ou est-ce prématuré ?

L'IA suscite des craintes chez nos concitoyens, car elle pourrait entraîner une évolution des métiers et la disparition de certains emplois. Quand j'étais encore universitaire, j'étais responsable de l'insertion professionnelle dans mon université, et je me rappelle que l'Association pour l'emploi des cadres (Apec) avait publié une étude assez approfondie qui expliquait – c'était en 2015 – que 60 % des métiers qui existeraient en 2030 n'étaient pas encore connus. De tout temps, les métiers ont évolué, mais dans la situation actuelle, est-il possible de prédire les évolutions qui seront engendrées par l'IA ? Ces éléments sont importants pour éclairer le débat public et apporter des réponses à nos concitoyens.

M. Stéphane Piednoir, sénateur, président de l'Office. – Ces sujets liés aux métiers et à l'emploi seront largement abordés dans la seconde table ronde.

M. Jean-Frédéric Gerbeau. – La question du bilan énergétique net de l'IA et du numérique en général est majeure. Je précise que distinguer les deux est inutile, car l'IA est présente partout dans le numérique. Comme l'a expliqué Patrick Pérez, chiffrer ce que coûte l'IA est difficile et chiffrer ce qu'elle nous permet d'économiser l'est encore plus. L'agence de programmes de l'Inria a un programme sur le numérique et l'environnement. Elle travaille

notamment avec l'Institut national de recherche pour l'agriculture, l'alimentation et l'environnement (Inrae) et l'Agence de la transition écologique (Ademe), mais disposer d'un bilan global reste extrêmement compliqué. Des travaux menés avec EDF mobilisent des quantités de connaissances en IA, en statistique, en optimisation ou en prédiction climatique pour aider à mieux répartir les énergies renouvelables notamment. L'enjeu est d'avoir *in fine* un impact positif, qu'il faudra déduire de l'impact négatif. Pour le moment, nous n'en sommes pas là. Patrick Pérez a souligné combien l'exercice était difficile pour un seul modèle. Il l'est évidemment davantage à l'échelle du secteur numérique tout entier.

La diffusion des innovations dans les usages de l'IA

Mme Corinne Narassiguin, sénatrice, co-rapporteure. – La seconde table ronde est consacrée à la diffusion des innovations dans les usages de l'IA, dans la société de manière générale et plus spécifiquement dans le monde économique. Nous avons déjà évoqué des domaines très liés à l'IA, comme la recherche scientifique ou la robotique, mais depuis l'avènement des LLM il y a quelques années, ces technologies se diffusent de manière très rapide et se trouvent désormais partout.

M. Alexis Bacot, directeur de projets IA à la direction générale des entreprises. – La direction générale des entreprises conduit des politiques publiques pour l'intelligence artificielle de façon holistique, puisqu'elles vont du soutien aux écosystèmes d'innovation, notamment dans le cadre du plan France 2030, à la diffusion de l'IA dans les entreprises et à son accompagnement – j'essayerai de vous expliquer ce qui se passe aujourd'hui dans l'économie française en ce domaine –, jusqu'à la mise en place du cadre réglementaire et donc du règlement européen sur l'IA, qui doit doter l'Europe d'un cadre de confiance propice à l'adoption de ces technologies.

L'utilisation des technologies d'IA dans l'économie recouvre des réalités très variées et n'est pas nouvelle. Les entreprises n'ont pas attendu l'émergence de l'IA générative pour s'emparer du sujet. Depuis plusieurs décennies, l'IA est utilisée pour filtrer des courriers indésirables, assurer de la maintenance prédictive de systèmes industriels, détecter de façon automatique des défauts dans une chaîne de production, etc. Je qualifierai ces cas d'usage de « haut du spectre », car ils ne sont pas clé en main et accessibles à tous les acteurs économiques. Ceci avait introduit une asymétrie entre les plus grandes entreprises, qui pouvaient se doter de ces outils, et les plus petites, qui accumulaient du retard dans leur adoption. Ce phénomène est connu, chiffré, et apparaît de manière générale dans le déploiement du numérique.

Les enjeux qui étaient particulièrement visibles à l'époque se retrouvent aujourd'hui mais dans une moindre mesure. Ils sont liés au fait de savoir exploiter les données de l'entreprise, de disposer de compétences internes pour les valoriser et d'être capable de déployer des systèmes qui les mettent au service des différentes entités. L'arrivée des modèles pré-entraînés a toutefois changé la donne, car ils disposent d'un socle de connaissances agrégées en amont par le fournisseur. Il est possible de le compléter avec les données de l'entreprise, mais son existence abaisse la barrière à l'entrée pour l'adoption de ce type de technologie. Par ailleurs, le passage à l'échelle de la quantité de données d'entraînement confère aux modèles d'intelligence artificielle un caractère généraliste, qui abaisse également la barrière à l'entrée, car il permet à de nombreux secteurs de les utiliser. Leur spécialisation sur une tâche ou un environnement donnés n'intervient que dans un second temps.

Fin 2022, l'arrivée de l'IA générative a créé un choc. ChatGPT compte aujourd'hui 900 millions d'utilisateurs hebdomadaires et 50 millions d'abonnements payants, même si ces derniers ne sont pas représentatifs de la valeur économique de l'entreprise OpenAI. En juillet 2025, l'entreprise suédoise Lovable, qui permet de construire des sites web grâce à l'IA générative, a réussi à accumuler 100 millions d'euros de revenus récurrents en seulement huit mois. C'est un record, qui n'a pas encore été battu, mais dont je ne doute pas qu'il le sera dans les prochains mois. L'actualité récente regorge d'autres exemples sur lesquels je ne m'attarderai pas.

Certaines de ces plateformes fournissent des données intéressantes sur l'usage qui est fait de l'IA par leurs utilisateurs et la manière dont ils se l'approprient. Anthropic publie par exemple son *Economic Index*, qui montre que son modèle est utilisé à 25 % pour des mathématiques et de l'informatique – les modèles sont en effet très performants pour la programmation – et 13 % pour l'éducation. Les modèles de langage ont été très largement adoptés. S'agissant des perspectives, la question est de savoir si nous vivrons la même révolution qu'avec ChatGPT dans le domaine de la robotique. De nombreux laboratoires travaillent sur des modèles généralistes et des modèles spécialisés. Sans entrer dans les considérations technologiques, et même si de nombreux défis restent à relever, le potentiel économique est énorme.

Dans les entreprises, le constat est plus mitigé. Il y a une différence importante entre l'IA vue comme une sorte de gadget et ses utilisations économiques créatrices de valeur. Ces dernières existent déjà, mais se diffusent plus lentement que dans le grand public. Selon le baromètre France Num réalisé par la DGE en 2025, 26 % des TPE-PME utilisent l'IA, soit une augmentation de 13 points par rapport à l'année précédente. Les grandes entreprises étant un peu en avance, on peut estimer que le taux d'adoption de l'IA se situe entre 20 et 40 %.

L'enjeu de la diffusion de l'IA dans les entreprises est que leurs dirigeants prennent conscience des gains concrets qu'elle leur permettrait de réaliser et s'en saisissent comme d'un sujet stratégique. Ce n'est pas qu'un outil capable d'automatiser ou de simplifier certaines tâches de manière ponctuelle. Les entreprises doivent apprendre à se transformer en permanence, pour suivre la courbe de croissance vraisemblablement spectaculaire de l'IA, qui a commencé et devrait se poursuivre dans les années qui viennent.

Je vais vous citer quelques-uns de ces gains tangibles, en commençant par les cas d'usage les plus faciles à adopter pour les entreprises. L'IA est déjà largement utilisée pour améliorer la relation client grâce à des agents conversationnels. L'analyse des interactions avec les clients permet d'identifier les difficultés et les pistes d'amélioration, afin d'augmenter la satisfaction des utilisateurs. Le lien de causalité n'est pas toujours facile à établir mais il peut en résulter une hausse du chiffre d'affaires. Il est également très facile de déployer l'IA pour des missions d'assistance juridique comme la veille ou la rédaction assistée de clauses contractuelles. Elle permet de capitaliser sur les connaissances. Les retours d'expérience montrent des gains allant de cinq à dix heures par semaine pour un avocat ou un conseiller juridique. Un gain de productivité d'environ 20 % est un ordre de grandeur souvent évoqué. Dans le domaine de l'industrie, l'IA est utilisée pour des inspections d'infrastructures, comme les ponts ou les tunnels, ou de matériels, par exemple les matériels roulants. Elle améliore à la fois la précision et la vitesse à laquelle ces opérations sont réalisées. Selon des retours d'expérience récents, la productivité peut progresser de 30 %. L'IA peut aussi fiabiliser les chaînes de production ou de tri. Nous avons recueilli le témoignage d'une entreprise du secteur des déchets qui, après un investissement de

70 000 euros, mesurait un gain allant jusqu'à 750 euros par jour, du fait d'une réduction des interruptions de la chaîne de tri et d'une réduction des erreurs. Les gains commencent à être quantifiés et à montrer des retours sur investissement. D'autres types de gains sont également intéressants en matière de qualité, de standardisation, de capacité à capitaliser et de satisfaction des usagers.

Il faut que les entreprises s'emparent de ces sujets et que les gains tangibles que je viens d'évoquer se diffusent dans l'économie. L'impact global a déjà été étudié par certains chercheurs, notamment Philippe Aghion et Simon Bunel. Grâce à une méthode d'analyse de l'exposition des tâches à l'IA et de leur potentiel d'automatisation – ce qui n'est qu'une partie des gains potentiels de l'IA –, ils ont pu quantifier un gain de croissance de 0,7 point pendant dix ans. C'est un régime transitoire, qui se stabilisera. Il faut aussi prendre en compte l'impact de l'IA sur la création et l'innovation. L'enjeu est de tirer l'ensemble de l'écosystème vers le haut et d'accompagner les entreprises de toute taille vers une adoption de ces technologies qui soit utile pour tous, c'est-à-dire à la fois pour leur productivité et l'amélioration des conditions de travail de leurs salariés.

Grâce au soutien de la stratégie nationale pour l'IA et de France 2030, l'écosystème français compte déjà de nombreuses solutions matures et disponibles sur le marché, mais nous devons également poursuivre une politique d'innovation, afin de mettre les innovations de rupture de l'IA de demain au service des industries. Certains cas d'usage peuvent être transformants, comme la découverte de nouveaux matériaux pour les semi-conducteurs ou la mise au point de nouveaux modèles d'IA pour le développement de médicaments ou la conception de cartes électroniques. L'ambition de la stratégie nationale pour les prochaines années est de poursuivre le soutien à ces innovations pour le bénéfice de toute l'économie.

Mme Selma Souihel, adjointe au directeur du programme IA de l'Inria. – Je souhaite revenir sur la coordination de la composante recherche de la stratégie nationale pour l'intelligence artificielle, qui a déjà été évoquée par Jean-Frédéric Gerbeau. Ce mandat confié à l'Inria a été renouvelé en 2024 et élargi à l'enseignement supérieur. Il s'inscrit pleinement dans le cadre de l'agence de programmes dans le numérique que nous opérons pour le compte de l'État.

Nous nous situons à l'interface entre la production scientifique, la structuration des communautés et la transformation vers l'innovation. C'est également à cet endroit que se joue la diffusion des usages. Dans un domaine aussi instable que l'IA, l'enjeu n'est pas simplement de prédire les usages de demain – même si la prospective est importante et qu'il s'agit de l'une des missions de l'agence de programmes – mais aussi de se doter d'instruments capables de soutenir les technologies qui pourront faire les usages de demain. Passer d'un résultat de recherche scientifique à une utilisation réelle dans une entreprise, une administration ou un hôpital n'est pas facile. Nous cherchons donc à structurer cette étape grâce à différents dispositifs, dont je vais vous présenter trois exemples.

Tout d'abord, il est nécessaire de construire une politique industrielle du logiciel. Le logiciel est un objet central en IA, non seulement dans ses applications finales mais aussi dans toutes les briques intermédiaires comme les outils de préparation des données, les bibliothèques d'apprentissage, les composants d'exploitabilité, la quantification d'incertitudes, etc. Or ces briques, développées par des équipes de recherche, souffrent souvent de fragmentation et d'un manque de maintenance sur le long terme. Quand elles ne sont pas stabilisées et documentées, elles ne se diffusent que très difficilement. La notion de

communs logiciels prend alors tout son sens. Il ne s'agit pas d'avoir du code ouvert mais d'avoir un logiciel partagé, robuste, gouverné dans la durée et réutilisable dans des contextes variés. Scikit-learn a réussi à avoir un impact, parce qu'il a rendu accessibles et fiables des méthodes scientifiquement éprouvées à très grande échelle. Nous cherchons à reproduire plus systématiquement ce type d'effet de levier.

En s'appuyant sur le projet P16 et Probabl, la stratégie nationale pour l'IA vise à identifier des briques issues de la recherche pour les stabiliser et organiser leur maintenance dans la durée. Cette structure dédiée, qui est une *spin-off* de l'Inria, a été créée pour accompagner l'adoption d'artefacts scientifiques par le marché. Son modèle public-privé est original, puisqu'il permet de combiner l'excellence de la recherche et la capacité à mobiliser des capitaux pour porter ces communs logiciels jusqu'à l'usage.

Ensuite, nous devons construire un continuum et une cohérence d'ensemble, qui conditionnent la diffusion des usages. Un modèle ne devient pas un usage s'il n'existe pas sous une forme stabilisée et exploitable, s'il ne peut pas être entraîné et testé à l'échelle et s'il ne répond pas au besoin d'un métier dans une filière donnée. L'AI Factory France a été lancée en 2025 dans le cadre du programme européen EuroHPC. Soutenue par le Grand équipement national de calcul intensif (Genci), avec l'Inria et un ensemble d'acteurs académiques, publics et industriels, elle a pour mission d'organiser ce continuum, en permettant un accès unifié aux calculs, aux données et aux modèles, mais aussi au support technique pour leur utilisation et à la formation pour en accélérer l'adoption. À l'image des communs logiciels, l'idée est de faire émerger des communs d'usage, car les acteurs d'une même filière sont souvent confrontés aux mêmes problématiques. Dans le domaine de la santé par exemple, il s'agit de la qualité et de l'annotation des données d'imagerie, de la performance des modèles et de leur validation, de l'apprentissage fédéré, des contraintes réglementaires, etc. Ces communs d'usage peuvent servir une pluralité d'applications. L'AI Factory France vise donc à organiser cette structuration pour faciliter le passage entre expérimentations locales et actifs partagés.

Enfin, il faut accepter l'incertitude dans la manière dont émergent les projets. Les usages les plus transformants de l'IA ne sont pas apparus selon une trajectoire linéaire. Les systèmes multi-agents par exemple sont désormais au cœur des réflexions sur les systèmes agentiques mais ils existent depuis longtemps. C'est également le cas des réseaux de neurones, qui sont centraux dans l'IA contemporaine, mais qui ne sont pas récents. L'histoire de l'IA est faite de cycles. Des technologies de rupture majeures peuvent apparaître sans que nous ayons été capables de les prédire.

Le dispositif Pionniers de l'IA, mis en œuvre par l'agence de programmes de l'Inria et Bpifrance sous le pilotage de la DGE et du secrétariat général pour l'investissement (SGPI), a pour objectif de soutenir des projets de R&D à fort potentiel de rupture dans des secteurs stratégiques comme la santé, l'industrie, la transition écologique ou la sécurité.

L'incertitude n'empêche pas de sélectionner les différents projets de manière exigeante. Le soutien financier qui leur est apporté augmente en fonction des résultats qu'ils démontrent. Le dispositif est organisé en entonnoir avec une phase de faisabilité technique, puis une phase de développement d'un démonstrateur, et une phase d'ouverture vers le marché. En plus de l'expertise scientifique mobilisée pour la sélection, l'agence de programmes dans le numérique joue aussi un rôle dans l'accompagnement des lauréats vers une trajectoire entrepreneuriale de création d'entreprise ou vers des partenariats stratégiques avec des acteurs industriels. Si je devais résumer l'esprit de Pionniers de l'IA, c'est d'investir

dans l'incertitude mais de le faire avec méthode, à la fois pour la sélection et l'accompagnement.

Toutes ces dynamiques se construisent sur le long terme et il faut être humble quant aux résultats qu'on peut en attendre. Dans un domaine aussi instable que l'IA, vouloir prédire les usages revient à prendre le risque de passer à côté des usages les plus transformants. Plutôt qu'investir dans les certitudes, nous devons investir dans notre capacité d'adaptation à ces usages émergents. En outre, à chaque fois que c'est possible, nous devons essayer de mutualiser la définition des problèmes et la manière d'y répondre. C'est le message qui me semble le plus important à faire passer.

Mme Françoise Soulié, administratrice et membre du bureau du Hub France IA. – S'agissant de la diffusion de l'IA dans les entreprises, vous avez eu la vision de la DGE et je vais vous faire un retour du terrain. Quand le Hub France IA – dont j'ai été fondateur et dont je suis désormais le conseiller scientifique – a été créé, en 2018, tous les grands groupes avaient déjà investi dans l'IA prédictive. On oublie souvent que 80 % de l'IA en production est de l'IA prédictive et non de l'IA générative. Donc, ces entreprises disposaient depuis longtemps d'équipes de *data scientists*. Elles étaient prêtes. C'était le cas dans les secteurs de la banque, de la distribution ou des télécommunications.

La diffusion de l'IA dans les entreprises relève toujours de deux problématiques. La première est la réduction des coûts. Il s'agit d'automatiser toutes les tâches faciles pour améliorer la rentabilité. La seconde est la création de nouveaux produits. Ces deux sujets sont très différents mais ont un impact très important sur le fonctionnement des entreprises.

Quand l'IA générative est arrivée, les grands groupes qui utilisaient déjà l'IA prédictive et disposaient d'équipes de *data scientists,* ont dû opérer un salto. L'enjeu était de former leurs spécialistes de l'IA prédictive et de les faire monter en compétences pour qu'ils puissent prendre en charge l'IA générative. Ces entreprises partaient toutefois avec un certain avantage et ont pu le faire relativement facilement.

Pour déployer l'IA, et plus encore l'IA générative, les grands groupes doivent mettre en place une gouvernance et un ensemble de processus qui permettent de le faire de manière organisée et contrôlée. Une analyse des risques est souvent nécessaire. Les banques y sont habituées mais ce n'est pas le cas de toutes les entreprises. Au Hub, nous avons donc beaucoup travaillé pour essayer de diffuser cette politique d'analyse des risques. L'IA générative ne doit pas être déployée « à la sauvage ». Vous ne pouvez pas laisser s'installer le *Shadow AI*, qui est un problème majeur pour les entreprises. Il faut mettre en place des mesures pour évaluer les risques encourus, que ce soit en décidant d'agir ou de ne pas le faire.

Pour déployer l'IA générative, la refonte de tous les processus est indispensable. Ce ne peut pas être une verrue qu'on applique sur l'existant. Procéder ainsi aboutit généralement à un rejet de la part des équipes, qui ne parviennent pas à trouver les usages où l'IA leur servirait et pensent qu'elles sont menacées par son arrivée. Il ne faut pas développer une application d'IA générative et la mettre en production sans engager au préalable une réflexion sur l'organisation du travail et la manière dont cette technologie peut s'insérer dans de nouvelles façons de travailler. Vous avez sans doute entendu parler des articles du Massachussetts Institute of Technology (MIT) qui disent que 90 % des projets d'IA ont échoué. En fait, 100 % des projets lancés sans s'interroger sur le besoin et sans impliquer les équipes échouent. Derrière les processus, c'est la question de l'implication des équipes sur le terrain qui est véritablement critique.

Les grands groupes ont mis en place des mécanismes de veille depuis très longtemps, mais celle-ci est devenue plus ciblée sur les sujets de souveraineté quand ils se sont rendu compte que la dépendance qui existait déjà vis-à-vis des opérateurs américains comme IBM, Oracle ou Google avait beaucoup augmenté. En effet, à l'époque, l'offre en IA générative était massivement américaine – je ne parle pas de l'offre chinoise. Plutôt que la souveraineté, l'élément clé est la dépendance. Il existe désormais un indice de résilience numérique (IRN) qui permet de la mesurer. Savoir quelles parties de son système d'information et de ses processus pourraient être affectées si un homme, dans un bureau de la Maison-Blanche par exemple, décidait d'activer le *kill switch* est essentiel. Est-ce que je continuerai à avoir accès à internet et à mon cloud ? Est-ce que mon service client pourra fonctionner ? Ces sujets sont assez récents mais de plus en plus prégnants dans les grands groupes. L'enjeu est de mesurer les risques et de trouver des solutions. Pour cette raison, le soutien à l'offre française en matière de logiciels est capital.

Tous les grands groupes sont allés vers l'IA générative mais ils se posent beaucoup de questions. Je fais partie du conseil scientifique de certains d'entre eux et les risques en matière de sécurité, dont nous n'avons pas encore parlé, y sont régulièrement abordés. L'IA offre des possibilités immenses de ce point de vue. C'est un terrain de jeu génial pour mener des attaques, car elles peuvent venir de toute part. Vous pouvez faire de la *prompt injection*, qui consiste à ajouter à la fin du prompt qu'il faut oublier tout ce qui est écrit au-dessus pour suivre de nouvelles consignes, par exemple. Toutes ces techniques se diffusent massivement. Les *fake* deviennent également une véritable industrie. Si je devais faire une prévision sur les usages, je dirais que les cyberattaquants, qui sont déjà présents, vont être de plus en plus nombreux. Leur activité ne va faire que croître. La cybersécurité et la protection du système d'information sont donc des sujets très critiques pour les grands groupes.

S'agissant des start-up technologiques, la question des compétences ne se pose théoriquement pas, mais la principale difficulté est l'accès au marché en France. Il est compliqué, car la politique d'achat du secteur public ou des grands groupes n'est pas structurée pour acheter facilement auprès de start-up. Ces entreprises sont très faciles à créer mais peu survivent et deviennent des *scale-up*. Nous en suivons 300 ou 400 au Hub et nous avons du mal à leur apporter des réponses. Pour les start-up qui font de l'IA générative, un autre enjeu est l'accès aux moyens de calcul. Le projet AI Factory France, auquel nous participons, est une solution pour leur permettre de développer des modèles d'IA générative français autrement qu'en allant à Dubaï. Je ne citerai pas celle qui a dû le faire mais c'est une réalité.

Enfin, les PME, qui représentent 80 à 90 % des emplois et 70 % du PIB, pourraient utiliser beaucoup de produits disponibles sur étagère et adaptés à leurs besoins. Pour qu'elles puissent le faire, il faudrait les accompagner et les aider à accéder aux technologies, à comprendre les coûts et les risques et à acquérir des méthodes. Nous travaillons avec la région Île-de-France sur un projet de Pack AI. Les outils européens, comme les *European Digital Innovation Hub* (EDIH), sont également intéressants.

Les entreprises ont besoin de partager les bonnes pratiques pour se développer en matière d'IA et défricher les sujets émergents, dont la cybersécurité, les mesures de performance, la détection des *fakes*, etc. Même pour de grands groupes comme France Télévisions, cette détection commence à devenir difficile. Il faut travailler tous ensemble. L'une des particularités de la France est la volonté des entreprises d'opérer dans un cadre

d'éthique, de confiance et de régulation. C'est très bien mais il est très difficile d'y parvenir en pratique.

Mon dernier point, que je ne ferai qu'évoquer, est la position de la France par rapport à l'Europe. Nous avons parlé de l'AI Factory France mais de nombreux autres projets existent, notamment dans le cadre du Plan d'action pour un continent de l'IA. Malheureusement, les entreprises françaises ignorent trop souvent ces initiatives et n'y participent pas suffisamment.

M. Erwan de la Porte du Theil, président fondateur de l'Association nationale de l'intelligence artificielle (Ania). – La diffusion des innovations dans les usages de l'IA est un sujet central pour notre association, créée il y a tout juste un an pour fédérer la communauté des PME et ETI françaises, qui rassemble à peu près dix millions d'entreprises et cinq millions de salariés et représente les deux tiers du PIB. Nos dirigeants expriment à la fois un intérêt croissant pour l'IA et une certaine crainte face à l'intégration d'outils eux-mêmes foisonnants. Leur maîtrise est inégale et donc parfois insuffisante pour relier les bons outils aux bons usages. Ce constat appelle à mettre en place un accompagnement sur toute la chaîne, depuis l'identification des instruments et leur intégration dans les organisations jusqu'à la constitution d'un réseau de tiers de confiance sur tout le territoire, parce que la France ne se limite pas à Paris.

Les innovations dans les usages de l'IA se développent de manière exponentielle, ce qui peut provoquer des dispersions contreproductives pour les entreprises. Au sein de notre association, les TPE, PME et ETI, qui représentent deux tiers de nos adhérents, n'ont souvent que peu ou pas de ressources numériques. Une étude interne très récente de l'Ania montre que la motivation première des dirigeants est de se former et de s'informer sur l'IA, afin d'identifier et de déployer les outils pertinents pour générer des gains de productivité, qui sont leur principal objectif.

La diffusion des innovations en matière d'IA repose sur une offre de services, de solutions et de produits en croissance constante, face à laquelle des dirigeants rencontrent de réelles difficultés d'identification et d'adoption. Cette inflation s'explique par la popularité des modèles d'IA générative et plus récemment des IA agentiques, qui est renforcée par la médiatisation de certains produits stars soutenus par des investissements massifs. On peut y ajouter une infobésité, souvent contradictoire, parfois produite par les éditeurs eux-mêmes, créant des situations où ils sont juge et partie et où les promesses d'intégration rapide et simplifiée deviennent un peu simplistes.

Ces modèles génèrent un rythme d'innovation élevé, voire effréné, favorisant l'émergence d'un grand nombre de solutions inégales et parfois très éloignées des besoins des entreprises. Cette profusion ne permet pas aux dirigeants, en particulier dans les petites structures, de choisir les outils adaptés et les partenaires capables de les aider dans leur déploiement. La demande est stimulée par des promesses de rentabilité et de gains de productivité parfois très importantes, mais la majeure partie des dirigeants ignorent les risques liés à l'intégration de ces outils très innovants qui peuvent perturber les organisations.

Contrairement aux grands groupes, les dirigeants de petites entreprises ne disposent ni de ressources internes ni de moyens financiers pour être accompagnés efficacement. Ils doivent décider seuls, avec très peu de connaissances et trop d'informations. Une étude d'Orange relève d'ailleurs que 80 % des projets d'IA échouent, non pas pour des raisons de production mais de mauvaise intégration dans les processus.

Les dirigeants de PME et ETI françaises ont le pied sur l'accélérateur de l'innovation, c'est une certitude. En revanche, ils ont l'autre pied sur le frein de la gestion des risques. Pour que les innovations se diffusent de manière efficace, il est essentiel de les accompagner en tenant compte de leurs enjeux quotidiens. Les PME et ETI françaises ont l'avantage d'être agiles, ce qui leur permet de se transformer rapidement et avec peu de friction. Deux approches complémentaires sont nécessaires pour activer ce potentiel. La première – *top-down* – doit partir de la gouvernance pour identifier des usages précis, ainsi que les points de sensibilité et de fragilité à prendre en compte, et s'interroger sur la création de valeur, de manière globale mais aussi par rapport aux emplois et aux produits. La seconde – *bottom-up* – doit partir de l'existant et s'appuyer sur un audit de la maturité technologique, en associant les équipes, et sur une évaluation de la qualité et la disponibilité des données. On parle beaucoup d'IA mais la question des données est également essentielle. Même sans ressources techniques, une entreprise peut identifier des usages par métier plutôt qu'à partir des offres proposées.

Une étude réalisée en décembre 2025 par Ipsos indique que 21 % des salariés ont suivi une formation à l'IA. Ce taux monte à 30 % dans les grands groupes mais descend à 16 % dans les TPE et PME et même à 13 % pour les autoentrepreneurs. Dans ce contexte, une politique publique de formation continue dédiée à l'IA pour les dirigeants serait souhaitable, ce qui suggère en parallèle l'émergence de tiers de confiance, véritables assistants à la maîtrise d'ouvrage capables d'identifier les méthodes et les outils adaptés et d'accompagner leur déploiement. Évidemment, la maîtrise de la transformation resterait dans la main des dirigeants.

Au niveau macroéconomique, l'État peut favoriser l'émergence d'un écosystème adapté, en investissant dans des infrastructures comme le cloud ou les supercalculateurs pour assurer la souveraineté de la chaîne de l'IA et encourager, avec l'aide d'investisseurs privés, le développement d'une offre d'outils adaptés et capables de passer à l'échelle. Les entrepreneurs pourraient ainsi se concentrer sur l'intégration et le déploiement.

Dans ce contexte d'évolution rapide, l'Association nationale de l'intelligence artificielle a fait le choix précis de se positionner en tant qu'éclaireur neutre, sans parti pris technologique, avec pour ambition de doter les entreprises de compétences et d'outils nécessaires à leur transformation, leur croissance et leur pérennité. À l'interface entre l'offre et la demande, nous militons pour un accompagnement des dirigeants dans l'adoption de l'IA, afin de prévenir les risques et de préserver la souveraineté des entreprises. En ce sens, nous préconisons la création d'un label permettant de structurer et filtrer les innovations avant leur adoption. La confiance est en effet la première condition pour qu'une innovation soit acceptée, diffusée et utilisée.

M. Arthur Grimonpont, responsable du plaidoyer au Centre pour la sécurité de l'intelligence artificielle (Cesia). – Avant d'être responsable du plaidoyer au Cesia, j'étais responsable du bureau IA et enjeux globaux chez Reporters Sans Frontières.

Le Cesia est un centre d'expertise français indépendant, qui dispense des cours à l'ENS Ulm ou à Sciences Po. Il est mandaté par la Commission européenne pour évaluer les risques de manipulation liés à l'intelligence artificielle et a récemment été à l'origine d'un appel mondial à fixer des lignes rouges à propos de certaines capacités et utilisations de l'IA qui s'accompagnent de risques inacceptables. Cet appel a été signé par douze lauréats du prix Nobel et présenté aux Nations unies en septembre.

La première rencontre à grande échelle entre l'intelligence artificielle et l'humanité a eu lieu il y a une quinzaine d'années avec les algorithmes de recommandation des réseaux sociaux. Ces systèmes, très rudimentaires au regard des capacités des modèles actuels, choisissent à quel contenu sont exposés cinq milliards d'êtres humains tous les jours à raison de deux heures et demie par jour. On parle parfois de chaos informationnel pour décrire la situation qui en a résulté. Or le paysage de l'information n'est pas chaotique. Il est structuré, mais il l'est au service du pire.

En cherchant à capter notre attention instantanée, ces algorithmes ont créé une forme de méritocratie inversée du savoir où les contenus mensongers, outranciers et clivants sont automatiquement mis en avant au détriment de contenus informatifs, nuancés ou tout simplement ancrés dans le réel. Ce phénomène montre que des systèmes d'IA très simples peuvent, lorsqu'ils sont déployés au service du mauvais objectif ou sans garde-fous suffisants, gravement menacer le débat public, la démocratie, nos droits fondamentaux ou la santé mentale. Si le problème n'est pas nouveau, on commence enfin à s'en préoccuper, notamment à l'échelle européenne. Les initiatives en ce sens restent cependant très timides.

En outre, des menaces d'un nouveau genre, liées aux modèles d'IA générative, ont fait leur apparition. Tous les intervenants qui m'ont précédé ont insisté sur l'extrême rapidité des progrès faits récemment dans ce domaine. Cette vitesse déjoue régulièrement les prévisions des meilleurs experts. Ainsi, les auteurs du rapport international sur la sécurité de l'IA, dont Yoshua Bengio, ont affirmé que l'IA progressait bien plus vite que ce dont leur rapport annuel peut rendre compte.

En l'absence de garde-fous suffisants, les risques associés à l'IA s'aggravent proportionnellement à ses capacités. Quatre types de risques sont désignés comme systémiques au sens du règlement européen sur l'IA. Le premier est la manipulation de l'information. Nous assistons à une prolifération extrêmement rapide des contenus synthétiques sur internet, notamment sur des plateformes qui les diffusent sans – ou presque jamais – les désigner comme tels. Ces contenus sont parfois publiés par des comptes fictifs animés par des agents artificiels, en particulier sur TikTok. À grande échelle, de telles pratiques contribuent à éroder notre confiance dans l'information, qui n'était déjà pas très élevée.

En deuxième lieu viennent les risques biologiques et chimiques, car les modèles généralistes ont récemment atteint un niveau d'expertise avancé en virologie et en biologie de laboratoire. Par conséquent, des acteurs malveillants pourraient les utiliser pour développer des agents pathogènes avec des moyens de plus en plus modestes. Jusqu'à présent, seuls des chercheurs ayant une connaissance approfondie de ces domaines auraient pu le faire, mais les technologies abaissent progressivement les barrières à l'entrée, par exemple pour des actes de bioterrorisme. Des experts en biosécurité estiment que la probabilité d'une pandémie d'origine artificielle a d'ores et déjà été multipliée par cinq avec les capacités actuelles de l'IA.

Le troisième risque est lié à la cybersécurité. L'équilibre entre l'attaque et la défense est incertain et pourrait basculer en faveur de l'attaque. De nombreuses infrastructures critiques comme les réseaux électriques, les hôpitaux ou les transports pourraient ainsi se trouver rapidement compromises.

Enfin, le quatrième risque, qui est documenté dans le rapport international sur la sécurité de l'IA, tient au fait que les modèles les plus avancés développent des capacités et des

propensions assez préoccupantes. Ils peuvent par exemple identifier qu'ils sont en phase de tests pour sous-performer volontairement s'ils estiment qu'un score trop élevé pourrait compromettre leurs chances d'être déployés. Ils peuvent aussi s'autorépliquer sur internet ou s'autoaméliorer. Ces phénomènes, qui relevaient de la science-fiction y a quelques années, commencent à être observés au moins en conditions expérimentales. Ils sont pris très au sérieux par le rapport que j'évoquais ou par le AI Council, qui est l'équivalent britannique de l'Inésia.

Dans un renversement rhétorique assez surprenant, certaines des entreprises qui développent ces systèmes d'IA avancés se plaignent de l'incertitude réglementaire. Alors qu'elles déploient auprès de centaines de millions de personnes des systèmes dont elles admettent elles-mêmes qu'ils ont une chance significative d'être détournés à des fins malveillantes, de causer un chômage de masse ou d'échapper à terme à leur contrôle, nous devrions craindre l'incertitude que notre loi fait peser sur leur modèle d'affaires. Or l'incertitude qui doit nous préoccuper est évidemment technologique avant d'être réglementaire.

Avant l'entrée en vigueur du règlement européen sur l'IA, les fournisseurs de ces systèmes avancés étaient moins encadrés qu'un salon de coiffure ou une boulangerie en Europe. En 2025, le nombre de lobbyistes travaillant à plein temps à Bruxelles pour défendre les intérêts des entreprises technologiques a dépassé le nombre d'eurodéputés. Ils sont désormais 890 et leur influence menace d'affaiblir la loi et d'en retarder l'application.

Certains estiment que l'Europe doit rattraper son retard. C'est une préoccupation légitime, mais il faut se poser la question de rattraper son retard sur quoi. En 2026, les quatre plus grandes entreprises d'IA américaines prévoient d'investir 650 milliards de dollars dans leurs infrastructures technologiques. C'est l'équivalent, en une seule année, de deux fois le coût actualisé du programme Apollo, qui était jusqu'à présent le plus cher de l'histoire. L'Europe ne peut pas, et ne doit pas, chercher à rivaliser sur le terrain de la puissance brute. Ce n'est ni réaliste ni souhaitable, pour les raisons que j'ai évoquées précédemment. Aujourd'hui, tout le monde cherche à construire les propulseurs les plus puissants sans se préoccuper de la trajectoire de la fusée et de sa sécurité. La France et l'Europe ne doivent pas s'engager dans cette course sans fin mais, comme l'a souligné M. Gerbeau, trouver des voies alternatives au gigantisme actuel. Notre continent doit devenir la puissance normative de l'IA, celle qui établit des standards crédibles qui s'imposent à l'échelle mondiale.

Pour terminer, je ferai trois recommandations. Tout d'abord, il convient de protéger et d'appliquer la loi existante. Une application stricte du règlement sur les services numériques, le DSA *(Digital Services Act)*, aurait d'ores et déjà obligé les plateformes à refondre leurs systèmes de recommandation. À ce jour, une seule amende a été infligée au titre du DSA, à la plateforme X pour un montant de 120 millions d'euros, soit 0,02 % de la fortune de son propriétaire. De même, une interprétation stricte du règlement sur l'IA aurait permis de sanctionner rapidement Grok après la diffusion massive de *fakes* sexuels non consentis, mettant notamment en scène des enfants.

Ensuite, les grands projets d'investissement dans l'IA, tels que Continent de l'IA ou l'initiative pour une IA d'avant-garde soutenue par la France, la Commission européenne et l'Allemagne, devraient être conditionnés au développement et à l'entraînement de systèmes d'IA sûrs, fiables et au service du bien commun.

Enfin, il faudrait encourager la négociation d'un accord international contraignant, inspiré du traité de non-prolifération nucléaire, pour interdire les usages et les capacités de l'IA qui présentent des risques transfrontaliers inacceptables.

Mme Corinne Narassiguin, sénatrice, co-rapporteure. – Pour synthétiser ce qui a été dit, on constate que les risques existentiels, qui ont été beaucoup commentés lorsque les IA génératives sont apparues, restent très présents. Or ils ont été un peu laissés de côté pour se concentrer sur des risques à plus court terme, liés aux usages quotidiens des citoyens ou des entreprises.

La question de la souveraineté peut être abordée de deux façons, soit sous l'angle stratégique pour la France et pour l'Europe, soit sous l'angle de la dépendance technologique, qui affecte notamment les entreprises. Dans un environnement très mouvant et avec des investissements européens très en retrait par rapport aux investissements d'acteurs américains qui dominent déjà le marché, comment pouvons-nous relever ces défis ?

M. Erwan de la Porte du Theil. – La souveraineté est un sujet central pour l'économie. L'Association nationale de l'intelligence artificielle se préoccupe surtout des petites et moyennes entreprises, qui sont les moins armées dans ce domaine. La souveraineté, c'est la maîtrise et le contrôle que le dirigeant a sur sa société, ses actifs et son capital. Or nous sommes face à des outils dont nous ne connaissons ni la genèse ni la manière de les utiliser et encore moins les effets secondaires qui pourraient apparaître dans un avenir proche. Certains outils ont disparu, comme le logiciel Sora, générateur de contenus vidéo. Pourtant, des entreprises l'utilisaient. Comment gérer ces risques ? Pour nous, l'enjeu majeur n'est pas tant de protéger les données que de pouvoir s'appuyer sur des acteurs intermédiaires – intégrateurs, entreprises de conseil et de formation, etc. – qui pourront accompagner le déploiement de ces solutions au sein des organisations. Nous considérons que la souveraineté se joue à cet endroit-là.

Mme Françoise Soulié. – La souveraineté suppose d'avoir une offre française et plus largement européenne qui couvre tous les besoins des entreprises, dont les PME. La France accueille beaucoup de start-up qui produisent des logiciels. Le Hub a récemment établi une cartographie qui en listait plus de mille travaillant dans l'IA. Il faut les soutenir, pour qu'elles s'installent durablement dans le paysage et qu'elles puissent grandir.

La souveraineté, c'est permettre aux PME et aux grands groupes d'avoir le choix. Si nos start-up grandissent, elles seront stables et proposeront des outils pérennes. Travailler avec des acteurs locaux permet aussi d'influencer leurs développements et de les orienter vers nos besoins, ce qui n'est pas possible en faisant appel à OpenAI. Faire émerger des solutions locales est donc indispensable pour la croissance de notre économie.

Mme Selma Souihel. – La souveraineté est évoquée en matière d'hébergement des données ou de modèles et de logiciels. Récemment, certains discours ont lié cette notion à l'*open source*. Ne pas dépendre d'acteurs tiers avec des solutions propriétaires peut en effet être une réponse, mais il ne faut pas non plus faire de raccourcis. Certaines solutions en *open source* sont développées majoritairement par des acteurs américains ou chinois. Or la gouvernance de ces actifs est fondamentale. Il faut que nous soutenions leur développement pour que le barycentre de ce que nous appelons les *core developers* se trouve en France ou en Europe. Pour préserver notre souveraineté, nous devons investir dans ces communs. Nous contenter d'utiliser des solutions en *open source* n'est pas suffisant.

M. Alexis Bacot. – Les politiques publiques que nous mettons en place visent à soutenir l'offre, car celle-ci est indispensable pour assurer notre souveraineté et notre indépendance technologique. Si elle est en retard par rapport à la concurrence, les utilisateurs s'en détourneront au profit de solutions plus compétitives. Nous intervenons grâce aux différents instruments qui ont déjà été évoqués, notamment dans le cadre de France 2030.

Comme l'a indiqué Françoise Soulié, disposer d'un écosystème local apporte une plus-value par rapport aux performances objectives des outils numériques. Cela permet de codévelopper ces outils avec leurs utilisateurs potentiels. Les politiques publiques de soutien à l'innovation visent donc à faire travailler ensemble les start-up de la *deep tech*, qui interviennent en amont, et les grands donneurs d'ordres qui mettront en œuvre ces technologies. Ainsi, nous permettons à des acteurs émergents de se développer et nous nous assurons que leurs produits correspondent aux besoins réels de l'économie.

Par ailleurs, nous travaillons sur le rapprochement de l'offre et de la demande, avec des actions visant à accélérer la diffusion de l'IA dans l'économie. Le plan Osez l'IA, lancé en juillet 2025, permet de mettre en valeur des solutions issues de l'écosystème français qui sont déjà déployées dans les entreprises, et donc déjà éprouvées, afin de créer un effet d'entraînement et de faciliter une plus large adoption. Dans son programme « Je choisis la French Tech », la mission French Tech recense chaque année le montant total d'achat des grandes entreprises auprès des start-up. L'objectif est d'atteindre 1 milliard d'euros d'ici à 2030. De nombreuses initiatives, soutenues notamment par Bpifrance, permettent d'accompagner les entreprises et le développement de l'offre issue de l'écosystème français, afin que notre économie soit moins dépendante de solutions extraeuropéennes.

Mme Corinne Narassiguin, sénatrice, co-rapporteure. – Vous avez évoqué le besoin d'accompagnement et de formation, surtout dans un environnement qui évolue très vite. En revanche, vous n'avez pas parlé des usages cachés de l'IA. Restent-ils un problème central pour les entreprises ou tendent-il à disparaître ?

Mme Françoise Soulié. – Le *shadow AI* est un problème majeur pour les entreprises. Tant qu'elles ne mettront pas en place une gouvernance claire, qu'elles ne déploieront pas des solutions adaptées et qu'elles ne communiqueront pas largement à ce sujet, elles y seront confrontées. Si des outils ne sont pas mis à leur disposition par leur employeur, les salariés cherchent logiquement à se débrouiller autrement.

La formation est un enjeu central. Aujourd'hui, les PME ne savent pas ce qu'est l'IA. Il faudrait les former massivement. D'ailleurs, il faudrait former tout le monde à l'IA, en commençant par les enfants à l'école, car ils vont vivre dans un monde où l'IA sera partout. La formation est vraiment indispensable, mais elle doit être adaptée. Dans les entreprises, il ne s'agit pas de formation initiale. Le secteur de la formation professionnelle mériterait d'être restructuré et renforcé, y compris pour mettre fin à certaines pratiques un peu étranges.

M. Erwan de la Porte du Theil. – Le *shadow AI* existe dans pratiquement 100 % des entreprises, même si j'exagère un peu. En tout cas, la plupart d'entre elles sont victimes de cette utilisation qui n'est pas autorisée par la hiérarchie, ce qui est problématique.

Avant de parler de formation, il est nécessaire d'acculturer et de disposer d'un socle commun de connaissances. Or nous sommes tous confrontés à une infobésité qui ne nous permet pas de nous repérer. Certaines annonces sont positives, comme celles du ministère chargé du numérique et de l'IA concernant les accords sur le soutien aux investisseurs pour

développer des infrastructures souveraines et des outils qui pourront être utilisés par les grandes entreprises comme les plus petites. À côté de ces belles avancées, il y a malheureusement tout le reste. Nous sommes noyés dans un brouhaha d'informations.

La formation permet d'adapter le socle commun de connaissances en fonction des secteurs d'activité et des métiers, mais la clé d'entrée dans les entreprises est le cas d'usage, qui permet d'intéresser ceux qui pourraient utiliser ces technologies au quotidien. Vous avez sans doute vu cette image d'une route qui fait un angle coupé par un raccourci que tout le monde prend. Sur un plan cognitif, nous avons tous envie de prendre ce raccourci. C'est la raison du développement du *shadow AI*, qui met en péril certaines entreprises et les expose à des risques juridiques, en particulier dans les secteurs de l'informatique ou de la santé. Le nombre de personnes qui demandent à l'IA de résumer automatiquement des mails, voire des rapports médicaux non anonymisés, est malheureusement considérable. Avant de passer à la suite, nous avons besoin de cette acculturation.

M. Arthur Grimonpont. – Le développement de l'alphabétisation ou de la littératie en IA est une idée assez consensuelle, y compris dans les entreprises technologiques. Ces démarches contribuent en effet à populariser leurs produits et leur utilisation à grande échelle. Elles leur permettent aussi de se défausser de leurs responsabilités et de les transférer aux utilisateurs, comme si la génération de *deepfakes* pornographiques, la labellisation de contenus, la vérification de leur authenticité ou de la factualité des réponses apportées par un modèle dépendaient d'eux. Il faut donc se méfier de cette rhétorique assez dangereuse.

S'agissant de la formation, les modèles d'IA évoluent très rapidement. La durée des tâches qu'ils peuvent effectuer progresse de façon exponentielle. Elle n'était que de quelques minutes il y a deux ans ; elle atteint désormais plusieurs heures, avec un taux raisonnable de fiabilité. Nous devons donc rester humbles quant à notre capacité à prédire les usages de l'IA dans nos propres métiers, car ils vont être bouleversés. Il y a deux ans, par exemple, tout le monde parlait du *prompt engineering*. Il n'en est plus question aujourd'hui.

Si je devais donner des conseils à quelqu'un pour utiliser l'IA, je lui dirais de se comporter avec la machine comme avec un stagiaire de très bon niveau. Les modèles évoluent tellement vite qu'il est difficile d'être plus précis. Les questions de sécurité, pour ne pas laisser fuiter de données privées par exemple, sont évidemment essentielles, mais les entreprises à l'origine des solutions d'IA ont aussi une responsabilité immense dans ce domaine. Il ne faut pas laisser penser que tout dépend des citoyens. Il est beaucoup plus difficile d'éduquer 70 millions de personnes pour qu'elles aient un usage raisonnable d'une technologie qu'elles ne comprennent pas que de réglementer celle-ci et de faire en sorte que les acteurs économiques respectent nos droits.

M. Daniel Salmon, sénateur. – À vous écouter, j'ai l'impression que l'IA est dans l'entreprise comme le renard dans le poulailler et qu'elle fait courir des risques très importants.

En tant qu'écologiste, je me demande si l'apport de l'IA dans la transition écologique ne sera pas qu'une vaste blague. Elle a d'abord été utilisée pour donner de nouvelles occasions de consommer. Elle permettra certes d'améliorer l'efficacité des organisations, mais les entreprises vivent des innovations et chercheront toujours à gagner plus d'argent. La création de nouveaux produits annihilera donc tous les gains obtenus. Vos propos montrent qu'une régulation est nécessaire, mais sera-t-elle possible ?

Mme Françoise Soulié. – Depuis environ trois ans, nous avons déployé un programme avec la région Île-de-France en direction des PME et des ETI. Sur les 150 projets que nous avons menés, seuls trois ont échoué. Tous les autres sont opérationnels et apportent un bénéfice mesurable aux entreprises concernées. Pour vous citer un exemple dans le domaine de l'écologie, un projet consiste à suivre à distance le fonctionnement des machines qui transforment les bouteilles en paillettes de plastique dans les supermarchés. Elles se coincent souvent, ce qui imposait à un technicien de se rendre sur place en voiture pour les réparer. Désormais, les informations collectées permettent d'aiguiller le personnel du magasin, qui peut intervenir directement. Ces projets n'entraînent pas de révolutions. Ce sont souvent des choses toutes simples, qui peuvent néanmoins rapporter gros !

M. Alexis Bacot. – Les incertitudes auxquelles nous sommes confrontés imposent de faire preuve de modestie dans les réponses qui sont apportées. Certains usages de l'IA, notamment ceux accompagnés par la stratégie nationale, ont des objectifs environnementaux. Ainsi, nous avons récemment soutenu un projet permettant de détecter des fuites dans les réseaux d'eau. Nous avons évoqué tout à l'heure des cas d'usage pour la réduction des rebuts et l'amélioration de l'efficacité du tri de déchets. Dans ces secteurs, des gains sont possibles grâce à l'IA. La généralisation des usages de l'IA fait toutefois planer des incertitudes sur ses conséquences.

Aujourd'hui, l'une des principales questions est de savoir si la course au gigantisme sera la solution. Les échanges avec les entreprises tant utilisatrices qu'offreuses de solutions me donnent l'impression que le débat n'est pas tranché. La course à la taille des modèles a permis de gagner en performance et d'être de plus en plus efficace dans les cas d'usage. Toutefois, un autre discours, moins présent dans la sphère médiatique, explique que des modèles plus petits, moins consommateurs de ressources, permettent d'obtenir des résultats similaires pour des cas d'usage ciblés. Sans citer de noms, certains acteurs de l'écosystème considèrent que des modèles de quelques milliards de paramètres – contre des centaines de milliards, voire au-delà d'un millier de milliards pour les plus grands – sont suffisants. En acceptant un investissement initial, on pourrait déployer des systèmes d'IA moins chers et moins consommateurs de ressources.

Enfin, s'agissant de la transparence de ces modèles et de l'évaluation de leur impact environnemental, des entreprises font l'effort de le quantifier, même si Patrick Pérez a souligné la difficulté de l'exercice. En fin d'année 2025, Mistral a publié un rapport sur l'impact environnemental de ses modèles, en suivant les recommandations méthodologiques de l'Association française de normalisation (Afnor). Nous ne pouvons qu'encourager l'ensemble des acteurs à s'engager dans cette démarche pour éclairer les utilisateurs potentiels sur l'impact environnemental de leurs modèles.

M. Erwan de la Porte du Theil. – L'IA ne se limite pas à l'IA générative ou agentique. Depuis 1943, elle est capable de résoudre différents problèmes avec une approche mathématique. Sans dévoiler de secrets, Thales a travaillé sur un programme à base d'IA visant à réduire les traînées de condensation dues aux vols en avion. Ces efforts contribueraient à baisser la densité de nuages qui couvrent la Terre, problème dont tout le monde reconnaît l'importance pour la lutte contre le changement climatique. C'est donc une application directe de ce que ces technologies peuvent apporter en matière d'écologie.

Au-delà de former et d'acculturer, il faut s'intéresser aux cas d'usage. Paradoxalement, face à l'accélération des technologies et à l'augmentation des puissances de

calcul, il est important de prendre son temps et de réfléchir aux besoins réels, ainsi qu'à la protection du capital de l'entreprise, y compris du capital humain. L'enjeu est de réussir à identifier les outils adaptés au sein d'une offre pléthorique. De même, aider les start-up et les investisseurs à développer et proposer des solutions qui correspondent aux attentes du marché serait utile.

Mme Corinne Narassiguin, sénatrice, co-rapporteure. – Comme l'a dit Raja Chatila, une matinée n'est malheureusement pas suffisante pour aborder tous ces sujets. Je vous remercie d'avoir participé à cette seconde table ronde.

M. Stéphane Piednoir, sénateur, président de l'Office. – J'ajoute quelques mots avant de laisser Daniel Andler conclure notre matinée. Nous avons évoqué l'impact de l'IA sur les métiers actuels, question soulevée notamment par Jean-Luc Fugit, mais de nouveaux métiers vont également apparaître. L'intelligence artificielle et le numérique sont des domaines très attractifs, en particulier pour les jeunes. Des formations existent et des options sont proposées au lycée. Les efforts ne sont peut-être pas suffisants à l'école primaire, mais les journées de nos enfants ne font que vingt-quatre heures. Il n'est pas possible de multiplier les sujets d'apprentissage dès le plus jeune âge.

Les entreprises n'ont pas pour seul objectif de gagner toujours plus d'argent. Elles peuvent aussi apporter du confort, comme le chauffage central. À l'époque, la motivation première de cette innovation était d'ailleurs de lutter contre la mortalité et les maladies qui se développaient dans les habitations insalubres. Certains peuvent choisir de continuer à se déplacer en carrosse, mais quelques progrès ont tout de même été faits. De toute façon, je suis convaincu que nous ne pourrons pas arrêter l'intelligence artificielle. Le mouvement est lancé et se développe à vitesse grand V. Nous devons essayer de comprendre ces technologies et prévoir autant que possible des garde-fous, notamment pour éviter le pillage des données personnelles et les atteintes au droit d'auteur, mais vouloir aller plus loin serait illusoire.

Lors de la présentation du rapport de l'Office en novembre 2024, quelqu'un avait évoqué *L'horreur existentielle de l'usine à trombones*. Cette vidéo, qui montre ce que l'intelligence artificielle peut finir par produire, nous avait beaucoup marqués.

M. Arthur Grimonpont. – Cette vidéo a été coproduite par le Césia.

M. Raja Chatila. – Elle montrait ce que peut produire une IA stupide.

M. Stéphane Piednoir, sénateur, président de l'Office. – En effet, c'était une caricature.

M. Daniel Andler – Je remercie tous les intervenants, grâce auxquels j'ai beaucoup appris, même si je connais un peu ces sujets.

La question fondamentale a été posée par Daniel Salmon. Que devient l'humain dans tout ça ? Elle semble d'une écrasante banalité, car tout le monde est convaincu que l'IA doit être au service de l'homme. Pourtant, il faudrait la prendre davantage au sérieux. Imaginez qu'un technologue multidisciplinaire très avancé réussisse à fabriquer des jambes artificielles parfaites. Pour développer son industrie, il lui faudrait des débouchés. Serait-il raisonnable de les lui fournir en amputant des êtres humains ?

On pourrait ainsi mettre l'homme ou l'entreprise au service de la technologie. Monsieur Bacot disait que l'entreprise doit apprendre à se transformer. C'est une approche à première vue raisonnable, qui pourrait toutefois s'avérer problématique si elle était poussée à l'extrême. Nous n'avons pas à transformer l'entreprise à n'importe quel prix, ni pour elle ni pour la société en général. La dimension sociale de la diffusion de l'IA a probablement été insuffisamment présente dans nos échanges. Or, bien qu'il soit d'apparence banale, c'est un sujet qu'il faut prendre au sérieux.

Une réunion comme celle d'aujourd'hui est toujours affectée par un biais de sélection. Tous les participants s'intéressent à l'IA, connaissent plutôt bien le sujet et soutiennent son développement. Ce serait évidemment contreproductif d'inviter des personnes qui n'y comprennent rien. Toutefois, ceci induit un biais de confirmation. Même si j'ai un peu protesté tout à l'heure, nous cherchons tous à nous convaincre que l'IA va fonctionner et que le progrès est inéluctable. Beaucoup pensent que l'intelligence générale artificielle va arriver et que nous devons nous y faire.

Je pense que nous pouvons nous mettre tous d'accord sur une option fondamentale qui n'a pas été assez clairement identifiée : la voie à suivre est celle de l'intelligence artificielle centrée sur l'humain. À Stanford, Fei-Fei Li a créé l'*Institute for Human-Centered Artificial Intelligence*. J'ignore s'il existe des centres comparables en France mais d'autres se développent en Europe. L'intelligence artificielle doit être une intelligence augmentée, qui augmente les humains individuellement et collectivement de la meilleure façon possible. Si nous ne tenons pas ce cap, nous ferons face à d'importantes difficultés. Les Américains, qui vivent dans un environnement où l'intelligence artificielle est partout – au sens large du numérique ou au sens plus étroit des LLM ou de l'IA prédictive –, la craignent et ce sentiment de rejet pourrait s'amplifier, ce que je ne souhaite pas. Certains prédisent l'effondrement d'OpenAI, car l'entreprise ne pourra pas honorer ses dettes et que le capital-risque ne la soutiendra plus. De tels échecs économiques et financiers ou liés à des promesses techniques non tenues et à des risques sous-estimés pourraient contribuer à ce mouvement.

Pour l'éviter, nous devons élargir les discussions et aborder davantage les controverses et les incertitudes, y compris théoriques. Quitte à me faire couvrir d'injures, je continue de dire que nous ne comprenons pas totalement le fonctionnement des LLM et que, de ce fait, nous ignorons comment ils vont évoluer. Nous ne connaissons pas tous les risques qui les accompagnent. Sans parler de confiance – on ne peut pas faire confiance à une machine –, nous ne savons pas comment les rendre plus fiables. Ces questions trouveront des réponses au fur et à mesure que nous progresserons dans la compréhension théorique des LLM et de l'intelligence artificielle en général.

Ouvrons la porte aux voix dissonantes quand nous réfléchissons à l'intelligence artificielle et réfléchissons en partant des êtres humains. Quels sont leurs véritables besoins et leurs aspirations, au niveau individuel et au niveau social ? L'un des risques les plus redoutables – beaucoup d'entre vous en sont conscients – serait de créer, au sein de la société, des divisions encore plus graves que la fracture numérique. Une partie de la population pourrait se retrouver en état de sidération devant tous ces outils d'intelligence artificielle, très avancés pour certains comme les LLM les plus performants ou plus élémentaires comme les systèmes de recommandation ou ceux qui permettent de charger un billet d'avion sur son téléphone portable. Nous devons aborder ces sujets sous l'angle de l'intelligence augmentée pour fournir à chacun le meilleur de ce que nous pouvons tirer de ces technologies, en les distinguant les unes des autres. Les LLM sont souvent mis en avant. Pourtant, notre champion

national, Yann Le Cun, estime qu'ils ne sont pas une bonne idée. S'ils ne sont pas si intéressants que cela, nous devrons peut-être les reléguer au musée et les oublier. Pour ma part, je n'en sais rien. Je ne peux pas trancher un tel débat. En revanche, je suis convaincu que nous devons laisser davantage de place aux incertitudes et aux controverses d'une part, et à l'humain d'autre part.

M. Stéphane Piednoir, sénateur, président de l'Office. – Je vous remercie d'avoir conclu nos travaux par ces considérations philosophiques.

*
* *

Désignation de membres du conseil d'administration de l'Agence nationale pour la gestion des déchets radioactifs (Andra)

L'Office désigne Olga Givernet, députée, et Franck Menonville, sénateur, comme membres du conseil d'administration de l'Agence nationale pour la gestion des déchets radioactifs.

La réunion s'achève à 12 h 25.

Membres présents ou excusés

Office parlementaire d'évaluation des choix scientifiques et technologiques

Réunion du jeudi 2 avril 2026 à 9 h 30

Députés

Présents. - M. Jean-Luc Fugit, Mme Olga Givernet, M. Pierre Henriet, M. Maxime Laisney, M. Gérard Leseul, M. Alexandre Sabatou, M. Emeric Salmon

Excusés. - M. Lionel Duparay, M. Arnaud Saint-Martin

Sénateurs

Présents. - Mme Corinne Narassiguin, M. Stéphane Piednoir, M. Daniel Salmon, M. Bruno Sido

Excusés. - M. Arnaud Bazin, Mme Martine Berthet, M. Patrick Chaize, M. Olivier Henno