

L'intelligence artificielle dans la science : Défis, opportunités et avenir de la recherche (Synthèse)

Le rapport complet est disponible en anglais :

OCDE (2023), *Artificial Intelligence in Science: Challenges, Opportunities and the Future of Research*, Éditions OCDE, Paris, <https://doi.org/10.1787/a8d820bd-en>.

L'accroissement de la productivité dans la recherche scientifique pourrait s'avérer le plus important de tous les bénéfices liés au recours à l'intelligence artificielle (IA), tant sur le plan économique que social. Bien que l'IA intervienne, désormais, dans toutes les disciplines scientifiques, et ce, à tous les stades, son potentiel est loin d'être pleinement exploité. Les pouvoirs publics et le monde de la recherche ont encore beaucoup à faire pour développer le recours à l'IA au service de la science et maximiser ses contributions aux progrès de la recherche. Avancer dans ce sens soutiendra le développement des pays de l'OCDE, renforcera leur capacité à innover et leur permettra de mieux répondre aux défis mondiaux, qu'il s'agisse du changement climatique ou de nouvelles pandémies.

Des programmes multidisciplinaires ambitieux pour favoriser le progrès

Pour surmonter les difficultés posées par le recours à l'IA, il est nécessaire de mettre en place de vastes programmes pluridisciplinaires réunissant des scientifiques de toutes spécialités, des informaticiens, des ingénieurs, des statisticiens, des mathématiciens, et tous les autres acteurs susceptibles d'apporter leur contribution. Pour ce faire, un financement public dédié est nécessaire, et doit être alloué dans un cadre favorisant une collaboration globale, plutôt qu'un financement cloisonné entre les différentes disciplines. **L'interaction entre les roboticiens et les experts des différents domaines doit devenir une priorité.** En effet, en permettant de réduire les coûts et d'accélérer considérablement le rythme des expérimentations, les robots utilisés en laboratoire pourraient révolutionner plusieurs domaines de la science.

Les gouvernements peuvent encourager et soutenir les initiatives tournées vers l'avenir et ayant un impact à long terme. À l'instar du Nobel Turing Challenge, qui vise à concevoir des systèmes autonomes capables d'effectuer des recherches de niveau mondial, les initiatives de ce genre dans le secteur scientifique encouragent à travailler ensemble. Or, la collaboration et la coordination sont essentielles pour pouvoir concentrer les efforts sur les défis mondiaux et favoriser un accord sur les normes relatives à l'IA, mais aussi pour susciter l'intérêt des jeunes scientifiques pour des projets de cette envergure.

En vue de faire progresser l'intelligence artificielle et la science, il est également important d'**améliorer l'accès du monde de la recherche au calcul de haute performance (CHP) et à ses logiciels.** La mise à disposition de ressources informatiques par les grandes entreprises technologiques est certes utile, mais elle comporte d'importantes lacunes, et les groupes de recherche moins bien financés risquent d'être distancés.

Espérer que les chercheurs universitaires soient performants en utilisant des ressources informatiques de pointe, fournis par des entreprises de services cloud à but lucratif suppose des coûts bien trop élevés, dans la plupart des cas, pour être réaliste. **Les laboratoires nationaux et leurs infrastructures informatiques, en collaboration avec le secteur privé et les universités, pourraient combler ces lacunes et contribuer à l'élaboration de matériel de formation pour les établissements d'enseignement supérieur.** Les pays les plus avancés dans ce domaine, notamment les États-Unis et certains pays de l'Union européenne, pourraient participer à l'élaboration des politiques visant à mettre toutes les ressources à disposition sur un espace commun.

Parallèlement, la mise à jour des programmes d'études pourrait être utile. **Par exemple, en utilisant des techniques d'IA déjà éprouvées, on pourrait apprendre aux étudiants à explorer la littérature scientifique existante à la recherche de nouvelles hypothèses.** Le cursus biomédical standard ne prévoit aucune formation de ce type. Cette mise à jour pourrait s'accompagner de nouveaux programmes de recherche industrielle et de formations doctorales interdisciplinaires, qui seraient basés sur la consolidation des connaissances et assistés par l'IA.

En outre, les pouvoirs publics peuvent prendre des mesures visant à **améliorer la disponibilité des données de la recherche ouverte et à mieux exploiter la puissance des données dans différents domaines, de la santé au climat.** L'espace européen des données de santé et GAIA-X, qui vise à construire une infrastructure de données fédérées pour l'Europe, en sont deux exemples concrets. Par ailleurs, il convient de **soutenir les centres de recherche en les aidant à adopter des systèmes d'apprentissage, tels que l'apprentissage fédéré, qui permettent d'appliquer l'intelligence artificielle à des données sensibles détenues par plusieurs parties, tout en garantissant la protection de la vie privée.** Enfin, les gouvernements doivent améliorer l'interopérabilité des instruments de laboratoire grâce à des interfaces standardisées, et pourraient, pour ce faire, **réunir les utilisateurs et fournisseurs des instruments de laboratoire avec ceux qui en développent les technologies.**

La R&D publique pour faire progresser l'IA

La recherche publique (R&D publique) peut cibler les domaines spécifiques pour lesquels des percées sont nécessaires pour optimiser les applications de l'IA, tant sur le plan théorique que pratique. Il faudrait, notamment, commencer par dépasser les modèles actuels, basés sur de grands ensembles de données et sur le CHP, et trouver des moyens d'automatiser, à grande échelle, la création de données dites « FAIR », c'est-à-dire faciles à trouver, accessibles, interopérables et réutilisables. En outre, on pourrait faire progresser AutoML (automatisation de la conception de modèles d'apprentissage automatique) pour contribuer à remédier à la rareté et au coût élevé de l'expertise existante en matière d'IA. On peut envisager, par exemple, le lancement de défis autour d'AutoML dans le monde de la recherche scientifique, et le financement de recherches sur l'application d'AutoML à la science pilotée par l'IA.

Le développement de plateformes libres, telles que OpenML et DynaBench, permettant de déterminer les modèles d'IA les plus performants pour un large éventail de problèmes, devrait aussi être encouragé. Le soutien des pouvoirs publics est fondamental pour faciliter l'utilisation simultanée de ces plateformes par un grand nombre de disciplines scientifiques.

La R&D publique devrait servir à favoriser une manière de penser novatrice, interdisciplinaire et ambitieuse. Par exemple, le traitement du langage naturel (TLN) peut aider à gérer le volume considérable de la littérature scientifique disponible de nos jours, même s'il est vrai que les performances actuelles du TLN sont exagérées. Les recherches existantes en matière de TLN offrent un cadre peu favorable à la formation d'idées osées et à haut risque, indispensables à toute découverte scientifique. On pourrait envisager des centres de recherche, des flux de financement et des systèmes de publication scientifique destinés à récompenser les nouvelles méthodes, même si celles-ci n'en sont qu'à leurs balbutiements.

Les bases de connaissances organisent le savoir mondial en cartographiant les liens entre différents concepts, grâce à des informations provenant de multiples sources. **Ce besoin ne pouvant pas être satisfait par le secteur privé, il est de la responsabilité des pouvoirs publics de soutenir un vaste programme visant à créer des bases de connaissances essentielles à l'IA dans le domaine scientifique.** Cela pourrait se traduire par la création d'un réseau de connaissances ouvert, qui servirait de ressource commune à l'ensemble de la communauté de la recherche pilotée par l'IA. Des fonds publics relativement modestes seraient suffisants pour réunir des experts en IA, des scientifiques issus de toutes les disciplines et des organisations professionnelles, y compris de bénévoles, dans le but de poser les fondements nécessaires à la mutualisation des connaissances pratiques et professionnelles.

La diversité des thèmes de la recherche en IA semble être en déclin et de plus en plus axée sur les approches à forte intensité de calcul et de données qui dominent dans les grandes entreprises technologiques. **Le renforcement du rôle des pouvoirs publics dans la R&D pourrait rétablir cette diversité et contribuer à accroître le vivier de talents.** Cela permettrait aux bailleurs de fonds d'accorder une attention particulière aux projets qui explorent des techniques nouvelles et des méthodes distinctes du paradigme dominant de l'apprentissage profond. **Dans le même temps, les décideurs politiques pourraient encourager le milieu de la recherche à examiner et quantifier les pertes engendrées par l'appauvrissement de la recherche en IA, en matière de résilience technologique, de créativité et d'inclusivité, et à s'intéresser aux possibles conséquences de la domination croissante du secteur économique.**

Une grande partie de l'IA dans la science nécessite une collaboration entre personnes ; cependant, les bailleurs de fonds pourraient aussi aider au développement d'outils destinés à améliorer la collaboration entre les équipes humaines et l'IA, et intégrer ces outils dans la science traditionnelle. La taille croissante des équipes et la dimension internationale des consortiums dans le domaine de la science impose, désormais, de combiner l'intelligence collective humaine à l'artificielle ; or, l'investissement dans ce domaine de recherche a pris du retard par rapport à d'autres thèmes.

Entre autres, des progrès sont nécessaires dans l'application de l'apprentissage automatique à l'imagerie médicale. Les échecs constatés lors de la COVID-19 ont été considérables. Comme pour d'autres utilisations de l'apprentissage automatique en science, des mesures incitatives sont nécessaires pour encourager la recherche sur des méthodes plus largement validées, dont le financement devrait être assorti de pratiques d'évaluation plus rigoureuses.

Questions relatives à la gouvernance de la recherche

Les organismes politiques devraient évaluer systématiquement les répercussions de l'IA sur la pratique scientifique au quotidien, à commencer par les interactions entre l'homme et l'IA, le travail, les parcours professionnels et la formation, pour lesquels d'importants changements sont à prévoir. Les appels de fonds pourraient **exiger de telles évaluations, et les bailleurs de fonds et décideurs politiques devraient mettre en place des procédures de suivi permettant d'agir sur la base des informations recueillies.** Entre autres mesures, ils pourraient créer et soutenir de nouveaux forums indépendants assurant une discussion permanente sur la nature changeante du travail scientifique, et ses incidences sur la productivité et la culture de la recherche.

Le déploiement de grands modèles de langage (« LLM », en anglais), tels que ChatGPT, doit se faire sous la supervision des pouvoirs publics, car les conséquences sont actuellement encore méconnues. Les LLM pourraient entraîner une augmentation du travail superficiel en le facilitant, brouiller les concepts de paternité et de propriété des idées et inventions, et même, à terme, engendrer de nouvelles inégalités entre les personnes dont la langue maternelle bénéficie d'investissements élevés en matière de recherche, et les citoyens parlant des langues à ressources faibles. D'un autre côté, les LLM et autres formes d'IA pourraient également être utilisés pour participer aux processus de gouvernance, par exemple

en facilitant l'évaluation par les pairs, une piste qui nécessite d'être étudiée davantage et soumise à des tests plus approfondis.

Enfin, **des politiques devraient être élaborées pour prévenir des dangers de double usage en lien avec la découverte de médicaments grâce à l'IA.** Peu d'attention a été accordée aux risques imminents liés à la capacité d'automatiser la conception, les essais et la production de molécules extrêmement mortelles. Il existe, d'ailleurs, d'autres risques de double usage à prendre en compte. Les pouvoirs politiques et le monde de la recherche doivent examiner et décider quel dispositif de gouvernance protégera le plus efficacement l'intérêt général.

Un besoin croissant de formation des décideurs politiques pour les aider à choisir les initiatives à soutenir

Les réseaux sociaux et plateformes existants pourraient servir à la diffusion des nouvelles pratiques. Les plateformes communautaires, telles que la communauté Loop ou Academia.edu, pourraient être utilisées comme des laboratoires d'expérimentation pour la collaboration entre l'homme et l'IA, destinée à la découverte de connaissances et à la génération et synthèse d'idées, et les réseaux sociaux permettraient la diffusion d'approches, telles que les découvertes basées sur la recherche documentaire.

Des mesures sont aussi nécessaires pour améliorer la reproductibilité de la recherche en IA. Les organismes publics de financement pourraient, notamment, exiger que les codes, les données et les métadonnées soient partagés librement avec des tiers, afin de permettre à ces derniers d'effectuer des expériences sur leur propre matériel.

Pour des raisons évidentes, l'Afrique subsaharienne, comme d'autres régions en développement, devrait bénéficier d'un financement dédié à l'application de l'IA dans le domaine scientifique, beaucoup plus important. Les projets de coopération au développement pourraient aider les pays concernés à faire progresser le champ de la science ouverte, à encadrer la législation sur la protection des données, à améliorer les infrastructures numériques, à renforcer les moyens nécessaires à l'application de l'IA, et à soutenir les initiatives émergentes de l'Afrique, à commencer par le développement local de données, de logiciels et de technologies. Les projets relatifs à l'IA au service de la science menés avec les pays en développement peuvent se révéler mutuellement bénéfiques ; des modèles de financement à faible coût ont, d'ailleurs, déjà fait leurs preuves. Enfin, la coopération au développement peut contribuer à créer et à encourager des centres de recherche d'excellence.



Extrait de :

Artificial Intelligence in Science

Challenges, Opportunities and the Future of Research

Accéder à cette publication :

<https://doi.org/10.1787/a8d820bd-en>

Merci de citer ce chapitre comme suit :

OCDE (2023), « Synthèse », dans *Artificial Intelligence in Science : Challenges, Opportunities and the Future of Research*, Éditions OCDE, Paris.

DOI: <https://doi.org/10.1787/238594f3-fr>

Ce document, ainsi que les données et cartes qu'il peut comprendre, sont sans préjudice du statut de tout territoire, de la souveraineté s'exerçant sur ce dernier, du tracé des frontières et limites internationales, et du nom de tout territoire, ville ou région. Des extraits de publications sont susceptibles de faire l'objet d'avertissements supplémentaires, qui sont inclus dans la version complète de la publication, disponible sous le lien fourni à cet effet.

L'utilisation de ce contenu, qu'il soit numérique ou imprimé, est régie par les conditions d'utilisation suivantes :

<http://www.oecd.org/fr/conditionsdutilisation>.