

Résumé

L'intelligence artificielle (IA) et la robotique constituent des technologies révolutionnaires majeures qui transforment l'économie et la société. Pour comprendre et anticiper ces transformations, les décideurs politiques doivent d'abord comprendre ce que ces technologies sont capables de faire. L'OCDE a lancé le projet *L'IA et les compétences de demain* afin de concevoir un programme qui permette d'évaluer les capacités de l'IA et de la robotique et leurs répercussions dans les domaines de l'éducation et du travail. Ce rapport constitue la première étape dans le développement de l'approche méthodologique du projet. Il examine les taxonomies et les tests existants dans le domaine de la psychologie et de l'informatique et évalue leurs points forts, leurs faiblesses, et la possibilité de les utiliser pour évaluer les capacités des machines.

L'évaluation des capacités de l'IA et de la robotique offre une base indispensable pour comprendre leurs répercussions dans le domaine de l'éducation, du travail et pour l'ensemble de la société.

Un programme continu d'évaluation de l'IA et de la robotique viendra enrichir d'un nouvel élément fondamental l'éventail de mesures comparatives internationales de l'OCDE qui permettent aux responsables de l'action publique de comprendre les compétences humaines. Le Programme international pour le suivi des acquis des élèves (PISA) s'intéresse au lien qui existe entre les systèmes éducatifs et le développement des compétences humaines, tandis que le Programme pour l'évaluation internationale des compétences des adultes (PIAAC) examine la relation entre ces compétences et le monde du travail et d'autres tâches importantes réalisées par les adultes. Un programme conçu pour évaluer les capacités de l'IA et de la robotique mettra en relation les compétences humaines et ces technologies essentielles, créant ainsi un pont entre l'IA et la robotique, leurs implications pour le domaine de l'éducation et du travail, et les transformations sociales qui en découleront dans les prochaines décennies.

Il existe un certain nombre de taxonomies et d'outils d'évaluation des compétences humaines. Ils permettent de comprendre les capacités de l'IA selon des points de vue et des contextes différents.

Les taxonomies issues des travaux sur la psychologie cognitive sont des modèles hiérarchiques des capacités cognitives générales, comme l'intelligence fluide, la mémoire, l'apprentissage, la perception visuelle et auditive, évaluées par des analyses factorielles des tests de capacités cognitives. Ces tests ont été largement utilisés et validés pour évaluer les compétences humaines.

L'intérêt de la recherche scientifique pour les compétences sociales et émotionnelles et leurs évaluations est en constante progression. Ces compétences portent essentiellement sur la personnalité, le tempérament, les attitudes, l'intégrité et les interactions personnelles des individus. Depuis peu, la recherche scientifique se penche non seulement sur les capacités individuelles mais aussi collectives. Elle

étudie les différents facteurs de l'intelligence collective et développe les évaluations qui permettent de les mesurer.

La recherche dans le domaine de l'éducation a également contribué à définir et façonner la compréhension des compétences humaines. Elle s'intéresse aux connaissances dans des matières spécifiques comme les mathématiques, la biologie ou l'histoire, aux compétences fondamentales comme la lecture et le calcul, et aux compétences transversales plus complexes comme la résolution de problèmes, la collaboration, la créativité et les compétences numériques et globales. Un large éventail de tests est disponible grâce aux évaluations pédagogiques à grande échelle, aussi bien internationales que nationales.

Les compétences peuvent être mises en lien avec les tâches et les emplois professionnels, et mesurées au moyen de tests professionnels complexes.

Un autre domaine important - la psychologie du travail et des organisations - met en lien les compétences à des tâches spécifiques de certaines professions. Les taxonomies relatives aux professions qui en résultent classifient les professions selon les tâches demandées et les compétences, connaissances et aptitudes requises pour les réaliser. Les classifications les plus fréquemment utilisées sont la base de données du ministère du Travail des États-Unis, l'Occupational Network (O*NET), et la classification européenne des aptitudes, compétences certifications et professions (ESCO). Les évaluations dans ce domaine comprennent un ensemble de tests sur les professions et les emplois.

Les adultes en bonne santé partagent certaines compétences de base que les systèmes d'IA ne présentent pas.

De nombreuses taxonomies utilisées pour l'évaluation des compétences négligent les compétences cognitives de faible niveau ou élémentaires, pourtant omniprésentes. Celles-ci sont rarement évaluées chez les adultes car, en l'absence de handicap majeur, il existe peu de différences significatives entre individus. Pourtant, les systèmes d'IA ne sont pas nécessairement dotés de ce genre de compétences (comme par exemple évoluer dans un environnement physique complexe, comprendre un langage de base ou connaître les lois fondamentales du monde). Les taxonomies et évaluations de ce type de compétences se retrouvent plutôt dans les domaines de l'intelligence animale, du développement de l'enfant et de la neuropsychologie. Un domaine apparu depuis peu évalue les compétences de base (de faible niveau) des systèmes d'IA en s'inspirant de ces branches de la psychologie.

L'évaluation des systèmes d'IA et robotiques est difficile et l'utilisation de tests initialement destinés aux humains peut être trompeuse.

L'évaluation de l'IA met l'accent sur les composants fonctionnels des systèmes intelligents, comme la représentation des connaissances, le raisonnement, la perception, la navigation et le traitement automatique du langage naturel. Ces composants sont en général associés à la technique sous-jacente d'un système d'IA et nombre d'entre eux peuvent être superposés avec des catégories de capacités humaines développées en psychologie. Cependant, cette superposition n'est pas toujours exacte, et de nombreuses capacités que l'IA est en train de développer, telles que l'identification du langage et la génération d'images réalistes, ne semblent pas bien couvertes par les taxonomies ou tests des compétences humaines.

De plus, la conception des tests humains part du principe que les participants aux tests partagent tous les caractéristiques de base de l'intelligence humaine, qui peuvent être radicalement différentes de celles de l'IA. Par exemple, l'intégration de compétences fondamentales, comme la compréhension du langage et

la reconnaissance des objets, est aisée pour les humains. Toutefois, la plupart des systèmes d'IA sont formés pour réaliser une tâche spécifique et restreinte, mais ils ne sont pas (ou rarement) capables d'intégrer cet apprentissage et de s'en servir pour réaliser un autre type de tâche. Ceci rend difficile de déduire, à partir des performances d'un système d'IA à un test d'aptitude spécifique destiné aux humains, une compétence sous-jacente de l'IA, et encore moins d'en déduire une intelligence générale.

Différents types d'évaluations empiriques peuvent mesurer les capacités de l'IA mais elles sont dispersées et non systématiques.

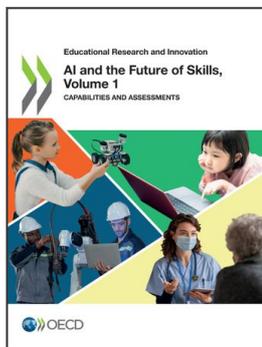
Une multitude d'évaluations et de compétitions évaluent et comparent les systèmes d'IA de façon empirique, mais celles-ci n'ont pas encore été classées systématiquement. De plus en plus d'organisations réalisent des campagnes d'évaluation rigoureuses pour mesurer les capacités des systèmes d'IA et robotiques. Elles comprennent l'évaluation de fonctions isolées, comme par exemple l'autolocalisation. Elles réalisent également l'évaluation de tâches complètes constituant une activité concrète, comme le pilotage autonome et la synthèse de texte. L'évaluation des systèmes d'IA est particulièrement bien développée dans certains domaines, comme la compréhension du langage ou encore la traduction automatique, ce dernier constituant notamment un domaine riche d'enseignements pour l'évaluation de l'IA.

Une évaluation systématique de l'IA nécessite l'adoption d'un cadre complet qui couvre toutes les compétences humaines nécessaires dans le domaine professionnel et dans la vie quotidienne.

Fournir des mesures valides, fiables et concrètes des capacités de l'IA et de la robotique nécessite l'adoption d'une approche globale qui rassemble différentes traditions de la recherche et des méthodologies complémentaires. L'objectif serait de couvrir l'ensemble des capacités humaines pertinentes; les capacités complémentaires nécessaires à prendre en considération pour l'IA (car elles sont difficiles pour l'IA et souvent négligées dans les listes des compétences humaines); et l'éventail complet des tâches valorisées dans l'éducation, le travail et la vie au quotidien.

Une méthodologie solide implique de comprendre comment l'IA et la robotique sont évaluées et de rassembler différentes méthodes d'évaluation.

Une approche pluridisciplinaire a besoin d'un fondement théorique qui tienne compte des difficultés posées par l'évaluation des capacités de l'IA et de la robotique relative aux compétences humaines. Les différentes approches disciplinaires peuvent être réparties en deux dimensions. Une première dimension consiste à déterminer si les taxonomies et les tests des compétences mesurent *principalement les capacités associées aux humains ou principalement celles associées à l'IA*. L'autre dimension consiste à déterminer si les approches mesurent des *capacités uniques (isolées) ou alors des tâches complexes* qui font appel à des capacités multiples. La future évaluation systématique des capacités de l'IA devrait associer différentes évaluations relatives à ces deux dimensions et tirer habilement parti de leur potentiel à déterminer efficacement les implications pour le travail et l'éducation de demain.



Extrait de :
AI and the Future of Skills, Volume 1
Capabilities and Assessments

Accéder à cette publication :
<https://doi.org/10.1787/5ee71f34-en>

Merci de citer ce chapitre comme suit :

OCDE (2022), « Résumé », dans *AI and the Future of Skills, Volume 1 : Capabilities and Assessments*, Éditions OCDE, Paris.

DOI: <https://doi.org/10.1787/39c4481a-fr>

Cet ouvrage est publié sous la responsabilité du Secrétaire général de l'OCDE. Les opinions et les arguments exprimés ici ne reflètent pas nécessairement les vues officielles des pays membres de l'OCDE.

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>.