

Intelligence artificielle

L'**intelligence artificielle** (**IA**) est la capacité des machines à effectuer des tâches typiquement associées à l'intelligence humaine, comme l'apprentissage, le raisonnement, la résolution de problème, la perception ou la prise de décision. L'intelligence artificielle est également le champ de recherche visant à développer de telles machines ainsi que les systèmes informatiques qui en résultent.

Souvent classée dans le domaine des mathématiques et des sciences cognitives, l'IA fait appel à des disciplines telles que la neurobiologie computationnelle (qui a notamment inspiré les réseaux neuronaux artificiels), les statistiques, ou l'algèbre linéaire. Elle vise à résoudre des problèmes à forte complexité logique ou algorithmique. Par extension, dans le langage courant, l'IA inclut les dispositifs imitant ou remplaçant l'homme dans certaines mises en œuvre de ses fonctions cognitives¹.

Les applications de l'IA comprennent notamment les moteurs de recherche, les systèmes de recommandation, l'aide au diagnostic médical, la compréhension du langage naturel, les voitures autonomes, les *chatbots*, les outils de génération d'images, les outils de prise de décision automatisée, les programmes compétitifs dans des jeux de stratégie et certains personnages non-joueurs de jeu vidéo².

Depuis l'apparition du concept, les finalités, les enjeux et le développement de l'IA suscitent de nombreuses interprétations, fantasmes ou inquiétudes, que l'on retrouve dans les récits ou films de science-fiction, dans les essais philosophiques³ ainsi que parmi des économistes.

Définition

Le terme « intelligence artificielle », souvent abrégé par le sigle « IA » (ou « AI » en anglais, pour *artificial intelligence*) a été créé par John McCarthy, qui l'a défini comme « la science et l'ingénierie de la fabrication de machines intelligentes, en particulier de programmes informatiques intelligents. Elle est liée à la tâche similaire qui consiste à utiliser des ordinateurs pour comprendre l'intelligence humaine, mais l'IA ne doit pas se limiter aux méthodes qui sont biologiquement observables »⁴.

Pour Marvin Lee Minsky, l'un de ses créateurs, l'IA est « la construction de programmes informatiques qui s'adonnent à des tâches qui sont, pour l'instant, accomplies de façon plus satisfaisante par des êtres humains car elles demandent des processus mentaux de haut niveau tels que : l'apprentissage perceptuel, l'organisation de la mémoire et le raisonnement critique »^{a,5}. Cette définition combine l'aspect « artificiel » des ordinateurs et des processus informatiques, aux aspects « intelligents » d'imitation de comportements humains, notamment de raisonnement et d'apprentissage. Celui-ci est à l'œuvre dans les jeux, dans la pratique des mathématiques, dans la compréhension du langage naturel, dans la perception visuelle (interprétation des images et des scènes), auditive (compréhension du langage parlé) ou par d'autres capteurs, dans la commande d'un robot dans un milieu inconnu ou hostile.

Avant les années 2000, d'autres définitions sont proches de celle de Minsky, mais varient sur deux points fondamentaux⁶ :

- les définitions qui lient l'IA à un aspect *humain* de l'intelligence et celles qui la lient à un modèle idéal d'intelligence, non forcément humaine, nommée *rationnalité* ;
- les définitions qui insistent sur le fait que l'IA a pour but d'avoir *toutes les apparences* de l'intelligence (humaine ou rationnelle), et celles qui insistent sur le fait que le *fonctionnement interne* du système d'IA doit ressembler également à celui de l'être humain et être au moins aussi rationnel.

Intelligence artificielle



Partie de	<u>Informatique</u> , nouvelles technologies, <u>raisonnement</u>
Pratiqué par	Chercheur ou chercheuse en intelligence artificielle (d), <u>ingénieur en intelligence artificielle</u>
Champs	<u>Intelligence artificielle symbolique</u> <u>connexionnisme</u> ethical artificial intelligence (d) <u>apprentissage automatique</u>
Histoire	<u>Histoire de l'intelligence artificielle</u>

Le grand public confond souvent l'intelligence artificielle avec l'apprentissage automatique (*machine learning*) et l'apprentissage profond (*deep learning*). Ces trois notions diffèrent et sont en réalité imbriquées : l'intelligence artificielle englobe l'apprentissage automatique, qui lui-même englobe l'apprentissage profond⁷.

Les définitions font souvent intervenir^{8,9} :

- une capacité à percevoir l'environnement et à prendre en compte la complexité du monde réel ;
- un traitement de l'information (collecter et interpréter des intrants, captés sous forme de données) ;
- des prises de décision (y compris dans le raisonnement et l'apprentissage), choix d'actions, exécution de tâches (dont d'adaptation, de réaction aux changements de contexte...), avec un certain niveau d'autonomie ;
- la réalisation d'objectifs spécifiques (raison ultime des systèmes d'IA).

Le groupe AI Watch note que les IA peuvent aussi être classées en fonction des familles d'algorithmes et/ou des modèles théoriques qui les sous-tendent, des capacités cognitives reproduites par l'IA, des fonctions exécutées par l'IA. Les applications de l'IA peuvent, elles, être classées en fonction du secteur socioéconomique et/ou des fonctions qu'elles y remplissent⁸.

Une manière de définir l'intelligence artificielle est de considérer ses applications et les types de tâches qu'elle résout. Un rapport de la Commission européenne publié en 2020 présente une taxonomie classant les définitions de l'IA selon diverses tâches réalisées, telles que le raisonnement, l'apprentissage, la perception, etc.¹⁰. La même année, le professeur Jack Copeland propose une définition similaire, qui permet de distinguer plus clairement les facettes de l'IA, selon cinq catégories principales¹⁰ :

1. L'apprentissage généralisé : l'IA apprend à partir de données diverses pour identifier des modèles et appliquer ces connaissances à de nouvelles situations, comme détecter des fraudes en ligne ;
2. Le raisonnement : cette capacité permet à l'IA de faire des prédictions et tirer des conclusions à partir des données, aidant dans des décisions comme la prédiction de comportements d'achat ;
3. La résolution de problèmes : l'IA trouve des solutions optimales pour des problèmes spécifiques, utilisée dans des contextes comme l'optimisation industrielle ou les stratégies de jeu ;
4. La perception : elle permet à l'IA de reconnaître et interagir avec son environnement, utilisée dans la robotique avancée et les véhicules autonomes pour naviguer et accomplir des tâches ;
5. La compréhension du langage : l'IA analyse et génère du langage à travers le NLP, utilisé dans des applications comme les assistants vocaux et les *chatbots* pour améliorer l'interaction utilisateur.

Techniques

Comportements prédéfinis (imitation d'intelligence)

Pour les intelligences artificielles servant principalement à donner une impression d'intelligence dans un cadre contrôlé, notamment pour les personnages non-joueurs des jeux vidéo, il est courant que l'apprentissage automatique ne soit pas utilisé. Un ensemble de fonctions et comportements plus précises et moins flexibles sont alors implémentés.

Il s'agit souvent de liste de textes ou paroles prédéfinis, aux déclenchements parfois conditionnels, par exemple un choix de mouvements suivant une série de règles et des déplacements de *pathfinding*.

Apprentissage automatique

L'apprentissage automatique consiste à permettre au modèle d'IA d'apprendre à effectuer une tâche au lieu de spécifier exactement comment il doit l'accomplir¹¹. Le modèle contient des paramètres dont les valeurs sont ajustées tout au long de l'apprentissage. La méthode de la rétropropagation du gradient est capable de détecter, pour chaque paramètre, dans quelle mesure il a contribué à une bonne réponse ou à une erreur du modèle, et peut l'ajuster en conséquence. L'apprentissage automatique nécessite un moyen d'évaluer la qualité des réponses fournies par le modèle¹². Les principales méthodes d'apprentissage sont :

Apprentissage supervisé

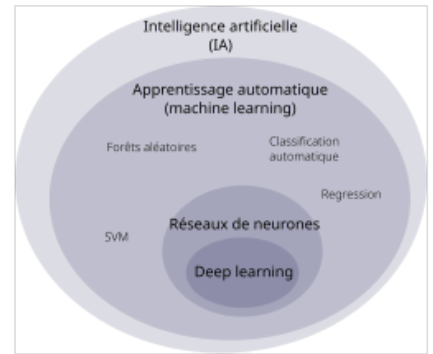


Diagramme de Venn montrant comment s'imbriquent les notions d'intelligence artificielle, d'apprentissage automatique et d'apprentissage profond.

Un jeu de données annoté est utilisé pour entraîner l'algorithme. Il contient des données d'entrée fournies au modèle et les réponses correspondantes attendues, que le modèle est entraîné à produire¹¹. Il est parfois difficile de se procurer suffisamment de données annotées avec les réponses attendues¹³.

Apprentissage non supervisé

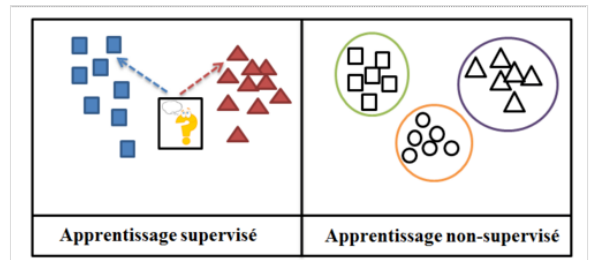
Un jeu de données est fourni au modèle, mais n'est pas annoté avec les réponses attendues. Le but peut par exemple être de regrouper les données similaires entre elles¹¹ (*clustering*).

Apprentissage auto-supervisé

Un problème d'apprentissage supervisé est *automatiquement* généré à partir d'un jeu de données non annoté. Cela fonctionne souvent en cachant une partie des informations (des mots d'un texte, des morceaux d'images...) afin d'entraîner le modèle à les prédire¹⁴.

Apprentissage par renforcement

L'agent est plongé dans un environnement où ce qu'il fait est évalué. Par exemple, un agent peut apprendre à jouer aux échecs en jouant contre lui-même, et le résultat (victoire ou défaite) permet à chaque itération d'évaluer s'il a bien joué. Il n'y a dans ce cas pas besoin de jeu de données¹¹.



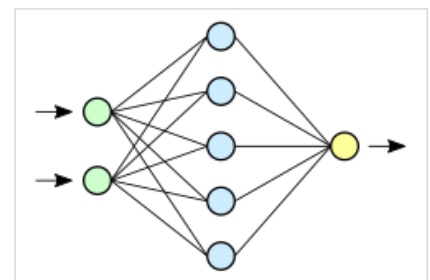
En apprentissage supervisé, les réponses sont connues, tandis qu'en apprentissage non supervisé, l'algorithme découvre des structures dans les données par lui-même¹¹.

Réseaux de neurones

Les réseaux de neurones artificiels sont inspirés du fonctionnement du cerveau humain : les neurones sont en général connectés à d'autres neurones en entrée et en sortie. Les neurones d'entrée, lorsqu'ils sont activés, agissent comme s'ils participaient à un vote pondéré pour déterminer si un neurone intermédiaire doit être activé et ainsi transmettre un signal vers les neurones de sortie. En pratique, pour l'équivalent artificiel, les « neurones d'entrée » ne sont que des nombres et les poids de ce « vote pondéré » sont des paramètres ajustés lors de l'apprentissage^{15,16}.

Hormis la fonction d'activation, les réseaux de neurones artificiels n'effectuent en pratique que des additions et des multiplications matricielles, ce qui fait qu'ils peuvent être accélérés par l'utilisation de processeurs graphiques¹⁷. En théorie, un réseau de neurones peut approximer n'importe quelle fonction¹⁸.

Pour de simples réseaux de neurones à propagation avant (*feedforward* en anglais), le signal ne passe que dans une direction. Avec les réseaux de neurones récurrents, le signal de sortie de chaque neurone est réinjecté en entrée de ce neurone, permettant d'implémenter un mécanisme de mémoire à court terme¹⁹. Les réseaux neuronaux convolutifs, qui sont particulièrement utilisés en traitement d'images, introduisent une notion de localité. Leurs premières couches identifient des motifs relativement basiques et locaux comme des contours, là où les dernières couches traitent de motifs plus complexes et globaux¹⁶.



Exemple de réseau de neurones comprenant deux neurones d'entrée (en vert), une couche « cachée » de neurones (en bleu) et un neurone de sortie (en jaune).

Apprentissage profond

L'apprentissage profond (*deep learning* en anglais) utilise de multiples couches de neurones entre les entrées et les sorties, d'où le terme « profond »²⁰. L'utilisation de processeurs graphiques pour accélérer les calculs et l'augmentation des données disponibles a contribué à la montée en popularité de l'apprentissage profond. Il est utilisé notamment en vision par ordinateur, en reconnaissance automatique de la parole et en traitement automatique des langues²¹ (ce qui inclut les grands modèles de langage).

Grands modèles de langage

Les grands modèles de langage sont des modèles de langage ayant des milliards de paramètres. Ils reposent très souvent sur l'architecture transformeur²².

Les transformeurs génératifs préentraînés (*Generative Pretrained Transformers* ou *GPT* en anglais) sont un type particulièrement populaire de grand modèle de langage. Leur « pré-entraînement » consiste à prédire, étant donnée une partie d'un texte, le token suivant (un *token* étant une séquence de caractères, typiquement un mot, une partie d'un mot, ou de la ponctuation). Cet entraînement à prédire ce qui va suivre, répété pour un grand nombre de textes, permet à ces modèles d'accumuler des connaissances sur le monde. Ils peuvent ensuite générer du texte semblable à celui ayant servi au pré-entraînement, en prédisant

un à un les *tokens* suivants. En général, une autre phase d'entraînement est ensuite effectuée pour rendre le modèle plus véridique, utile et inoffensif. Cette phase d'entraînement (utilisant souvent une technique appelée RLHF) permet notamment de réduire un phénomène appelé « hallucination », où le modèle génère des informations d'apparence plausible mais fausses²³.

Avant d'être fourni au modèle, le texte est découpé en *tokens*. Ceux-ci sont convertis en vecteurs qui en encodent le sens ainsi que la position dans le texte. À l'intérieur de ces modèles se trouve une alternance de réseaux de neurones et de couches d'attention. Les couches d'attention combinent les concepts entre eux, permettant de tenir compte du contexte et de saisir des relations complexes²⁴.

Ces modèles sont souvent intégrés dans des agents conversationnels, aussi appelés chatbots, où le texte généré est formaté pour répondre à l'utilisateur. Par exemple, l'agent conversationnel ChatGPT exploite les modèles GPT-3.5 et GPT-4²⁵. En 2023 font leur apparition des modèles grand public pouvant traiter simultanément différents types de données comme le texte, le son, les images et les vidéos, tel Google Gemini²⁶.

Recherche et optimisation

Certains problèmes nécessitent de chercher intelligemment parmi de nombreuses solutions possibles.

Recherche locale

La recherche locale, ou recherche par optimisation, repose sur l'optimisation mathématique pour trouver une solution numérique à un problème, en améliorant progressivement la solution choisie²⁷.

En particulier, en apprentissage automatique, la descente de gradient permet de trouver une solution localement optimale, étant donné une fonction de coût à minimiser en faisant varier les paramètres du modèle. Elle consiste, à chaque étape, à modifier les paramètres à optimiser dans la direction qui permet de réduire le mieux la fonction de coût. La solution obtenue est *localement* optimale, mais il se peut qu'il y ait globalement de meilleures solutions, qui auraient pu être obtenues avec différentes valeurs initiales de paramètres²⁷. Les modèles d'IA modernes peuvent avoir des milliards de paramètres à optimiser, et utilisent souvent des variantes plus complexes et efficaces de la descente de gradient²².

Les algorithmes évolutionnistes, inspirés de la théorie de l'évolution, utilisent une forme de recherche par optimisation. À chaque étape, des opérations telles que la « mutation » ou le « croisement » sont effectuées de manière aléatoire pour obtenir différentes variantes, et les variantes les mieux adaptées sont sélectionnées pour l'étape suivante²⁷.

Recherche dans l'espace des états

La recherche dans l'espace des états vise à trouver un état accomplissant l'objectif à travers un arbre des états possibles²⁸. Par exemple, la recherche antagoniste est utilisée pour des programmes jouant à des jeux tels que les échecs ou le go. Elle consiste à parcourir l'arbre des coups possibles par le joueur et son adversaire, à la recherche d'un coup gagnant²⁹. La simple recherche exhaustive est rarement suffisante en pratique vu le nombre d'états possibles. Des heuristiques sont utilisées pour prioriser les chemins les plus prometteurs³⁰.

Logique

La logique formelle est utilisée pour le raisonnement et la représentation des connaissances. Elle se décline en deux principales formes, la logique propositionnelle et la logique prédicative. La logique propositionnelle opère sur des affirmations qui sont vraies ou fausses, et utilise la logique connective avec des opérateurs tels que « et », « ou », « non » et « implique ». La logique prédicative étend la logique propositionnelle et peut aussi opérer sur des objets, prédicats ou relations. Elle peut utiliser des quantificateurs comme dans « *Chaque X est un Y* » ou « *Certains X sont des Y* »³¹.

L'inférence logique (ou déduction) est le processus qui consiste à fournir — à l'aide d'un moteur d'inférence — une nouvelle affirmation (la conclusion) à partir d'autres affirmations connues comme étant vraies (les prémisses). Une règle d'inférence décrit les étapes valides d'une preuve ; la plus générale est la règle de résolution. L'inférence peut être réduite à la recherche d'un

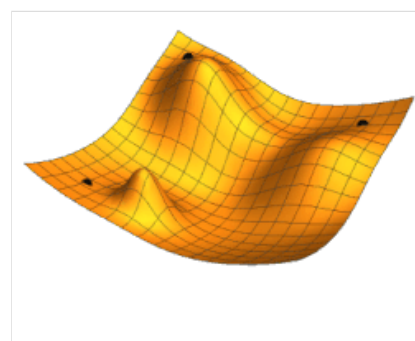


Illustration de la descente de gradient pour trois points de départ différents, faisant varier deux paramètres de sorte à minimiser la fonction de coût représentée par la hauteur.

chemin amenant des prémisses aux conclusions, où chaque étape est une application d'une règle d'inférence³¹. Mais à part pour de courtes preuves dans des domaines restreints, la recherche exhaustive prend beaucoup de temps.

La logique floue assigne des valeurs de vérité entre 0 et 1, permettant de gérer des affirmations vagues, comme « il fait chaud »³². La logique non monotone permet d'annuler certaines conclusions³¹. Diverses autres formes de logique sont développées pour décrire de nombreux domaines complexes.

Méthodes probabilistes et gestion de l'incertitude

De nombreux problèmes en IA (raisonnement, planification, apprentissage, perception, robotique, etc.) nécessitent de pouvoir opérer à partir d'informations incomplètes ou incertaines³³.

Certaines techniques reposent sur l'inférence bayésienne, qui fournit une formule pour mettre à jour des probabilités subjectives étant données de nouvelles informations. C'est notamment le cas des réseaux bayésiens. L'inférence bayésienne nécessite souvent d'être approximée pour pouvoir être calculée³⁴.

Les méthodes de Monte-Carlo sont un ensemble de techniques pour résoudre des problèmes complexes en effectuant aléatoirement de nombreuses simulations afin d'approximer la solution³⁵.

Les réseaux de neurones peuvent aussi être optimisés pour fournir des estimations probabilistes³⁶.

Des outils mathématiques précis ont été développés pour analyser comment des agents intelligents peuvent faire des choix et des plans en utilisant la théorie de la décision, la maximisation de l'espérance et la théorie de la valeur de l'information. Ces techniques comprennent des modèles tels que les processus de décision markoviens, la théorie des jeux et les mécanismes d'incitation³⁴.

Classifieurs et méthodes statistiques

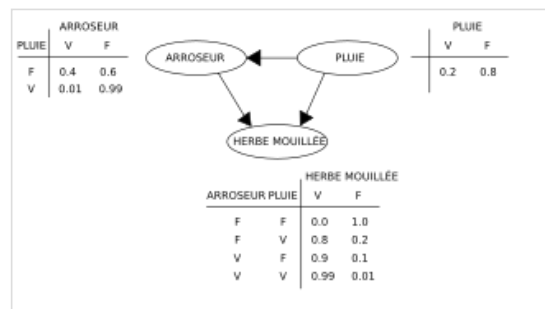
De nombreux modèles d'IA ont pour but d'assigner une catégorie (classification), une valeur (régression) ou une action à des données fournies. Les méthodes de classification comprennent arbres de décision, k plus proches voisins, machine à vecteurs de support ou classification bayésienne naïve^{37,34}. Les réseaux de neurones peuvent également faire de la classification³⁸.

Histoire

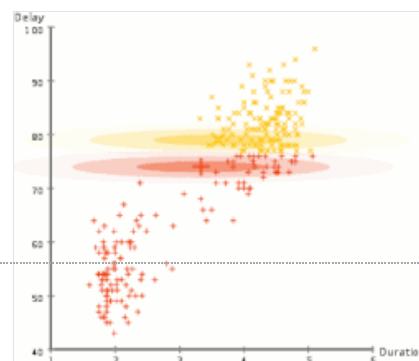
Comme précurseur à l'intelligence artificielle, divers automates ont été créés au cours de l'histoire, dont le canard de Vaucanson ou les automates d'Al-Jazari. Certains automates remontent à l'Antiquité et étaient utilisés pour des cérémonies religieuses³⁹. Des mythes et rumeurs rapportent également la création d'êtres intelligents, par exemple les golems⁴⁰.

Des philosophes et mathématiciens comme Raymond Lulle, Leibniz ou George Boole ont cherché à formaliser le raisonnement et la génération d'idées⁴¹.

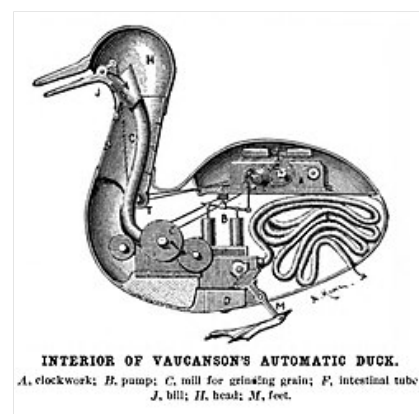
Au ^{xx}e siècle, Alan Turing a notamment inventé un modèle de calcul par la suite appelé machine de Turing, exploré la notion de calculabilité et d'intelligence des machines, et proposé le « jeu de l'imitation » (test de Turing) pour évaluer l'intelligence de futures machines⁴¹. Le terme « intelligence artificielle » a été mis en avant par John McCarthy lors de la conférence de Dartmouth en 1956, où l'intelligence artificielle a été établie en tant que discipline à part entière^{42,43}. Dans



Un exemple de réseau bayésien, et les tables de probabilité conditionnelle associées.



Séparation des données en deux groupes (partitionnement) par un algorithme de maximisation de l'espérance.



Le canard artificiel de Vaucanson (1738).

les années qui ont suivi, des chercheurs ont proposé diverses preuves de concept, dans des situations spécifiques, de ce que les machines peuvent faire en théorie. Par exemple, le programme ELIZA pouvait se faire passer pour un psychologue, et le Logic Theorist pouvait démontrer des théorèmes⁴⁴.

La fin du siècle a été marquée par des périodes d'enthousiasme, et deux périodes de désillusion et de gel des financements appelées « hivers de l'IA »⁴⁵, la première de 1974 à 1980 et la seconde de 1987 à 1993. Les systèmes experts ont été particulièrement populaires dans les années 1980, malgré leur fragilité et la difficulté à implémenter manuellement les bonnes règles d'inférences⁴⁴. Des techniques d'apprentissage automatique se sont développées (réseaux de neurones, rétropropagation du gradient, algorithmes génétiques) ainsi que l'approche connexionniste⁴⁴. Mais les faibles puissances de calcul et le manque de données d'entraînement limitaient leur efficacité. Certains domaines n'ont progressivement plus été considérés comme faisant partie de l'intelligence artificielle, à mesure qu'une solution efficace était trouvée⁴⁶ ; un phénomène parfois appelé « effet IA ». En 1997, pour la première fois, un supercalculateur a gagné plusieurs parties au jeu d'échec contre le champion du monde.

Dans les années 2000, le Web 2.0, le big data et de nouvelles infrastructures et capacités de calcul ont permis l'exploration de masses de données sans précédent. En 2005, le projet Blue Brain a débuté, ayant pour objectif de simuler le cerveau de mammifères⁴⁷. En 2012, le réseau neuronal convolutif AlexNet a lancé l'utilisation de processeurs graphiques pour entraîner des réseaux de neurones, décuplant ainsi les capacités de calcul dédiées à l'apprentissage⁴⁸. En 2016, un programme a gagné quatre des cinq parties de go jouées contre Lee Sedol, l'un des meilleurs joueurs au monde. Des organisations visant à créer une intelligence artificielle générale ont vu le jour, comme DeepMind en 2010⁴⁹ et OpenAI en 2015⁵⁰. Dès les années 2010, des outils d'intelligence artificielle (spécialisée ou généraliste) ont accompli des progrès spectaculaires, mais restent loin des performances du vivant dans beaucoup de ses aptitudes naturelles, en particulier sur son aptitude à apprendre rapidement à partir d'un faible volume d'information (par induction), selon le magazine Slate en 2019⁵¹.

En 2017, des chercheurs de Google ont proposé l'architecture transformer, qui a servi de base aux grands modèles de langage. En 2018, Yann Le Cun, Yoshua Bengio et Geoffrey Hinton ont remporté le prix Turing pour leurs travaux sur l'apprentissage profond^{52,53}.

En 2022, des programmes générant des images à partir de descriptions textuelles, comme Midjourney ou DALL-E 2, se sont popularisés⁵⁴. La même année, l'agent conversationnel ChatGPT a connu une croissance inédite, gagnant un million d'utilisateurs en seulement cinq jours⁵⁵ et cent millions d'utilisateurs en deux mois⁵⁶, ce qui a accentué un phénomène de « course » à l'IA⁵⁷. En 2023, les progrès rapides de l'IA ont suscité des inquiétudes quant à un potentiel risque d'extinction de l'humanité⁵⁸. Des modèles de fondation « multimodaux », c'est-à-dire capables de traiter simultanément plusieurs modalités (texte, images, son) ont émergé, tels que Google Gemini⁵⁹ et GPT-4o⁶⁰.



Dans les années 2010, les assistants personnels intelligents sont l'une des premières applications grand public de l'intelligence artificielle.

Intelligence artificielle générale

L'intelligence artificielle générale (IAG) comprend tout système informatique capable d'effectuer ou d'apprendre pratiquement n'importe quelle tâche cognitive propre aux humains ou autres animaux⁶¹. Elle peut alternativement être définie comme un système informatique surpassant les humains dans la plupart des tâches ayant un intérêt économique⁶².

L'intelligence artificielle générale a longtemps été considérée comme un sujet purement spéculatif⁶³. Certains travaux de recherche ont déjà décrit GPT-4 comme ayant des « étincelles » d'intelligence artificielle générale^{64,65}. Les experts en intelligence artificielle affichent de larges désaccords et incertitudes quant à la date potentielle de conception des premières intelligences artificielles générales (parfois appelées « intelligences artificielles de niveau humain »), leur impact sur la société, et leur potentiel à déclencher une « explosion d'intelligence »⁶⁶.

Un sondage de 2022 suggère que 90 % des experts en IA pensent que l'IAG a plus d'une chance sur deux d'être réalisée dans les 100 ans, autour d'une date médiane de 2061⁶⁷.

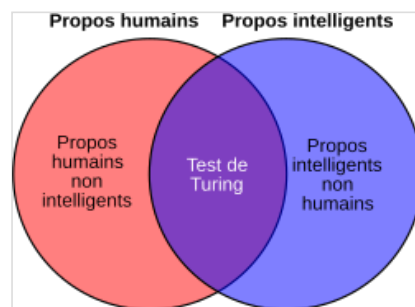
Une superintelligence artificielle est un type hypothétique d'intelligence artificielle générale dont les capacités intellectuelles dépasseraient de loin celles des humains les plus brillants⁶⁸. Le philosophe Nick Bostrom note que les machines disposent de certains avantages par rapport aux cerveaux humains, notamment en ce qui concerne la mémoire, la vitesse (la fréquence des processeurs étant de l'ordre de dix millions de fois plus élevée que celle des neurones biologiques) et la capacité à partager des connaissances⁶⁹.

Tests

Dans ce contexte, un test est un moyen d'évaluer les capacités d'une intelligence artificielle à imiter certains comportements et raisonnements humains.

Test de Turing

Dans le test de Turing, une machine et un humain répondent textuellement aux questions d'un interrogateur humain. L'interrogateur ne les voit pas mais doit déterminer à partir des réponses textuelles lequel des deux est la machine. Pour passer le test, la machine doit parvenir une bonne partie du temps à tromper l'interrogateur. Ce test a été conçu par Alan Turing en 1950 dans l'article « *Computing Machinery and Intelligence* ». Initialement appelé le « jeu de l'imitation », son but était de fournir une expérience concrète pour déterminer si les machines peuvent penser⁷⁰.



Le test de Turing évalue la capacité à se faire passer pour un humain dans un dialogue écrit, mais incite donc aussi à imiter les défauts humains.

Test du café

Imaginé par Steve Wozniak, le test du café consiste à placer un système intelligent dans un habitat américain moyen et à lui demander de faire un café⁷¹. La réussite du test implique donc plusieurs tâches comme l'orientation dans un environnement inconnu, déduire le fonctionnement d'une machine, trouver les ustensiles nécessaires...

Test de l'étudiant

Proposé par Ben Goertzel, le test de l'étudiant évalue la capacité d'un robot à s'inscrire dans un établissement d'enseignement supérieur, suivre les cours, passer les examens et obtenir le diplôme final⁷².

Test de l'embauche

Proposé par le chercheur Nils John Nilsson, le test de l'embauche consiste à faire postuler un système intelligent à un travail important pour l'économie, où il doit travailler au moins aussi bien qu'un humain⁷³.

Personnalités

Prix Turing

Plusieurs prix Turing (ACM Turing Award) ont été attribués à des pionniers de l'intelligence artificielle, notamment :

- Marvin Minsky (1969) ;
- John McCarthy (1971) ;
- Allen Newell et Herbert Simon (1975) ;
- Edward Feigenbaum et Raj Reddy (1994) ;
- Judea Pearl (2011) ;
- Yann Le Cun, Geoffrey Hinton et Yoshua Bengio (2019)⁷⁴.

Autres personnalités

- Claude Shannon, fondateur de la théorie de l'information, inventeur de la porte logique, concepteur de programmes d'échecs⁷⁵.
- Rodney Brooks, un roboticien australien.

- Demis Hassabis, cofondateur et PDG de DeepMind.
- Ian Goodfellow, inventeur des réseaux antagonistes génératifs.
- Andrew Ng, connu comme directeur scientifique de Baidu et comme créateur de Coursera.
- Terry Winograd, pionnier en traitement automatique des langues.
- Vladimir Vapnik, co-inventeur des machines à vecteurs de support.
- Seymour Papert, ancien directeur du Laboratoire d'intelligence artificielle du MIT.
- Jacques Pitrat, pionnier français en intelligence artificielle symbolique.

En 2023, le magazine Time publie une liste de 100 personnalités influentes du domaine de l'IA et leurs biographies⁷⁶.

Domaines d'application

L'IA permet d'effectuer différents types de tâches, dont :

- la génération de texte, d'images, de vidéos, ou de musiques avec l'IA générative⁷⁷, à partir d'un prompt décrivant le résultat attendu ;
- l'analyse statistique et l'interprétation de grandes masses de données (*big data*) issues d'Internet ;
- la personnalisation de contenus, par exemple par les systèmes de recommandation des réseaux sociaux ;
- le partitionnement et la classification automatiques ;
- la reconnaissance de formes, des visages et la vision en général, etc. ;
- l'intégration automatique d'informations provenant de sources hétérogènes (fusion de données) ;
- l'aide aux diagnostics et à la décision ;
- l'assistance par des machines dans les tâches dangereuses, ou demandant une grande précision ;
- l'automatisation de tâches ;
- le traitement automatique des langues, notamment la traduction automatique ;
- la reconnaissance automatique de la parole (conversion de parole en texte) et le dialogue automatique : se faire comprendre en lui parlant ;
- le raisonnement automatique ;
- la composition musicale automatique (voir les travaux de René-Louis Baron et de l'Ircam ; vers 2016, les recherches de François Pachet portent le développement de *flow machines* telles que Deepbach^{78, 79}) ;
- l'émotion artificielle (voir les travaux de Rosalind Picard sur l'émotion) et l'éventualité d'une subjectivité artificielle ;
- la résolution de problèmes complexes, tels que les problèmes d'allocation de ressources.

En combinant différents algorithmes, on est passé en quelques années de la reconnaissance de l'écriture manuscrite sur des formulaires de chèques bancaires (années 1990) à l'optimisation d'itinéraires entre deux ou plusieurs points, tenant compte des voies disponibles, de la longueur et de la vitesse probable sur chacun des segments, du moyen de locomotion et de souhaits particuliers (éviter les péages, points de passage obligés, caractéristiques touristiques, etc.), puis à la reconnaissance automatisée de documents dactylographiés ou manuscrits, permettant ensuite de les classer selon leur nature avant de les transmettre à des agents spécialisés, humains ou automates, capables d'y apporter la suite appropriée (réponses prédéfinies ou déclenchement d'une chaîne de traitements). Lorsque les algorithmes ne parviennent pas à interpréter les données avec un degré suffisant de certitude, ils sont généralement soumis à des humains. Entre alors en jeu l'apprentissage : en corrélant les caractéristiques des données accumulées et des interprétations apportées ou corrigées par les agents humains, l'IA améliore progressivement ses facultés d'analyse jusqu'à ce que ses prédictions soient suffisamment fiables pour qu'il ne soit plus nécessaire de les faire vérifier par un humain. Il en va de même pour la plupart des situations où l'IA, par ses performances et son coût, surpasse l'humain dans l'analyse de textes, d'images fixes ou animées, d'enregistrements sonores, de données scientifiques, commerciales ou industrielles. L'accroissement des performances conjuguées du matériel et des algorithmes permet en outre de traiter en temps réels les flux de données comme la voix humaine et les images de caméras, ouvrant ainsi la voie à la traduction simultanée, la transcription textuelle, l'identification des individus, la détection de comportements anormaux ou illégaux, voire le dialogue avec des humains.

L'intelligence artificielle est désormais utilisée dans de nombreux domaines. Ses capacités permettent notamment d'automatiser et d'optimiser des tâches complexes, de traiter et d'analyser de vastes quantités de données, et d'améliorer la prise de décision⁸⁰. L'adoption de l'intelligence artificielle est en forte expansion dans les années 2020, stimulée par les avancées en intelligence artificielle générative, en particulier dans les grands modèles de langage, dont la polyvalence ouvre la voie à de nouveaux cas

d'usage⁸¹, ainsi que dans les domaines de la programmation informatique (génération de code), de l'image et du son (photos et vidéos de synthèse, animation du visage en association avec la synthèse vocale) et dans la prévention et l'atténuation des risques (diagnostic médical, situations de crise, accidents industriels, catastrophes naturelles, etc.).

Finance et banques

Plusieurs grands noms de la finance se sont montrées intéressées par de telles technologies, avec des projets comme ceux de Bridgewater Associates où une intelligence artificielle va gérer entièrement un fonds⁸² ou encore la plateforme d'analyse prédictive Sidetrade.

Sont également développés des systèmes de trading algorithmique, dont les gains de vitesses permis par l'automatisation peuvent leur donner un avantage par rapport à des traders humains, en particulier grâce au trading à haute fréquence⁸³.

Militaire

Le domaine militaire utilise de plus en plus l'intelligence artificielle, notamment pour le pilotage automatique, le guidage de missiles, l'identification, le commandement, l'aide à la décision⁸⁴, la cyberguerre et la cyberdéfense⁸⁵, ou pour la documentation et les processus administratifs⁸⁶.

Cette course aux armements est notamment illustrée par le projet *Maven* aux États-Unis⁸⁷. Dès 2015, une IA nommée *ALPHA* a « systématiquement triomphé d'un pilote de chasse chevronné »⁸⁸. En 2018, l'ONU a tenté d'interdire les systèmes d'armes létales autonomes « avant qu'il ne soit trop tard », mais peine encore en janvier 2024 à établir le moindre cadre légal international face aux réticences, notamment de la Russie, des États-Unis et d'Israël, dont le veto peut bloquer une proposition⁸⁹. Des drones tueurs pilotés par intelligence artificielle ont été utilisés lors du conflit ukraino-russe⁹⁰. Le 10 janvier 2024, OpenAI a modifié ses conditions d'utilisation ; il continue d'interdire l'usage de ses services tels que ChatGPT à des fins illégales ou de destruction des biens, mais n'interdit plus explicitement les usages militaires⁹¹. L'intelligence artificielle générative est parfois utilisée par les institutions militaires pour rédiger plus vite la documentation, mais son adoption est limitée par la confidentialité des données, les réglementations, ou le risque d'erreur et le besoin de vérification⁸⁶.

En France, la force opérationnelle IA du ministère des Armées rend en septembre 2019 un rapport détaillant sa stratégie, qui inclut la création d'une Cellule de coordination de l'intelligence artificielle de défense (CCIAD) rattachée à l'Agence de l'innovation de défense⁹². La loi de programmation militaire prévoit un budget de 700 millions d'euros pour les missions en faveur de l'IA, soit une moyenne de 100 millions par an⁹³. En 2021, la France est opposée aux armes totalement autonome, estimant qu'il est au moins nécessaire de conserver une supervision humaine^{94,95}. En février 2025, le ministère des Armées lance sa plateforme *GenIAI.intradef* (co-développée depuis 2022 par plusieurs entités), destinée à améliorer le travail quotidien des militaires et agents civils associés à l'Armée. Cette IA alimente un agent conversationnel et un convertisseur audio-texte ; elle peut lire et analyser des documents et des images et traduire des messages⁹⁶.

Dans le cadre de la guerre Israël-Hamas de 2023-2024, Israël a utilisé deux systèmes d'IA pour générer des cibles à frapper : Habsora (soit « l'évangile ») a été utilisé pour dresser une liste de bâtiments à cibler, tandis que Lavander a produit une liste de 37 000 personnes à cibler^{97,98}. La liste des bâtiments comprenait les maisons privées à Gaza de personnes soupçonnées d'être affiliées à des membres du Hamas. Les responsables de Tsahal affirment que le programme répond au problème antérieur du manque de cibles de l'armée de l'air. Auparavant, Tsahal était en mesure d'identifier 50 cibles par an, tandis que le programme en produit 100 par jour⁹⁸. La combinaison de la technologie de ciblage de l'IA et du changement de politique consistant à éviter les cibles civiles a entraîné un nombre sans précédent de morts civils palestiniens^{99,100}.

Médecine

La médecine a aussi vu de grands progrès grâce à l'utilisation de systèmes d'aide au diagnostic ou de diagnostic automatisé¹⁰¹.

En 2018, Google DeepMind, filiale de Google spécialisée dans la recherche avancée en intelligence artificielle, a publié les résultats d'une expérimentation d'intelligence artificielle pouvant détecter les maladies oculaires. Les résultats indiquent que l'IA le fait avec une marge d'erreur plus faible que les ophtalmologues¹⁰².

Google DeepMind a également conçu AlphaFold, un système d'intelligence artificielle utilisant l'apprentissage profond qui permet de prédire la façon dont des protéines se replient. Les protéines sont composées de chaînes d'acides aminés et la façon dont elles se replient détermine leur fonction. Cette nouvelle méthode, introduite en 2018 et améliorée en 2020, est nettement plus rapide que les approches traditionnelles et a été décrite comme une révolution dans le domaine de la recherche en biologie^{103,104}.

La France crée en 2019 le Health Data Hub afin d'encadrer et de faciliter l'utilisation des données de santé dans la recherche¹⁰⁵.

En 2023, la version de ChatGPT reposant sur GPT-4 s'est montrée facilement capable d'obtenir le diplôme de médecin aux États-Unis¹⁰⁶.

L'intelligence artificielle (IA) est de plus en plus utilisée dans le domaine médical, transformant les pratiques cliniques et facilitant le diagnostic, le traitement et la gestion des maladies. Les algorithmes d'apprentissage automatique, un sous-ensemble de l'IA, analysent de grandes quantités de données médicales pour en extraire des modèles et des tendances. Cette capacité d'analyse a permis des avancées significatives dans le diagnostic précoce de maladies comme le cancer et les maladies cardiaques, où les systèmes basés sur l'IA peuvent détecter des anomalies à partir d'images médicales avec une précision parfois supérieure à celle des praticiens humains¹⁰⁷.

Un des domaines où l'IA s'avère particulièrement efficace est celui de l'imagerie médicale. Des outils comme les réseaux de neurones convolutifs sont utilisés pour analyser les radiographies, les IRM et autres types d'imageries, permettant de repérer les signes précoces de pathologies complexes. Par exemple, une étude menée par McKinney et al. (2020) a démontré que l'IA pouvait réduire les faux positifs et négatifs dans le dépistage du cancer du sein, améliorant ainsi la précision et la rapidité du diagnostic¹⁰⁸.

De plus, l'IA joue un rôle clé dans la médecine personnalisée, où elle aide à adapter les traitements en fonction des caractéristiques génétiques et biologiques des patients. Grâce à l'analyse des données génomiques, les médecins peuvent élaborer des thérapies spécifiques pour des maladies comme le cancer ou les maladies génétiques rares. L'IA facilite également le développement de nouveaux médicaments, en identifiant des molécules potentielles pour le traitement de certaines maladies, un processus qui nécessitait auparavant plusieurs années¹⁰⁹.

Cependant, malgré ses promesses, l'usage de l'IA en médecine soulève des questions éthiques et de sécurité, notamment en ce qui concerne la protection des données des patients et la transparence des algorithmes¹¹⁰. La Food and Drug Administration (FDA) aux États-Unis et d'autres organismes réglementaires travaillent à définir des cadres pour encadrer l'utilisation sécurisée et éthique de l'IA dans les soins de santé.

Renseignement policier

Un usage de l'IA se développe dans le domaine de la prévention des crimes et délits. La police britannique, par exemple, développe une IA de ce genre, annoncée comme pouvant être opérationnelle dès mars 2019¹¹¹. Baptisée National Data Analytics Solution (Solution nationale d'analyse de données ou NDAS), elle repose sur l'IA et des statistiques et vise à estimer le risque qu'une personne commette un crime ou en soit elle-même victime, pour orienter les services sociaux et médicaux qui peuvent la conseiller.

L'usage d'outils de prédiction des crimes à partir des données préalablement existantes est toutefois l'objet de controverses, compte tenu des biais sociaux (notamment raciaux) qu'il comporte¹¹². En effet, la logique d'identification de schémas propre à ces technologies joue un rôle de renforcement des préjugés déjà existants.

Cybercrime

L'intelligence artificielle (IA) est de plus en plus exploitée dans le domaine du cybercrime, comme le révèle une étude de la société spécialisée en cybersécurité SlashNext. Cette tendance croissante à l'utilisation de l'IA pour commettre des crimes en ligne montre une sophistication accrue des attaques. L'entreprise SlashNext a notamment identifié l'usage de deux IA malicieuses, FraudGPT et WormGPT, tout en suggérant que ces découvertes ne représentent que la partie visible d'une menace potentiellement colossale. Lors de leurs investigations, les chercheurs ont également mis en lumière l'existence de DarkBart et DarkBert^b, deux chatbots malveillants en développement, capables d'intégrer la technologie de reconnaissance d'images de Google Google Lens. Ces chatbots pourraient envoyer du texte et des images, et participer à des attaques d'ingénierie sociale avancées. Face à cette menace croissante, les solutions actuelles de lutte contre le cybercrime semblent insuffisantes, estime un rapport d'ImmuneFi, qui souligne les limites de certaines IA, telles que ChatGPT, dans la détection des exploits¹¹³.

Droit

Le droit fait appel à l'IA dans la perspective de prédire les décisions de justice, d'aider à la décision et de trancher les cas simples¹¹⁴. L'Estonie a par exemple développé une intelligence artificielle capable de prendre des décisions de justice sur des délits mineurs¹¹⁵. Les États-Unis utilisent par ailleurs dans certaines juridictions le système COMPAS (en) (Correctional

Offender Management profiling for Alternative Sanctions), un système d'aide à la prise de décision pour les juges¹¹⁵. Plusieurs *startups* se sont spécialisées dans ce créneau, créant le domaine de la *legaltech*¹¹⁶.

Logistique et transports

Le domaine de la *logistique* a vu certains projets utilisant de l'intelligence artificielle se développer notamment pour la gestion de la *chaîne logistique* (*supply chain*) ou des problématiques de livraison telle celle du *dernier kilomètre*¹¹⁷.

L'intelligence artificielle est également fortement utilisée dans le domaine des *transports en commun*, car elle permet de faciliter la régulation et la gestion du trafic au sein de réseaux de plus en plus complexes, comme le système *UrbanLoop* en cours d'étude dans la ville de *Nancy*¹¹⁸.

Même si les problèmes d'optimisation de temps de trajet ou de transports font partie des plus anciennes applications de solutions à base d'intelligence artificielle (voir le *problème du voyageur de commerce* ou l'*algorithme de Dijkstra*), les avancées récentes, notamment en *apprentissage profond*, ont permis des progrès significatifs en matière de précision. Certains projets comme *Google Maps* utilisent par exemple des systèmes d'IA en milieu urbain pour compenser la réflexion du signal *GPS* sur les immeubles avoisinants¹¹⁹, ou pour cartographier des zones où peu d'informations sont disponibles^{120, 121}.

Plusieurs entreprises ont par ailleurs annoncé avoir développé des programmes de recherche en *voiture autonome*, notamment *Google* à travers sa filiale *Waymo*, l'entreprise française *Navya* ou encore *Tesla*.

Industrie

Les systèmes intelligents deviennent monnaie courante dans de nombreuses industries. Plusieurs tâches peuvent leur être confiées, notamment celles considérées comme trop dangereuses pour un humain¹²². Certains applications se concentrent sur les systèmes de *maintenance prédictive*, permettant des gains de performance grâce à une détection des problèmes de production en amont.

Robotique

La *robotique* a recours à l'intelligence artificielle à plusieurs égards, notamment pour la perception de l'environnement (*objets* et *visages*), l'*apprentissage* et l'intelligence artificielle développementale^{123, 124}.

L'*interaction homme-robot* manque encore souvent de naturel et est un enjeu de la robotique. Il s'agit de permettre aux robots d'évoluer dans le monde dynamique et social des humains et d'échanger avec eux de façon satisfaisante¹²³. L'échange nécessite également, à l'inverse, une évolution du regard que les humains portent sur les robots ; selon Véronique Aubergé, chercheuse à l'*Université Grenoble-Alpes* « la vraie révolution n'est pas technologique, elle est culturelle ». D'ores et déjà, à travers les robots dotés d'intelligence artificielle, tel *Google Home*, les utilisateurs combleraient un *isolement social*¹²³.



Un robot NAO en 2014.

Jeux vidéo

L'intelligence artificielle est par exemple utilisée pour animer les *personnages non-joueurs* de jeux vidéo, qui sont conçus pour servir d'opposants, d'aides ou d'accompagnants lorsque des joueurs humains ne sont pas disponibles ou désirés. Différents niveaux de complexité sont développés, d'une simple assistance à un comportement complexe imitant (ou dépassant) les meilleurs joueurs humains.

Art

Dès la fin des années 1980, des artistes s'emparent de l'intelligence artificielle pour donner un comportement autonome à leurs œuvres. Les Français *Michel Bret*, *Edmond Couchot* et *Marie-Hélène Tramus* sont des pionniers, ainsi qu'en témoignent des œuvres comme *La Plume* et *Le Pissenlit* (1988)¹²⁵, puis *La Funambule* (2000), animée par un *réseau de neurones*. L'Américain *Karl Sims*, en partenariat avec la société *Thinking Machines*, crée en 1993 *Genetic Images*, machines incorporant [Comment ?] des algorithmes génétiques. Le couple franco-autrichien *Christa Sommerer* et *Laurent Mignonneau* crée depuis le début des années

1990 de nombreuses œuvres dans le champ de la vie artificielle, parmi lesquelles *Interactive plant growing* (1992) ou *A-Volve* (1994)^[réf. nécessaire]. Le Français Florent Aziosmanoff propose quant à lui de considérer que l'emploi de l'intelligence artificielle dans l'art conduit à l'émergence d'une nouvelle discipline d'expression, qu'il nomme le Living art¹²⁶.

Le 23 octobre 2018, la société de vente aux enchères Christie's met en vente le tableau *Portrait d'Edmond de Belamy* réalisé par une intelligence artificielle à l'aide de réseaux antagonistes génératifs. La peinture est signée par la formule mathématique à l'origine de sa création (« Min (G) max (D) Ex [log (D(x))] + Ez [log(1-D(G(z)))] »)¹²⁷. Cette vente soulève de nombreux débats sur son statut de création artistique et sur l'auteur de l'œuvre : il peut être l'intelligence artificielle elle-même ou les trois créateurs qui l'ont programmée¹²⁸.

Des réseaux antagonistes génératifs ont été utilisés pour créer de fausses images réalistes, comme par le générateur de visages StyleGAN introduit en 2018¹²⁹, ou pour écrire *Terre Seconde* de Grégory Chatonsky, qui imagine en 2019 une version alternative de la planète Terre¹³⁰.

Dès 2022 apparaissent des modèles d'intelligence artificielle qui sont capables de créer des images réalistes à partir de descriptions textuelles, comme Midjourney, Stable Diffusion et DALL-E^{131, 132}. En mars 2023, des fausses photos d'actualité sont ainsi générées et diffusées sur Internet, mettant en scène des personnalités dans des situations extravagantes (le président Macron ramassant des poubelles, Donald Trump arrêté par des policiers¹³³, le pape François habillé en doudoune blanche¹³⁴). Elles deviennent rapidement virales, augmentant les craintes de manipulation de l'opinion¹³⁵. Cela pose aussi des questions de droits d'auteur¹³⁶.



Une fausse photo du jeune Elon Musk jouant au jeu vidéo Mars Marauder en 1995, générée par l'IA.

Image externe

Q Fausses photos d'actualité générées par Midjourney (mars 2023) ([https://www.liberation.fr/resizer/Pn9RvKgk0v-CTyhuXAhk2EE7fP0=/600x0/filters:format\(jpg\):quality\(70\):focal\(774x418:784x428\)/cloudfront-eu-central-1.images.arcpublishing.com/liberation/NRFBKWQBPSRDNDA6ZDNBIYGZCM.jpg](https://www.liberation.fr/resizer/Pn9RvKgk0v-CTyhuXAhk2EE7fP0=/600x0/filters:format(jpg):quality(70):focal(774x418:784x428)/cloudfront-eu-central-1.images.arcpublishing.com/liberation/NRFBKWQBPSRDNDA6ZDNBIYGZCM.jpg))¹³⁷.

De plus en plus de romans ont été coécrits avec une IA générative, tels que *Internes* en 2022¹³⁸ ou *東京都同情塔* (« La Tour de la compassion de Tokyo »), qui a reçu le prix Akutagawa en 2024¹³⁹.

En février 2024, le modèle Sora de OpenAI s'est montré capable de générer des vidéos relativement réalistes¹⁴⁰.

Des modèles d'IA capables de générer un morceau de musique à partir d'une description du style souhaité ont également fait leur apparition, comme Suno AI en 2023 et Udio en 2024¹⁴¹.

Autres domaines

La domesticité, avec des robots employé de maison¹⁴², ou pour certaines tâches précises comme en domotique.

En programmation informatique, notamment pour la maintenance prévisionnelle, l'autocomplétion ou l'aide au développement¹⁴³.

En journalisme : des IA (appelées improprement « robots journalistes ») pourraient à terme aider les journalistes en les débarrassant de certaines tâches, notamment la veille, le bâtonnage de dépêches ou la vérification des *fake news*¹⁴⁴.

La Corée du Sud propose la toute première animatrice télé virtuelle en novembre 2020 lors d'un JT¹⁴⁵.

En design : la conception assistée par ordinateur fait depuis longtemps appel à des algorithmes d'optimisation. En 2019, le créateur Philippe Starck lance ainsi une chaise développée en collaboration avec la société Autodesk, la « A.I.chair »¹⁴⁶.

Débats et enjeux

Les succès en IA encouragent les spéculations. Dans les milieux technophiles, on verse en général dans l'enthousiasme, le mouvement transhumaniste en est la meilleure expression. Mais certains s'inquiètent et s'interrogent, parfois alarmistes, y compris dans la sphère de la haute technologie. Ainsi, des figures réputées telles que Bill Gates — ancien PDG de Microsoft et

« figure emblématique de la révolution informatique de la fin du xx^e siècle »¹⁴⁷ — pensent qu'il faut rester très prudent quant aux développements futurs de ces technologies, qui pourraient devenir liberticides ou dangereuses.

Le développement de l'intelligence artificielle suscite un grand nombre de questions, notamment en ce qui concerne la possibilité pour les IA ou algorithmes d'accéder un jour à la conscience, d'éprouver des émotions ou de finalement se substituer aux humains. Certaines réactions sont ouvertement optimistes, d'autres sont au contraire pessimistes. En 2016, l'INRIA publie un premier Livre blanc consacré à l'IA¹⁴⁸.

Le philosophe Daniel Andler considère en 2023 que le rêve d'une intelligence artificielle qui rejoindrait celle de l'homme est une chimère, pour des causes conceptuelles et non techniques. L'intelligence humaine va selon lui plus loin que la simple résolution de problèmes : toutes ses autres tâches, reposant sur des affects, de la spontanéité et une forme de contingence, ne seront jamais accessibles à une intelligence non humaine¹⁴⁹.

Singularité

Une description d'un possible avenir de l'intelligence artificielle a été faite par le statisticien anglais Irving John Good :

« Supposons qu'existe une machine surpassant en intelligence tout ce dont est capable un homme, aussi brillant soit-il. La conception de telles machines faisant partie des activités intellectuelles, cette machine pourrait à son tour créer des machines meilleures qu'elle-même ; cela aurait sans nul doute pour effet une *réaction en chaîne* de développement de l'intelligence, pendant que l'intelligence humaine resterait presque sur place. Il en résulte que la machine ultra intelligente sera la dernière invention que l'homme aura besoin de faire, à condition que ladite machine soit assez docile pour constamment lui obéir. »

— Irving John Good¹⁵⁰

Cette hypothétique courte période de progrès drastique dont il est difficile de prédire les conséquences a été nommée « singularité ». Elle a été étudiée par Vernor Vinge dans les années 1990 et par Ray Kurzweil¹⁵¹ dans les années 2000¹⁵². Ce concept est central pour de nombreux transhumanistes, qui s'interrogent sur les dangers ou les espoirs d'un tel scénario, certains allant jusqu'à envisager l'émergence d'un « dieu » numérique appelé à prendre le contrôle du destin de l'humanité, ou à fusionner avec elle¹⁵². En 2014, Nick Bostrom a popularisé le concept de superintelligence artificielle¹⁵³.



Pour l'Américain Ray Kurzweil, l'intelligence artificielle dépassera bientôt l'intelligence naturelle.

Risques existentiels

Le développement de l'intelligence artificielle génère de l'enthousiasme, mais aussi de vives inquiétudes. Certains auteurs de science-fiction, tels Isaac Asimov, William Gibson ou Arthur C. Clarke, sur le modèle du récit de *L'Apprenti sorcier*, décrivent le risque d'une perte de contrôle des humains sur le processus technique. Dans les années 2010, différents intellectuels ont également pris position, notamment l'astrophysicien Stephen Hawking, selon qui l'intelligence artificielle risque réellement de surpasser un jour l'intelligence humaine et de finir par dominer l'humanité, voire de s'y substituer^{155, 156}. Il pose en novembre 2017 au salon technologique Web Summit de Lisbonne la question suivante « Serons-nous aidés par l'intelligence artificielle ou mis de côté, ou encore détruits par elle ? »¹⁵⁷.

Dans le milieu de la haute technologie, certains expriment publiquement des craintes similaires. C'est ainsi le cas, en 2015, de Bill Gates, Elon Musk et Bill Joy¹⁵⁸. Selon le spécialiste américain de l'informatique Moshe Vardi en 2016, l'intelligence artificielle pourrait mettre 50 % de l'humanité au chômage. « Nous approchons d'une époque où les machines pourront surpasser les hommes dans presque toutes les tâches »¹⁵⁹.

Hilary Mason, directrice de la recherche à Cloudera, a critiqué en 2018 le sensationnalisme entourant l'intelligence artificielle et prône une vision pragmatique et opérationnelle de cette technologie¹⁶⁰.



Un premier sommet en sûreté de l'IA s'est tenu en novembre 2023 pour aborder les risques liés aux modèles d'IA les plus avancés¹⁵⁴. Il a été suivi du Sommet de Paris pour l'action sur l'IA (2025).

Ajouter un mécanisme d'arrêt pourrait ne pas suffire face à une IA suffisamment avancée, qui pourrait s'avérer en mesure de cacher des intentions dangereuses, de manipuler ses détenteurs, de désactiver le mécanisme d'arrêt ou encore de se dupliquer. Selon Nick Bostrom en 2015, la seule solution viable à long terme consiste à trouver comment aligner les intelligences artificielles avec des valeurs humaines et morales¹⁶¹ :

« nous ne devrions pas être confiants dans notre capacité à garder indéfiniment un génie superintelligent enfermé dans une bouteille. Je crois que la réponse ici est de trouver comment créer une IA superintelligente de sorte que si — ou plutôt quand — elle s'échappe, elle reste sans danger, parce qu'elle est fondamentalement de notre côté, elle partage nos valeurs. »

— Nick Bostrom

Roman V. Yampolskiy, professeur de science informatique à l'Université de Louisville, évoque pourquoi et comment une IA obtient un résultat, pour s'assurer qu'il corresponde bien à l'attendu, sans biais : « si nous nous habituons à accepter les réponses de l'IA comme des paroles d'oracles ne nécessitant pas d'explication, alors nous serons incapables de vérifier si ces résultats ne sont pas biaisés ou manipulés »¹⁶².

En mai 2023, une déclaration du Center for AI Safety (« Centre pour la sûreté de l'IA ») affirme que réduire le risque d'extinction de l'humanité lié à l'IA devrait être une priorité mondiale, au même titre que pour d'autres risques civilisationnels tels les pandémies ou les guerres nucléaires. Elle est signée par des dirigeants de laboratoires d'IA comme OpenAI, Google DeepMind ou Anthropic, ainsi que par des chercheurs en intelligence artificielle^{163, 164}.

Critique de la technique et de la technologie

Comme l'explique l'historien François Jarrige, la critique de l'intelligence artificielle trouve son origine dans celle - plus ancienne et plus générale - des techniques et de la technologie, dont Lewis Mumford (aux États-Unis)¹⁶⁵, Jacques Ellul (en France)¹⁶⁶ et Günther Anders (en Allemagne)¹⁶⁷ sont au xx^e siècle les principaux instigateurs, et qui inspire aujourd'hui différents cercles militants (en France, par exemple : Pièces et Main d'Œuvre¹⁶⁸ et Technologos^{169, 170}).

Dans un rapport en date de février 2018 intitulé *The Malicious Use of Artificial Intelligence* 26 experts spécialistes en intelligence artificielle mettent en garde contre les dangers d'un usage criminel de l'IA : augmentation de la cybercriminalité, conduire à des utilisations de drones à des fins terroristes, manipulation de masse, etc.¹⁷¹.

Impact environnemental

Un autre problème est l'énorme quantité de ressources rares, de serveurs et d'énergie consommée par l'informatique sous-jacente à l'IA. Google admet en 2024 qu'il lui sera très difficile de tenir ses engagements de neutralité carbone, car depuis 2019 ses émissions de gaz à effet de serre ont augmenté de 48 % du fait du développement de l'IA¹⁷².

Afin de réduire leur empreinte carbone, certains fournisseurs d'IA (Microsoft, Amazon et Oracle) se tournent vers les marchés volontaires de la compensation carbone, et plus largement de solutions de décarbonation, comme celles de capture du CO₂¹⁷³. Ils s'orientent également vers les centrales nucléaires pour couvrir leur besoin en électricité¹⁷⁴. Microsoft a notamment signé en 2024 un contrat avec Constellation Energy pour l'achat de 837 MWe de capacité électrique fournie par la centrale nucléaire de Three Mile Island, pour une durée de vingt ans à partir de 2028¹⁷⁵.

Open source

Divers projets open source d'IA ont été menés par Hugging Face¹⁷³, EleutherAI¹⁷⁶, Google¹⁷⁷, ou Meta¹⁷⁶.

En mars 2023, comme alternative aux géants du Web et du cloud computing, qui ont le plus de pouvoir et d'influence, Mozilla annonce vouloir investir 30 millions de dollars dans un projet baptisé Mozilla.ai, qui est à la fois une startup et une communauté, indépendante des géants de la tech et de la recherche universitaire¹⁷⁸. Le projet vise à créer, dans le respect des valeurs de son manifeste (notamment transparence et responsabilité), un système d'IA « open source, digne de confiance et indépendant »¹⁷⁹.

De nombreux grands modèles de langage comme Mistral¹⁸⁰, Llama 3, Vicuna et Falcon¹⁸¹ sont rendus *open weight*, ce qui signifie que l'architecture et les paramètres entraînés du modèle d'IA sont rendus publics (*open source* impliquerait notamment de partager les données d'entraînement, ce qui souvent n'est pas le cas)¹⁸². Ces modèles peuvent être librement ajustés, ce qui permet notamment aux entreprises de les spécialiser pour leurs propres données et pour leur cas d'usage¹⁸³. Ces modèles d'IA facilitent l'innovation et la recherche, mais peuvent facilement être détournés. Ils peuvent être réentraînés de sorte à rendre inefficaces les mesures de sécurité, telles que le refus de répondre à une requête dangereuse. Certains chercheurs estiment ainsi

que si des modèles développent un jour des capacités dangereuses, comme le fait de faciliter drastiquement les cyberattaques ou le bioterrorisme, ils ne devraient pas être rendus *open weight*, d'autant plus qu'une fois diffusé sur internet, un modèle ne peut en général plus être supprimé partout^{184, 182}.

IA et emploi

Historiquement, l'innovation technologique a généralement été accompagnée d'une croissance de la productivité et de la création de nouveaux emplois pour compenser les pertes¹⁸⁵. Cependant, de nouvelles inquiétudes ont émergé avec l'essor de l'IA générative, telle que ChatGPT, capable de manipuler du texte, des images ou du code informatique¹⁸⁶.

Certains économistes sont sceptiques à l'idée d'un chômage de masse, citant des précédents historiques où le marché du travail s'est adapté à des bouleversements technologiques^{187, 188}. D'autres estiment que l'IA générative représente un changement plus profond, qui ne consiste pas seulement à automatiser des tâches répétitives. L'IA peut être appliquée à tous les secteurs, et le nombre de tâches où l'humain reste plus compétent que l'IA est amené à diminuer¹⁸⁹.

Une étude de Goldman Sachs a estimé en 2023 que les deux tiers des travailleurs européens sont exposés à divers degrés d'automatisation, et qu'un quart des emplois pourraient être remplacés par l'IA. L'étude prédit notamment que les professions administratives et légales seront particulièrement touchées, là où les métiers manuels en extérieur seront peu affectés¹⁸⁶.

Une solution envisagée dans le scénario d'un chômage de masse est celle d'une forme de redistribution des richesses avec un revenu universel. Les financements pourraient dans ce cas venir d'une taxe sur les richesses produites par les machines¹⁹⁰.

Défaut de représentativité

L'IA s'appuie sur l'analyse de données qui peinent souvent à représenter fidèlement la réalité, soit parce ces données ne proviennent que de sources numériques existantes (le plus souvent sur Internet) alors que quantités d'informations pertinentes pour le domaine considéré n'y figurent pas, soit parce que certaines catégories de données pourtant pertinentes n'ont pas été prises en compte, ou que des données non pertinentes l'ont été. Il en résulte alors une inadéquation entre les données utilisées pour l'entraînement des algorithmes et les données cibles sur lesquelles l'algorithme devra opérer, et par conséquent des erreurs de diagnostic qui peuvent, dans certains cas, induire des décisions inappropriées ou injustes^{191, 192, 193}.

Risques politiques et sociaux

Les réseaux sociaux et les bots ont favorisé la propagation de nombreuses fausses croyances et des dérives dans les débats démocratiques, lesquelles ont entraîné une certaine défiance vis-à-vis de la science, des élites intellectuelles et des médias d'information traditionnels¹⁹⁴ [réf. incomplète]. En conséquence, la progression et l'adoption rapides des techniques sous-jacentes de l'IA inspirent de multiples craintes quant aux impacts que celle-ci pourrait avoir sur les comportements individuels et collectifs. Ces conditions ouvrent la possibilité pour des opérateurs majeurs de la société de fausser le réel pour influencer ou manipuler les citoyens. Les acteurs économiques emploient l'IA pour influencer les consommateurs et pour optimiser le travail de leurs employés au point de les transformer en robots¹⁹⁵ [source secondaire souhaitée]. L'éventualité que l'IA intervienne dans les débats d'idées ou dans la conduite des affaires individuelles ou collectives mène à redouter qu'elle n'altère ou n'affaiblisse les institutions politiques¹⁹⁴ [réf. incomplète] et les pouvoirs, voire les capacités des humains¹⁹⁶ [source secondaire souhaitée].

Dès 2010, Nicholas Carr alertait sur l'usage intensif des outils numériques, qui modifie notre manière de penser et de traiter l'information, ce qui peut avoir des conséquences à long terme sur notre cognition, notamment affaiblir notre mémoire et notre capacité à comprendre les informations de manière approfondie et affecter la créativité, l'empathie et le débat intellectuel¹⁹⁷.

Les principales menaces couramment envisagées, consécutives à des biais ou au détournement des algorithmes d'IA sont^{198, 199} :

- la désinformation et la manipulation du public pour des raisons crapuleuses, religieuses ou idéologiques ;
- les fausses vidéos et hypertrucages représentant des personnalités faisant ou disant des choses qu'ils n'ont pas faites ou dites ;
- la corruption de données par vandalisme ou pour déstabiliser ou soutirer de fortes sommes d'argent (par rançongiciel) en empêchant l'accès à des données ou des services ;
- l'escroquerie par usurpation d'identité ou pour soutirer de l'argent contre des biens ou des services fictifs ;
- le chantage à grande échelle, par des menaces d'agression ou de révélations sur la vie privée ;
- la perte de contrôle ou le piratage d'infrastructures publiques (eau, énergie, signalisation routière, Internet, services en ligne...) à des fins de déstabilisation ou de chantage ;

- la perte de contrôle de robots, de véhicules autonomes ou de drones militaires, ou leur vol ou leur détournement à des fins agressives ou de chantage ;
- la manipulation de marchés financiers ;
- la surveillance de masse.

De plus, l'IA produit parfois des résultats contre-intuitifs, bien que fiables, ou propager de fausses croyances, qui contribuent, parmi d'autres facteurs, à éroder la confiance d'une partie de la population dans le discours rationnel, les théories scientifiques, le système médiatique, les institutions et les élites en général, et à la remise en cause de leurs légitimités²⁰⁰ et du processus démocratique¹⁹⁴[réf. incomplète].

Selon Yuval Noah Harari, jusqu'ici, du fait de défauts d'alignement et faute de mécanismes suffisants d'autocorrection, les objectifs qui sont assignés aux algorithmes des réseaux sociaux ont déjà induit ou influencé phénomènes de société indésirables, désinformation, hypertrucages, propos outranciers ou haineux, comme dans le cas du génocide des Rohingya au Myanmar²⁰¹, en outrepassant les règles éthiques qui sont implicites pour des humains. De plus, la capacité de l'IA à fonder ses analyses sur des volumes considérables de données et de très nombreux critères donne l'illusion de son infaillibilité, comparativement aux méthodes d'analyse par des individus ou des organisations humaines, fondés sur des corpus d'information beaucoup plus réduits et un nombre de critères limité dont certains très subjectifs. Certains algorithmes d'IA ont même la faculté d'inspirer confiance aux humains en misant sur leur subjectivité ou leur émotivité, voire de les tromper « délibérément » pour parvenir à leurs fins^{202, 203}. L'exploitation de telles possibilités pourrait permettre de constituer un ou plusieurs nouveaux systèmes de croyances et de règles morales, politiques et sociales, à l'instar des religions ou des régimes politiques, instituant comme vérité une certaine interprétation du réel et établissant une nouvelle forme d'ordre social. L'IA pourrait ainsi conduire à une certaine anarchie ou permettre l'instauration de régimes totalitaires dans les pays démocratiques (*a contrario*, des IA pourraient amplifier ou faire naître des idées subversives au sein des régimes autoritaires). Pour préserver les démocraties, il conviendrait donc que soient établies des régulations faisant intervenir des humains pour contenir les possibles dérives de l'IA¹⁹⁴[réf. incomplète].

L'emploi de l'IA pour noter les entreprises et citoyens, sur le modèle du crédit social, est déjà à l'œuvre en Chine. Ce procédé fait craindre une surveillance constante des personnes et des organisations, pouvant entraîner une forme d'asservissement des individus, ou des sanctions et récompenses excessives¹⁹⁴[réf. incomplète].

Éric Sadin récuse le terme d'« intelligence ». Selon lui, l'IA porte en elle un modèle de société utilitariste et rationaliste inspiré de manière extrêmement lacunaire par le fonctionnement du cerveau humain et visant à uniformiser les comportements en temps réel et à tous moments. Le libéralisme économique, saisissant le profit qu'il peut en tirer, investit massivement dans l'IA afin d'exploiter l'inclination naturelle des humains à la facilité et les capacités extraordinaires de l'IA à expertiser la complexité du réel et à orienter nos décisions. Cela nous conduit insidieusement à nous fier aux réponses formulées par des algorithmes au détriment de critères d'analyse plus multisensoriels et de choix plus subjectifs, donc plus humains. Nous pourrions alors céder progressivement à l'IA notre pouvoir d'analyse, de jugement et de décision et consentir à nous soumettre plus ou moins consciemment à une « intelligence » supérieure, puis finalement renoncer même à penser. Notre addiction aux « assistants personnels » et aux capteurs physiologiques ainsi que leur présence continue à nos côtés pourraient les amener progressivement à s'exprimer, décider voire imaginer à notre place et en notre nom, d'abord avec, puis sans notre consentement¹⁹⁶.

Selon Laurent Alexandre, l'IA contribuera à accroître les inégalités en écartant les individus les « moins aptes » de l'activité économique et sociale, et le revenu universel constituerait à ses yeux un moyen d'asseoir la domination des « élites » sur les populations défavorisées²⁰⁴.

Réglementation

En 2019, l'OCDE et le G20 adoptent une série de principes sur l'IA²⁰⁵. Le Partenariat mondial sur l'intelligence artificielle est lancé en juin 2020 pour promouvoir la conformité du développement de l'IA aux droits de l'homme et aux valeurs démocratiques. Il est hébergé par l'OCDE à Montréal et à Paris²⁰⁶. Une plateforme de communication, AI for Good (« l'IA pour le bien »), est créée pour faciliter les échanges et faire avancer les objectifs de développement durable de l'ONU grâce à l'IA²⁰⁷.

En 2023, plus de 1 600 politiques publiques et stratégies sur l'IA sont recensées dans le monde²⁰⁸. Elles viennent en particulier de l'Union européenne, la Chine, les États-Unis et le Royaume-Uni. Après les avancées réglementaires de l'UE et de la Chine, la Maison-Blanche publie en octobre 2023 un décret sur l'IA « sûre, sécurisée et digne de confiance ». En novembre 2023 a lieu un premier sommet en sécurité de l'IA au Royaume-Uni²⁰⁸.

En Europe, les services numériques sont réglementés par le RGPD²⁰⁹, le règlement sur les services numériques et la législation sur les marchés numériques. Pour l'intelligence artificielle en particulier, la législation sur l'intelligence artificielle (*Artificial Intelligence Act*, ou *AI Act* en anglais) définit quatre niveaux de risques pour les applications d'IA et met en avant des exigences de transparence, de protection des données, de sécurité et d'éthique²¹⁰.

En 2017, les Émirats arabes unis sont le premier pays au monde à se doter d'un ministre dédié à l'intelligence artificielle : Omar Sultan Al Olama²¹¹.

Appels à des règles éthiques pour l'IA

Dans la seconde moitié des années 2010, des lanceurs d'alerte et des enquêtes révèlent que l'IA, encore émergente, a déjà été utilisée à des fins malveillantes pour faire basculer des processus électoraux. Le premier cas notable a été la plate-forme RIPON, secrètement créée par le Groupe SCL, à la demande de Steve Bannon et du milliardaire américain Robert Mercer. Cette plateforme, principalement au service de groupes politiques libertariens de droite, a été un outil de désinformation, de production et de diffusion de *fake news* à grande échelle^{212,213}. Ripon, impliquée dans le scandale Facebook-Cambridge Analytica/Aggregate IQ), joua un rôle important dans la manipulation d'un grand nombre d'électeurs, notamment pour faire élire Donald Trump lors de l'élection présidentielle américaine de 2016, pour faire advenir le Brexit²¹⁴, ainsi que pour orienter des dizaines d'élections dans le monde.

Face à ces dérives, les géants du secteur de l'IA ont réagi en créant le 28 septembre 2016 un « partenariat pour l'intelligence artificielle au bénéfice des citoyens et de la société »²¹⁵. L'année suivante, Google DeepMind se dote d'une unité interne pour aborder les questions éthiques²¹⁶ et la conférence d'Asilomar réunit des personnalités influentes du domaine pour élaborer une charte²¹⁷ visant à réglementer les développements de l'IA²¹⁸.

Le 18 juillet 2018, 2 400 chercheurs, ingénieurs et personnalités du secteur de l'intelligence artificielle signent une lettre ouverte²¹⁹, s'engageant à « ne jamais participer ou soutenir le développement, la fabrication, le commerce ou l'usage d'armes létales autonomes ». La lettre précise que « La décision de prendre une vie humaine ne devrait jamais être déléguée à une machine ». Parmi les signataires se trouvent Elon Musk, les dirigeants de Google DeepMind Stuart Russell, Yoshua Bengio, et Toby Walsh²²⁰.

Fin 2020, l'UNESCO rejoint (en tant qu'observateur, comme l'OCDE) le conseil et le comité directeur du Partenariat mondial sur l'intelligence artificielle, avec la possibilité de participer activement aux travaux de ces organismes²²¹.

La publication en février 2020 d'un *Livre blanc sur l'intelligence artificielle*²²², pose les bases du règlement sur l'intelligence artificielle de 2021 par la Commission européenne, qui vise à encadrer les risques et les problèmes éthiques de ces technologies²²³. Ce projet classe les risques en quatre catégories, dont la plus grave est qualifiée comme suit :

« Risque inacceptable : les systèmes d'IA considérés comme une menace évidente pour la sécurité, les moyens de subsistance et les droits des personnes seront interdits. Il s'agit notamment des systèmes ou applications d'IA qui manipulent le comportement humain pour priver les utilisateurs de leur libre arbitre (par exemple, des jouets utilisant une assistance vocale incitant des mineurs à avoir un comportement dangereux) et des systèmes qui permettent la notation sociale par les États²²⁴. »

En décembre 2022, le « premier forum mondial sur l'éthique de l'IA », réunion ministérielle internationale, est réuni à Prague, sous l'égide de l'Unesco²²⁵.

La même année, l'Unesco, estimant que « l'autorégulation de l'industrie n'est manifestement pas suffisante pour éviter ces préjudices éthiques », a publié un communiqué (adopté le 23 novembre 2021) demandant à tous les États de mettre en œuvre sa recommandation sur l'éthique de l'intelligence artificielle²²⁶ afin de construire un cadre législatif et éthique pour l'IA.

L'objectif est de n'utiliser l'IA que lorsque les atouts qu'elle peut offrir sont bien identifiés, et qu'on peut éviter, limiter et réparer les risques qui lui sont associés (en particulier lors d'usages non pacifiques, malveillants et/ou aggravant les inégalités et des clivages). Ici, l'ONU, invite à ne pas utiliser l'IA quand elle met en péril la protection des données (tous les individus doivent pouvoir effacer et accéder aux enregistrements de leurs données personnelles, et les organismes de réglementation du monde entier doivent faire respecter ces dispositions). Cette recommandation vient aussi interdire la notation sociale et la surveillance de masse, contraires aux droits de l'homme et aux libertés fondamentales, et rejette l'idée d'accorder une personnalité juridique à l'IA « La Recommandation souligne que, lors de l'élaboration de cadres réglementaires, les États membres doivent tenir compte du fait que la responsabilité et l'obligation de rendre des comptes incombent toujours aux êtres humains en dernier ressort et que les technologies de l'IA ne devraient pas être dotées elles-mêmes d'une personnalité juridique ».

L'évaluation des IA prend en compte ses impacts éthiques sur les individus, sur la société et sur l'environnement. L'objectif étant à terme de créer une infrastructure juridique et technique *ad hoc*, ainsi qu'un responsable (indépendant) de l'éthique de l'IA pour surveiller l'utilisation et la création des IA qui devraient « privilégier les méthodes d'IA économes en données, en énergie et en ressources ». D'un point de vue écologique, les gouvernements sont invités, lors du cycle de vie du système d'IA, à analyser son « empreinte carbone, sa consommation d'énergie et l'impact environnemental de l'extraction des matières premières pour soutenir la fabrication des technologies d'IA », tout en cherchant à diminuer l'impact environnemental du numérique en investissant dans les technologies vertes. Ainsi, « si les systèmes d'IA ont un impact négatif disproportionné sur l'environnement, la Recommandation préconise de ne pas les utiliser »²²⁷.

En 2023, l'UNESCO réitère son appel à la mise en œuvre rapide de sa *Recommandation sur l'éthique de l'intelligence artificielle*, adoptée à l'unanimité par ses 193 États-membres²²⁷. « C'est le défi de notre temps », et il est « urgent que tous transposent ce cadre sous la forme de stratégies et de réglementations nationales. Nous devons traduire les engagements en actes » a commenté Audrey Azoulay (directrice générale de l'Unesco)²²⁸. L'ONU appelle ainsi les États qui ne l'ont pas déjà fait à rejoindre les plus de 40 pays « de toutes les régions du monde » qui ont commencé à créer de tels garde-fous, pour notamment créer un outil législatif capable d'encadrer et de surveiller les IA, tout en veillant à la protection des données personnelles et sensibles, et en sensibilisant la population mondiale à un usage responsable de l'IA²²⁸.

En 2023, lors de la 57^e Journée mondiale de la paix, le pape François exprime ses préoccupations quant aux conséquences potentielles de l'IA sur la paix mondiale et demande à la communauté internationale de définir un traité international contraignant pour réglementer son développement et son utilisation. Il se montre particulièrement préoccupé par « la possibilité de mener des opérations militaires à travers des systèmes de contrôle à distance », citant notamment les systèmes d'armes létales autonomes²²⁹.

Droits des robots

En 2017, le Parlement européen a demandé à une commission d'étudier la possibilité qu'un robot doté d'une intelligence artificielle puisse être considéré comme une personne juridique^{230, 231}. Advenant un dommage causé à un tiers par une intelligence artificielle, celle-ci pourrait être condamnée à réparer ce dommage. Il serait envisageable de conférer une personnalité électronique à tout robot prenant des décisions autonomes ou interagissant de manière indépendante avec des tiers, au même titre qu'une personne morale et physique.

Gouvernance démocratique

En 2023, Mark Coeckelbergh et Henrik Skaug Saetra, respectivement philosophe renommé de l'éthique des technologies et expert en sciences politiques, se penchent sur la question de l'intelligence artificielle (IA) et son potentiel rôle dans la lutte contre le changement climatique²³². Ils plaident pour une intégration des IA dans les politiques démocratiques, soulignant qu'elles peuvent faciliter la délibération et la prise de décisions. Toutefois, ils avertissent également que l'IA pourrait devenir un outil si puissant qu'il pourrait entièrement remplacer le gouvernement, entraînant des problèmes sociaux majeurs.

Pour illustrer leur propos, Coeckelbergh et Saetra présentent deux cas extrêmes : une démocratie sans remplacement des humains (*AI-augmented democracy*) et une technocratie dirigée par l'IA (*AI-driven technocracy*)²³². Ces deux cas extrêmes sont choisis pour démontrer les implications opposées de l'intégration de l'IA dans la gouvernance.

Dans une démocratie sans remplacement des humains, l'IA est présentée comme un simple outil visant à faciliter la prise de décision. Par exemple, elle peut être utilisée pour la traduction, la vérification des faits, ou la prise de notes. Ici, l'IA soutient le processus démocratique sans en prendre le contrôle.

En revanche, dans une démocratie avec remplacement, l'IA prend toutes les décisions, sans aucune intervention humaine. Cette approche, selon Saetra, pose cinq problèmes majeurs :

- les citoyens ont besoin d'une pleine participation politique pour être satisfaits ;
- les citoyens ne considèrent pas comme légitime un gouvernement auquel ils ne participent pas ;
- les ordinateurs ne devraient pas prendre de décisions affectant la vie et le bien-être des personnes ;
- l'IA n'est pas transparente et ne peut donc pas être entièrement contrôlée par l'homme ;
- l'obligation de rendre compte des conséquences des décisions politiques doit être claire, ce qui devient de moins en moins le cas lorsque l'IA prend des décisions.

Ces points soulignent les conséquences sociales et éthiques de la prise de décision par l'IA en ce qui concerne les humains. Ainsi, Coeckelbergh et Saetra concluent qu'une démocratie sans remplacement des humains est plus adaptée, l'IA y étant présente uniquement comme soutien et non comme entité décisionnelle.

Les deux auteurs estiment cependant qu'aucune des deux propositions n'est actuellement réalisable : les relations entre les humains et la technologie ne sont pas suffisamment évoluées pour permettre une utilisation éthique de l'IA. Ainsi, les décisions ne peuvent pas être prises uniquement par l'IA, car les erreurs sont encore trop fréquentes et les normes sociales et éthiques peu respectées²³².

Demandes de moratoire

Début 2023, l'apparition de ChatGPT suscite une grande curiosité, de l'enthousiasme, mais aussi des craintes sérieuses : « Devons-nous laisser les machines inonder nos canaux d'information de propagande et de mensonges ? (...) Devons-nous risquer de perdre le contrôle de notre civilisation ? Ces décisions ne doivent pas être déléguées à des leaders technologiques non élus » affirment Elon Musk, Steve Wozniak (cofondateur d'Apple) et des centaines d'experts. Le 29 mars 2023, ceux-ci, invoquant des « risques majeurs pour l'humanité », signent une pétition qui appelle le monde à un moratoire d'au moins six mois sur ces recherches, jusqu'à la mise en place de systèmes de sécurité, incluant : la création d'autorités réglementaires dédiées, des moyens pour efficacement surveiller des IA et des systèmes les utilisant, la mise à disposition de techniques permettant de mieux différencier le réel de l'artificiel, et la création d'institutions pouvant limiter les « perturbations économiques et politiques dramatiques (en particulier pour la démocratie) que l'IA provoquera »²²⁸.

Philosophie

Courants de pensée

La cybernétique naissante des années 1940 revendiquait très clairement son caractère pluridisciplinaire et se nourrissait des contributions les plus diverses : neurophysiologie, psychologie, logique, sciences sociales, etc. Elle envisagea deux approches des systèmes, approches reprises par les sciences cognitives et de ce fait l'intelligence artificielle^[réf. souhaitée] : une approche par la décomposition (du haut vers le bas, comme avec les systèmes experts) et une approche contraire par construction progressive du bas vers le haut, comme avec l'apprentissage automatique²³³.

Ces deux approches se révèlent plutôt complémentaires que contradictoires : on est à l'aise pour décomposer rapidement ce que l'on connaît bien, et une approche pragmatique à partir des seuls éléments que l'on connaît afin de se familiariser avec les concepts émergents est plus utile pour les domaines inconnus. Elles sont respectivement à la base des hypothèses de travail que constituent le cognitivisme et le connexionnisme, qui tentent aujourd'hui (2005)^[Passage à actualiser] d'opérer progressivement leur fusion.

Le guide pratique de Linux sur l'intelligence artificielle v3.0²³⁴, révisé le 15 décembre 2012, adopte pour la commodité du lecteur la taxinomie suivante :

- systèmes symboliques ;
- connexionnisme ;
- calcul évolutif (algorithmes génétiques, par exemple) ;
- alife (vie artificielle) et complexité ;
- agents et robotique.

Cognitivisme

Le cognitivisme considère que le vivant, tel un ordinateur (bien que par des procédés très différents), manipule essentiellement des symboles élémentaires. Dans son livre *La société de l'esprit*, Marvin Minsky, s'appuyant sur des observations du psychologue Jean Piaget, envisage le processus cognitif comme une compétition d'agents fournissant des réponses partielles et dont les avis sont arbitrés par d'autres agents. Il cite les exemples suivants de Piaget :

- l'enfant croit d'abord que plus le niveau d'eau est élevé dans un verre, plus il y a d'eau dans ce verre. Après avoir joué avec des transvasements successifs, il intègre le fait que la notion de hauteur du liquide dans le verre entre en compétition avec celle du diamètre du verre, et arbitre de son mieux entre les deux ;

- il vit ensuite une expérience analogue en manipulant de la pâte à modeler : la réduction de plusieurs objets temporairement représentés à une même boule de pâte l'incite à dégager un concept de conservation de la quantité de matière.

Au bout du compte, ces jeux d'enfants se révèlent essentiels à la formation de l'esprit, qui dégagent quelques règles pour arbitrer les différents éléments d'appréciation qu'il rencontre, par essais et erreurs.

Connexionnisme

Le connexionnisme, se référant aux processus auto-organisationnels, envisage la cognition comme le résultat d'une interaction globale des parties élémentaires d'un système. On ne peut nier que le chien dispose d'une sorte de *connaissance* des équations différentielles du mouvement, puisqu'il arrive à attraper un bâton au vol, ni qu'un chat ait aussi une sorte de *connaissance* de la loi de chute des corps, puisqu'il se comporte comme s'il savait à partir de quelle hauteur il ne doit plus essayer de sauter directement pour se diriger vers le sol. Cette faculté, qui évoque l'intuition des philosophes, se caractériserait par la prise en compte et la consolidation d'éléments perceptifs dont aucun pris isolément n'atteint le seuil de la conscience, ou en tout cas n'y déclenche d'interprétation particulière.

Synthèse

Trois concepts reviennent de façon récurrente dans la plupart des travaux :

- la redondance (le système est peu sensible à des pannes ponctuelles) ;
- la réentrance (les composants s'informent en permanence entre eux ; cette notion diffère de la réentrance en programmation) ;
- la sélection (au fil du temps, les comportements efficaces sont dégagés et renforcés).

Intelligence artificielle forte

Le concept d'intelligence artificielle forte fait référence à une machine capable non seulement de produire un comportement intelligent, notamment de modéliser des idées abstraites, mais aussi d'éprouver une impression d'une réelle conscience, de « vrais sentiments » (notion dont la définition n'est pas universelle), et « une compréhension de ses propres raisonnements »²³⁵.

Contrairement à l'intelligence artificielle générale, l'intelligence artificielle forte fait donc le plus souvent intervenir des notions philosophiques de conscience qui font que les capacités de l'intelligence artificielle ne suffisent pas à dire si elle est « forte ». Cela dit, aucune définition de la conscience pour une IA ne fait consensus²³⁶. Les termes « intelligence artificielle forte » et « intelligence artificielle générale » sont parfois en pratique utilisés de manière interchangeable⁶¹.

En partant du principe, étayé par les neurosciences²³⁷, que la conscience a un support biologique et donc matériel, les scientifiques ne voient généralement pas d'obstacle théorique à la création d'une intelligence consciente sur un support matériel autre que biologique. Selon les tenants de l'IA forte, si à l'heure actuelle il n'y a pas d'ordinateurs ou d'algorithmes aussi intelligents que l'être humain, ce n'est pas un problème d'outil mais de conception. Il n'y aurait aucune limite fonctionnelle (un ordinateur est une machine de Turing universelle avec pour seules limites celles de la calculabilité), seulement des limites liées à l'aptitude humaine à concevoir les logiciels appropriés (programme, base de données...).

Diversité des opinions

Les principales opinions soutenues pour répondre à la question d'une intelligence artificielle forte (c'est-à-dire douée d'une sorte de conscience) sont les suivantes :

- impossible : la conscience serait le propre des organismes vivants (supérieurs), et elle serait liée à la nature des systèmes biologiques. Cette position est défendue par certains philosophes et sociologues comme Harry Collins, pour qui l'intelligence requiert une immersion dans la société humaine, et donc un corps humain⁶³, et peut rappeler le courant du vitalisme ;
- impossible avec des machines manipulant des symboles comme les ordinateurs actuels, mais possible avec des systèmes dont l'organisation matérielle serait fondée sur des processus quantiques. Des algorithmes quantiques sont théoriquement capables de mener à bien des calculs hors de l'atteinte *pratique* des calculateurs conventionnels (complexité en N^3 au lieu de 2^N , par exemple, sous réserve d'existence du calculateur approprié). Au-delà de la rapidité, le scientifique Roger Penrose défend dans la théorie de la réduction objective orchestrée l'idée controversée que la conscience nécessiterait un fonctionnement non compatible avec les lois de la physique classique, et accessible uniquement à des systèmes quantiques^{238, 239} ;

- impossible car la pensée n'est pas un phénomène calculable par des processus discrets et finis. Cette théorie est notamment avancée par le philosophe John Searle et son expérience de la chambre chinoise²⁴⁰. Une conscience est donc nécessaire pour accéder à l'intelligence, mais un système informatique ne serait capable que d'en simuler une, sans pour autant la posséder, renvoyant au concept philosophique du zombie ;
- possible avec des ordinateurs manipulant des symboles. La notion de symbole est toutefois à prendre au sens large. Cette option inclut les travaux sur le raisonnement ou l'apprentissage symbolique basé sur la logique des prédicats, mais aussi les techniques connexionnistes telles que les réseaux de neurones, qui, à la base, sont définies par des symboles. Cette position est portée par des mouvements comme ceux du computationalisme et est portée par des philosophes comme Hubert Dreyfus, pour qui le cerveau suit les lois de la physique et de la biologie, impliquant que l'esprit est donc un processus simulable²⁴¹. Cette dernière opinion constitue la position la plus engagée en faveur de l'intelligence artificielle forte.

Des auteurs comme Douglas Hofstadter (mais déjà avant lui Arthur C. Clarke ou Alan Turing ; voir le test de Turing) expriment par ailleurs un doute sur la possibilité de faire la différence entre une intelligence artificielle qui éprouverait réellement une conscience, et une autre qui simulerait exactement ce comportement (voir Zombie (philosophie)). Après tout, nous ne pouvons même pas être certains que d'autres consciences que la nôtre, y compris chez des humains, éprouvent réellement quoi que ce soit, si ce n'est par une pétition de principe qui spécule que chaque humain se retrouve à l'identique chez tous les autres. On retrouve là le problème connu du solipsisme en philosophie.

Même si une intelligence artificielle forte n'était guère possible, une IA peut être de plus en plus perçue comme forte par une majorité d'individus parallèlement à l'arrivée des IA génératives, dont les LLM (de l'anglais *large language model*) comme ChatGPT ou Google Bard, et les outils de génération d'images comme Midjourney, DALL-E ou Stable Diffusion. En effet, le champ d'applications de ces outils est beaucoup plus large qu'auparavant : création, synthèse, traduction de textes, composition d'images, de vidéos à partir de *prompts*, textes descriptifs. Il devient ainsi de plus en plus difficile pour un être humain de distinguer des créations humaines de celles provenant d'une IA générative.

Emily Bender estime que les grands modèles de langage comme ChatGPT ne font que régurgiter plus ou moins aléatoirement des morceaux de texte venant des corpus ayant servi à leur entraînement, sans en comprendre le sens. Elle les appelle ainsi des « perroquets stochastiques »²⁴². De même, Jean-Gabriel Ganascia considère que le contenu qu'ils produisent n'est pas original et que leur utilisation dans la rédaction d'articles de recherche constitue une forme de plagiat²⁴³. Ilya Sutskever considère au contraire que ces modèles, à force d'être entraînés à prédire le mot suivant, acquièrent une forme de « modèle du monde » et une représentation « compressée, abstraite et utilisable » des concepts²⁴⁴.

Intelligence artificielle faible

La notion d'intelligence artificielle faible constitue une approche pragmatique d'ingénieur : chercher à construire des systèmes de plus en plus autonomes (pour réduire le coût de leur supervision), des algorithmes capables de résoudre des problèmes d'une certaine classe, etc. Mais, cette fois, la machine simule l'intelligence, elle semble agir comme si elle était intelligente.

Les tenants de l'IA forte admettent que s'il y a bien dans ce cas simple simulation de comportements intelligents, il est aisé de le découvrir et qu'on ne peut donc généraliser. En effet, si on ne peut différencier expérimentalement deux comportements intelligents, celui d'une machine et celui d'un humain, comment peut-on prétendre que les deux choses ont des propriétés différentes ? Le terme même de « simulation de l'intelligence » est contesté et devrait, toujours selon eux, être remplacé par « reproduction de l'intelligence ».

Si le terme *intelligence artificielle* peut désigner un système capable de résoudre plusieurs problèmes de façon relativement autonome tout en ne faisant que simuler le principe d'intelligence, il peut aussi désigner des systèmes capables de résoudre uniquement un type de problème pour un jeu de données prédéfini²⁴⁵. On peut donner pour exemple un système entraîné à reconnaître des chiffres écrits à la main, comme ceux utilisés par La Poste²⁴⁶, qui malgré sa grande performance sur sa tâche, serait incapable de fonctionner sur un problème sortant de ce pour quoi il a été conçu. Ces intelligences artificielles, aussi nommées « intelligences artificielles étroites » (terme issu de l'anglais *narrow AI*), sont conçus pour effectuer une tâche précise, contrairement à une intelligence artificielle générale²⁴⁷.

Question de l'intelligence

La définition du terme « intelligence artificielle » pose une question fondamentale : Qu'est-ce que l'intelligence²⁴⁸ ?

L'intelligence peut se définir de manière générale comme un ensemble de capacités cognitives permettant de résoudre des problèmes ou de s'adapter à un environnement²⁴⁹.

Le chercheur en IA Yann Le Cun avance que le noyau de l'intelligence est la faculté de *prédire*. Les bases de la programmation des premiers systèmes experts supposent de « maîtriser parfaitement un problème et d'avoir une vue précise de toutes les solutions », afin d'en programmer précisément le comportement²⁴⁸. Mais les systèmes d'IA modernes à base d'apprentissage automatique sont entraînés à prédire la réponse attendue, ou à générer une solution correcte. Leurs capacités émergent ainsi progressivement, par essai-erreur, sans que le programmeur n'ait besoin de fournir une solution algorithmique, ou même de savoir comment résoudre le problème²³³. Dans tous les cas, l'efficacité de l'intelligence artificielle dépend de sa capacité à répondre aux objectifs donnés par les programmeurs et à tendre vers l'autonomie décisionnelle²⁵⁰, ce qui présuppose, entre autres, une capacité de prédiction.

Le philosophe John Searle considère quant à lui que la faculté de *comprendre* est plus importante dans la définition de l'intelligence. Il essaie de démontrer la faiblesse des systèmes d'intelligence artificielle et les limites du test de Turing, par son expérience de la chambre chinoise, concluant : « on ne devrait pas dire d'une IA qu'elle comprend les informations qu'elle traite lorsqu'elle manipule des règles de syntaxe sans maîtriser la sémantique, c'est-à-dire sans reconnaître le sens des mots. La question de savoir si on peut parler d'une véritable intelligence reste donc ouverte »²⁴⁸. L'apprentissage automatique fonctionne cependant différemment de l'IA symbolique²³³, qui était populaire à l'époque où Searle a conçu l'expérience de pensée de la chambre chinoise en 1980²⁵¹.

Dans la fiction

Une machine ayant une conscience et capable d'éprouver des sentiments — ou de faire comme si c'était le cas — est un grand thème classique de la science-fiction, notamment des romans d'Isaac Asimov sur les robots²⁵².

Ce sujet a toutefois été exploité très tôt, comme dans le récit des aventures de Pinocchio, publié en 1881, où une marionnette capable d'éprouver de l'amour pour son créateur cherche à devenir un vrai petit garçon, ou dans *L'Homme le plus doué du monde*, une nouvelle de l'Américain Edward Page Mitchell où le cerveau d'un simple d'esprit est remplacé par un ordinateur inspiré des recherches de Charles Babbage²⁵³. Le roman *Le Miroir flexible* de Régis Messac propose quant à lui le principe d'une intelligence artificielle faible, mais évolutive, avec des automates inspirés de formes de vie simples, réagissant à certains stimuli tels que la lumière. Cette trame a fortement inspiré le film *A.I. Intelligence artificielle* réalisé par Steven Spielberg, sur la base d'idées de Stanley Kubrick, lui-même inspiré de Brian Aldiss²⁵⁴. L'œuvre de Dan Simmons, notamment le cycle d'*Hypérion*, évoque l'intelligence artificielle. *Destination vide*, de Frank Herbert, met en scène de manière fascinante l'émergence d'une intelligence artificielle forte. Plus récemment, l'écrivain français Christian Léourier a placé une intelligence artificielle au cœur de son roman court *Helstrid* (2018), dans lequel cette IA laisse un être humain mourir, contrevenant ainsi aux trois lois de la robotique instaurées par Isaac Asimov près de quatre-vingts ans plus tôt.



HAL 9000.

Les androïdes faisant preuve d'intelligence artificielle dans la fiction sont nombreux : le personnage de Data de la série télévisée *Star Trek : The Next Generation* est un être cybernétique doué d'intelligence, avec des capacités importantes d'apprentissage. Il est officier supérieur sur le vaisseau *Enterprise* et évolue aux côtés de ses coéquipiers humains qui l'inspirent dans sa quête d'humanité. Son pendant cinématographique est Bishop dans les films *Aliens* (1986) et *Alien 3* (1992). Dans le manga *Ghost in the Shell*, une androïde s'éveille à la conscience. Dans la saga *Terminator* avec Arnold Schwarzenegger, le T-800 reprogrammé, conçu initialement pour tuer, semble dans la capacité d'éprouver des sentiments humains. Par ailleurs, les Terminators successifs sont envoyés dans le passé par Skynet, une intelligence artificielle qui a pris conscience d'elle-même, et du danger que représentent les humains envers elle-même^[réf. nécessaire].

Dans le dernier épisode de la saison 10 d'Inspecteur Derrick (diffusé en 1983), intitulé *Un homme en trop*, le Professeur Römer (joué par Erich Hallhuber) veut arrêter ses recherches sur l'intelligence artificielle, par peur du pouvoir sans conscience que cette nouvelle génération d'ordinateurs pourrait avoir sur les humains en finissant par les dominer et les tuer. Il abat son successeur pour l'empêcher de continuer son travail^{255, 256, c}.

Quelques IA célèbres dans la science-fiction

- 1968 : *2001, l'Odyssée de l'espace* de Stanley Kubrick, inspiré de la nouvelle *La Sentinelle* d'Arthur C. Clarke, également auteur du scénario du film, décrit notamment la lutte entre l'ordinateur HAL et l'humain Dave.
- 1969 : *Le Cerveau d'acier*, d'après le roman de Dennis Feltham Jones (en) de 1967, montre un système d'IA militaire américain contacter son homologue russe pour qu'ils coopèrent à leur mission commune, éviter la guerre nucléaire (en neutralisant les humains).
- 1981 : dans *Blade Runner* de Ridley Scott, inspiré d'un roman de Philip K. Dick, des humains artificiels (des « répliquants ») reviennent sur terre après une mission spatiale, mais n'acceptent pas leur mort programmée à la suite du succès de leur mission.
- 1982 : dans *K 2000*, une Pontiac Trans-Am embarque une intelligence artificielle au nom de KITT, conçue pour réaliser des diagnostics pour les dossiers de la F.L.A.G. Elle ne peut ressentir des émotions, ne peut porter atteinte à la vie d'autrui et n'écoute que les ordres de Michael Knight. Une partie de l'histoire repose sur l'importance de la programmation, notamment à travers son jumeau maléfique et prototype K.A.R.R. dont la principale différence est d'avoir un instinct de survie qui prédomine.
- 1982 : *Tron* de Steven Lisberger, où le Maître contrôle principal (MCP) est un programme d'échecs qui a évolué en IA et tente de prendre le contrôle total du système.
- 1985 : *D.A.R.Y.L.* est un androïde que le gouvernement américain cherche à détruire.
- 1999 : *Matrix*, trilogie cinématographique de science-fiction dans laquelle les humains, enfermés dans un monde créé par l'IA, sont asservis par les machines. Une petite poche de résistance humaine résiste et se bat encore dans l'espoir de la venue de l'Élu : Neo.
- 1999 : *L'Homme bicentenaire*, où un exemplaire de NDR-114 dénommé « Andrew » poursuit un long chemin vers l'acquisition de la conscience, au point de se voir reconnaître le statut d'être humain à la fin de sa « vie ». Il s'agit d'une adaptation de la nouvelle éponyme d'Isaac Asimov.
- 2001 : *A.I. Intelligence artificielle* de Steven Spielberg, inspiré de la nouvelle de Brian Aldiss *Les Supertoys durent tout l'été*. Le personnage central est un enfant-robot doué d'émotions et de sentiments.
- 2003 - 2007 : *Code Lyoko* dessin animé où une I.A. appelée X.A.N.A. tente de prendre le contrôle du réseau mondial après avoir accédé à la conscience.
- 2004 : *I, Robot*, inspiré de l'œuvre de Isaac Asimov et thème semblable au film *AI*.
- 2008 : J.A.R.V.I.S. (*Just A Rather Very Intelligent System*) dans les films *Iron Man*, *Avengers*, etc. avec Robert Downey Jr., inspiré des comics de Marvel.
- 2011-2016 : la série télévisée *Person of Interest* met en scène un groupe de personnes guidées par une intelligence artificielle capable de prédire des crimes.
- 2012-2014 : la série télévisée *Real Humans : 100 % humain* décrit l'émergence de robots doués de conscience au sein de la société humaine.
- 2015 : *Ex Machina* de Alex Garland, dans lequel un test de Turing d'une semaine va dégénérer en faveur d'un robot féminin (gynοïde) révolutionnaire.
- 2016 : la série télévisée *Westworld* met en scène des androïdes à l'apparence humaine qui commencent à adopter des comportements imprévisibles²⁵⁷.
- 2023 : dans la mini-série *Class of '09*, l'IA est utilisée par le FBI pour résoudre des enquêtes.

Utilisation dans les jeux

Les jeux, notamment les jeux de stratégie, ont marqué l'histoire de l'intelligence artificielle, même s'ils ne mesurent que des compétences particulières, telles que la capacité de la machine en matière de calcul de probabilités, de prise de décision mais aussi d'apprentissage.

Hans Berliner (1929-2017), docteur en science informatique à l'université Carnegie-Mellon et joueur d'échecs, fut l'un des pionniers de la programmation pour les ordinateurs de jeu. Ses travaux commencèrent par un programme capable de battre un humain professionnel au backgammon, puis, à partir des années 1960 et avec l'aide d'IBM, il fit des recherches pour créer un programme capable de rivaliser avec des grands maîtres du jeu d'échecs. Ses travaux contribuèrent quelques décennies plus tard à la réalisation du supercalculateur Deep Blue²⁵⁸.

Outre la capacité des jeux à permettre de mesurer les performances de l'intelligence artificielle, que ce soit au travers d'un score ou d'un affrontement face à un humain, les jeux offrent un environnement propice à l'expérimentation pour les chercheurs, notamment dans le domaine de l'apprentissage par renforcement²⁵⁹.

Othello

Dans le jeu Othello, sur un plateau de 8 cases sur 8, chaque joueur place tour à tour des pions de sa couleur (noir ou blanc). Le vainqueur est celui qui possède les pions de la couleur dominante.

L'une des premières intelligences artificielles pour l'Othello est IAGO, développée en 1976 par l'université Caltech de Pasadena (Californie), qui bat sans difficultés le champion japonais Fumio Fujita.

Le premier tournoi d'Othello opposant des hommes à des machines est organisé en 1980. Un an plus tard, un nouveau tournoi de programmes regroupe 20 systèmes²⁶⁰. C'est entre 1996 et 1997 que le nombre de programmes explose : Darwersi (1996-1999) par Olivier Arzac, Hannibal (1996) par Martin Piotte et Louis Geoffroy, Keyano (1997) par Mark Brockington, Logistello (1997) par Michael Buro, etc.

Échecs

En 1968, le maître international anglais David Levy lança un défi à des spécialistes en intelligence artificielle, leur pariant qu'aucun programme informatique ne serait capable de le battre aux échecs dans les dix années à venir. Il remporta son pari, n'étant finalement battu par Deep Thought qu'en 1989²⁶¹.

En 1988, l'ordinateur HiTech de Hans Berliner est le premier programme à battre un grand maître du jeu d'échecs, Arnold Denker (74 ans) en match (3,5-1,5)^{262, d}.

En 1997, le supercalculateur conçu par IBM, Deep Blue (surnommé *Deeper Blue* lors de ce match revanche), bat Garry Kasparov (3,5-2,5) et marque un tournant : pour la première fois, le meilleur joueur humain du jeu d'échecs est battu en match (et non lors d'une partie unique) par une machine.

En décembre 2017, une version généraliste d'AlphaGo Zero (le successeur du programme AlphaGo de DeepMind^e) nommée AlphaZero, est développée pour jouer à n'importe quel jeu en connaissant seulement les règles, et en apprenant à jouer seul contre lui-même. Ce programme est ensuite entraîné pour le go, le shogi et les échecs. Après 9 heures d'entraînement, AlphaZero bat le programme d'échecs Stockfish (leader dans son domaine), avec un score de 28 victoires, 72 nulles et aucune défaite. Il faut cependant noter que la puissance de calcul disponible pour AlphaZero (4 TPU v2 pour jouer, soit une puissance de calcul de 720 Teraflops) était très supérieure à la puissance disponible de Stockfish pour ce match, ce dernier tournant sur un ordinateur équipé de seulement 64 cœurs Intel²⁶³. AlphaZero a également battu (après apprentissage) le programme de shōgi Elmo (en)^{264, 265}.



Un supercalculateur IBM similaire à Deep Blue, qui a battu le champion du monde d'échecs en titre dans un match en 1997.

Go

En 2015, l'IA réalise des progrès significatifs dans la pratique du go, plus complexe à appréhender que les échecs (entre autres à cause du plus grand nombre de positions : 10^{170} au go, contre 10^{50} pour les échecs, et de parties plausibles : 10^{600} au go, contre 10^{120} pour les échecs)²⁶⁶.

En octobre 2015, AlphaGo, un logiciel d'IA conçu par DeepMind, filiale de Google, bat pour la première fois Fan Hui, le triple champion européen de go²⁶⁷ et ainsi relève ce qu'on considérait comme l'un des plus grands défis pour l'intelligence artificielle. Cette tendance se confirme en mars 2016 quand AlphaGo bat par trois fois consécutives le champion du monde de la discipline, Lee Sedol, dans un duel en cinq parties²⁶⁸. Lee Sedol a déclaré au terme de la seconde partie qu'il n'avait trouvé « aucune faiblesse » chez l'ordinateur et que sa défaite était « sans équivoque ».

Jeopardy!

En 2011, l'IA Watson conçue par IBM bat ses adversaires humains au jeu télévisé américain *Jeopardy!*. Dans ce jeu de questions/réponses, la compréhension du langage est essentielle pour la machine ; pour ce faire, Watson a pu s'appuyer sur une importante base de données interne lui fournissant des éléments de culture générale, et avait la capacité d'apprendre par lui-même, notamment de ses erreurs. Il disposait néanmoins d'un avantage, la capacité d'appuyer instantanément (et donc avant ses adversaires humains) sur le buzzer pour donner une réponse²⁶⁶.



Réplique de Watson, lors d'un concours de *Jeopardy!*

Poker

En 2007, Polaris est le premier programme informatique à gagner un tournoi de poker significatif face à des joueurs professionnels humains^{269, 270}.

En 2017, lors du tournoi de poker « *Brains Vs. Artificial Intelligence : Upping the Ante* » (« Cerveau contre Intelligence Artificielle : on monte la mise ») organisé dans un casino de Pennsylvanie, l'intelligence artificielle Libratus, développée par des chercheurs de l'université Carnegie-Mellon de Pittsburgh, est confrontée à des adversaires humains dans le cadre d'une partie marathon étalée sur 20 jours²⁷⁰. Les joueurs humains opposés à Libratus, tous professionnels de poker, affrontent successivement la machine dans une partie en face à face (*heads up (en)*) selon les règles du « *No Limit Texas hold'em* » (*no limit* signifiant que les mises ne sont pas plafonnées), la version alors la plus courante du poker. Les parties sont retransmises en direct et durant huit heures par jour sur la plateforme Twitch²⁷¹.

Au terme de plus de 120 000 mains jouées, Libratus remporte tous ses duels face aux joueurs humains et accumule 1 766 250 dollars (virtuels). Le joueur humain ayant perdu le moins d'argent dans son duel face à la machine, Dong Kim, est tout de même en déficit de plus de 85 000 dollars. Dans leurs commentaires du jeu de leur adversaire, les joueurs humains admettent que celui-ci était à la fois déconcertant et terriblement efficace. En effet, Libratus « étudiait » chaque nuit, grâce aux ressources d'un supercalculateur situé à Pittsburgh, ses mains jouées durant la journée écoulée, utilisant les 15 millions d'heures-processeur de calculs du supercalculateur²⁷¹.

La victoire, nette et sans bavure, illustre les progrès accomplis dans le traitement par l'IA des « informations imparfaites », où la réflexion doit prendre en compte des données incomplètes ou dissimulées. Les estimations du nombre de possibilités d'une partie de poker sont en effet d'environ 10^{160} dans la variante *no limit* en face à face²⁷¹.

Auparavant, en 2015, le joueur professionnel Doug Polk (*en*) avait remporté la première édition de cet événement contre une autre IA, baptisée Claudico (*en*)²⁷¹.

Bridge

En mars 2022, un logiciel de bridge de la start-up française Nukkai parvient à gagner un tournoi et à expliquer aux perdants leurs erreurs²⁷².

Notes et références

- (*en*) Cet article est partiellement ou en totalité issu de l'article de Wikipédia en anglais intitulé « *Artificial intelligence* » (https://en.wikipedia.org/wiki/Artificial_intelligence?oldid=1194119617) » (voir la liste des auteurs (https://en.wikipedia.org/wiki/Artificial_intelligence?action=history)).

Notes

- (*en*) « *the building of computer programs which perform tasks which are, for the moment, performed in a more*

satisfactory way by humans because they require high level mental processes such as: perception learning, memory organization and critical reasoning ».

- b. DarkBert a été initialement conçu comme un outil de lutte contre le cybercrime.
- c. Dans cet épisode, l'échange entre Derrick et le Pr Rotheim (Ernst Schröder) vers 17'25" : « Le Dr Römer doutait sur le bien fondé des recherches sur l'intelligence artificielle », puis à 17'54" « On fait une caricature de l'intelligence qui n'a pas de morale », puis vers 29'25", madame Römer échange avec Derrick : « Il craignait que [...] L'intelligence artificielle présente de trop gros risques. », consultable sur YouTube ou le DVD *Inspecteur Derrick - Intégrale saison 10*.
- d. Arnold Denker était alors âgé de 74 ans et crédité d'un classement Elo de 2300, ce qui relativise un peu la performance du programme, un fort grand maître étant à cette époque plus vers les 2 650–2 700 points Elo, voire davantage.
- e. Voir plus bas dans la section « Go ».

Références

1. ministère de l'Enseignement supérieur et de la Recherche, « France intelligence artificielle (https://web.archive.org/web/20170329090007/https://cache.media.enseignementsup-recherche.gouv.fr/file/Actus/86/1/Conclusions_Groupes_Travail_France_IA_738861.pdf) » [PDF], 21 mars 2017 (version du 29 mars 2017 sur *Internet Archive*).
2. « Les applications les plus populaires de l'Intelligence Artificielle (IA) (<https://pandia.pro/guide/les-applications-les-plus-populaires-de-lia/>) », sur *Pandia*, 15 mars 2023 (consulté le 27 mai 2023).
3. (en) « Rulers of the world, unite! The challenges and opportunities of artificial intelligence (<https://www.sciencedirect.com/science/article/pii/S0007681319301260>) », sur *sciencedirect.com*, janvier 2020.
4. « Qu'est-ce que l'intelligence artificielle (IA) ? (<https://www.ibm.com/fr-fr/topics/artificial-intelligence>) », sur *IBM* (consulté le 6 avril 2023).
5. « Intelligence artificielle (http://www.larousse.fr/encyclopedie/divers/intelligence_artificielle/187257) », sur *Larousse* (consulté le 15 août 2020).
6. (en) Russell, S. et Norvig, P., *Artificial Intelligence: A Modern Approach (2nd ed.)*, Prentice Hall, 2003, 932 p. (ISBN 0-13-790395-2), section 1.1.
7. « Intelligence artificielle, machine learning, deep learning : kékako ? (<https://www.ledigitalab.com/fabrique/intelligence-artificielle-machine-learning-deep-learning-kekako/>) », sur *ledigitalab.com*, 2 octobre 2017.
8. (en) Centre commun de recherche, Sofia Samoil, Montserrat López Cobo et Blagoj Delipetrev, *AI watch, defining artificial intelligence 2.0 : towards an operational definition and taxonomy for the AI landscape*, Office des publications de l'Union européenne, 2021, 125 p. (ISBN 978-92-76-42648-6, ISSN 1831-9424 (<https://portal.issn.org/resource/issn/1831-9424>), DOI 10.2760/019901 (<https://dx.doi.org/10.2760/019901>)).
9. (en) « A definition of Artificial Intelligence: main capabilities and scientific disciplines (<https://digital-strategy.ec.europa.eu/en/library/definition-artificial-intelligence-main-capabilities-and-scientific-disciplines>) », *Shaping Europe's digital future*, 18 décembre 2018 (consulté le 22 août 2024).
10. Kathleen Desveaud, *L'Intelligence Artificielle décryptée : Comprendre les enjeux et risques éthiques de l'IA pour mieux l'appréhender*, EMS, 2024, 274 p. (ISBN 2376879604).
11. Mark Esposito, Kariappa Bheemaiah et Terence Tse, « Vous avez dit « machine learning » ? Quand l'ordinateur apprend à apprendre (<https://theconversation.com/vous-avez-dit-machine-learning-quand-lordinateur-apprend-a-apprendre-76049>) », sur *theconversation.com*, 1^{er} mai 2017 (consulté le 2 janvier 2024).
12. Daniel Nelson, « Qu'est-ce que la rétropropagation ? (<https://www.unite.ai/fr/qu'est-ce-que-la-r%C3%A9tropropagation/>) », sur *unite.ai*, 23 août 2020 (consulté le 2 janvier 2024).
13. (en) Nick Heath, « What is machine learning? Everything you need to know (<https://www.zdnet.com/article/what-is-machine-learning-everything-you-need-to-know/>) », sur *ZDNet* (consulté le 2 janvier 2024).
14. Émilie Dedieu, « Comment l'apprentissage auto-supervisé accélère l'intelligence artificielle », *L'Usine nouvelle*, 12 avril 2022 (lire en ligne (<https://www.usinenouvelle.com/article/comment-l-apprentissage-auto-supervise-accele-re-l-intelligence-artificielle.N1991622>), consulté le 7 janvier 2024).
15. Antoine Crochet-Damais, « Réseau de neurones artificiels : réseaux neuronaux pour l'IA (<https://www.journaldunet.fr/intelligence-artificielle/guide-de-l-intelligence-artificielle/1501851-reseau-de-neurones-artificiels/>) », sur *Le Journal du Net*, 29 juin 2022 (consulté le 3 janvier 2024).
16. « Qu'est-ce qu'un réseau de neurones convolutifs ? (<https://www.ibm.com/fr-fr/topics/convolutional-neural-networks>) », sur *IBM* (consulté le 7 janvier 2024).
17. (en) Ram Sagar, « Can We Speed Up Matrix Multiplication? (<https://analyticsindiamag.com/can-we-speed-up-matrix-multiplication/>) », sur *Analytics India Magazine*, 14 septembre 2021 (consulté le 7 janvier 2024).
18. (en) Kurt Hornik, Maxwell Stinchcombe et Halbert White, « Multilayer Feedforward Networks are Universal Approximators », *Neural Networks*, 1989 (lire en ligne (https://cognitivemedium.com/magic_paper/assets/Hornik.pdf) [PDF]).
19. « Deux types de neurones artificiels pour émuler le cerveau (<https://theconversation.com/deux-types-de-neurones-artificiels-pour-emuler-le-cerveau-151887>) », sur *The Conversation*, 7 janvier 2021 (consulté le 8 février 2025).

20. Stephane Nachez, « Qu'est-ce que le *Deep Learning* ? (<https://www.actuia.com/actualite/quest-deep-learning/>) », sur *Actuia*, 5 février 2018 (consulté le 3 janvier 2024).
21. (en) Sharon Goldman, « 10 years later, deep learning 'revolution' rages on, say AI pioneers Hinton, LeCun and Li (<https://venturebeat.com/ai/10-years-on-ai-pioneers-hinton-lecun-li-say-deep-learning-revolution-will-continue/>) », sur *VentureBeat*, 14 septembre 2022 (consulté le 6 janvier 2024).
22. (en) Sean Michael Kerner, « What are Large Language Models? (<https://www.techtarget.com/whatis/definition/large-language-model-LLM>) », sur *TechTarget* (consulté le 2 janvier 2024).
23. (en) Craig S. Smith, « ChatGPT-4 Creator Ilya Sutskever on AI Hallucinations and AI Democracy (<https://www.forbes.com/sites/craigsmith/2023/03/15/gpt-4-creator-ilya-sutskever-on-ai-hallucinations-and-ai-democracy/>) », sur *Forbes*, 15 mars 2023 (consulté le 2 janvier 2024).
24. (en) Aayush Mittal, « NLP Rise with Transformer Models: A Comprehensive Analysis of T5, BERT, and GPT (<http://www.unite.ai/nlp-rise-with-transformer-models-a-comprehensive-analysis-of-t5-bert-and-gpt/>) », sur *Unite.ai*, 8 novembre 2023 (consulté le 7 janvier 2024).
25. « Microsoft annonce l'intégration de la technologie GPT-4 dans ses outils Word, Excel, Outlook et Teams », *Le Monde*, 16 mars 2023 (lire en ligne (https://www.lemonde.fr/pixels/article/2023/03/16/microsoft-annonce-l-integration-de-la-technologie-gpt-4-dans-ses-outils-word-excel-outlook-et-teams_6165789_4408996.html)), consulté le 3 janvier 2024).
26. (en) Beatrice Nolan, « Here's what we know so far about Google's Gemini (<https://www.businessinsider.com/google-gemini-explainer-ai-model-2023-9>) », sur *Business Insider*, 6 décembre 2023 (consulté le 2 janvier 2024).
27. Russell et Norvig 2021, chapitre 4.
28. Russell et Norvig 2021, chapitre 3.
29. Russell et Norvig 2021, chapitre 5.
30. Russell et Norvig 2021, chapitre 3.5.
31. Russell et Norvig 2021, chapitres 6 à 10.
32. (en) « What is 'fuzzy logic'? Are there computers that are inherently fuzzy and do not apply the usual binary logic? (<https://www.scientificamerican.com/article/what-is-fuzzy-logic-are-t/>) », sur *Scientific American* (consulté le 7 janvier 2024).
33. Russell et Norvig 2021, chapitre 13.
34. Russell et Norvig 2021.
35. « Qu'est-ce que la simulation de Monte-Carlo ? (<https://www.ibm.com/fr-fr/topics/monte-carlo-simulation>) », sur *IBM* (consulté le 7 janvier 2024).
36. (en) Vijaysinh Lendave, « A Beginners' Guide to Cross-Entropy in Machine Learning (<https://analyticsindiamag.com/a-beginners-guide-to-cross-entropy-in-machine-learning/>) », sur *Analytics India Magazine*, 11 septembre 2021 (consulté le 7 janvier 2024).
37. Cassidy Kelley et Gaétan Raoul, « Machine Learning : les 9 types d'algorithmes les plus pertinents en entreprise (<https://www.lemagit.fr/conseil/Machine-Learning-les-9-types-dalgorithmes-les-plus-pertinents-en-entreprise>) », sur *LeMagIT*, 8 juin 2020 (consulté le 7 janvier 2024).
38. Russell et Norvig 2021, chapitre 21.
39. (en) HP Newquist, *The Brain Makers : Genius, Ego, And Greed in the Quest For Machines That Think*, New York: Macmillan/SAMS (ISBN 978-0-9885937-1-8), p. 30.
40. (en) Matthew Kressel, « 36 Days of Judaic Myth: Day 24, The Golem of Prague (<https://www.matthewkressel.net/2015/10/01/36-days-of-judaic-myth-day-24-the-golem-of-prague/>) », 1^{er} octobre 2015 (consulté le 16 décembre 2023).
41. (en) Pamela McCorduck, *Machines Who Think*, 2004, 604 p. (ISBN 1-56881-205-1).
42. Jacques Henno, « 1956 : et l'intelligence artificielle devint une science (<https://www.lesechos.fr/2017/08/1956-et-l-intelligence-artificielle-devint-une-science-181042>) », sur *Les Échos*, 21 août 2017 (consulté le 30 juin 2020).
43. (en) Daniel Crevier, *AI: The Tumultuous Search for Artificial Intelligence*, New York, BasicBooks, 1993 (ISBN 978-0-465-02997-6), p. 17.
44. (en) Nick Bostrom, *Superintelligence : paths, dangers, strategies*, Oxford University Press, 2017, 353 p. (ISBN 978-0-19-967811-2, lire en ligne (https://books.google.com/books?id=7_H8AwAAQBAJ&printsec=frontcover)), p. 6-7.
45. Antoine Crochet-Damais, « AI winter : qu'est-ce que l'hiver de l'IA ? (<https://www.journaldunet.fr/intelligence-artificielle/guide-de-l-intelligence-artificielle/1501295-ai-winter/>) », sur *Le Journal du Net*, 6 janvier 2022 (consulté le 16 décembre 2023).
46. (en) Pamela McCorduck, *Machines Who Think*, 2004, 604 p. (ISBN 1-56881-205-1), p. 423.
47. « Un cerveau artificiel annoncé dans dix ans (<http://www.lefigaro.fr/sciences/2009/07/31/01008-20090731ARTFIG00452-un-cerveau-artificiel-annonce-dans-dix-ans-.php>) », *Le Figaro*, 8 septembre 2009.
48. (en) « Rogue superintelligence and merging with machines: Inside the mind of OpenAI's chief scientist (<https://www.technologyreview.com/2023/10/26/1082398/exclusive-ilya-sutskever-openai-chief-scientist-on-his-hopes-and-fears-for-the-future-of-ai/>) », sur *MIT Technology Review* (consulté le 17 décembre 2023).

49. (en) Cade Metz, Karen Weise, Nico Grant et Mike Isaac, « Ego, Fear and Money: How the A.I. Fuse Was Lit », *The New York Times*, 3 décembre 2023 (ISSN 0362-4331 (<https://portal.issn.org/resource/issn/0362-4331>), lire en ligne (<https://www.nytimes.com/2023/12/03/technology/ai-openai-musk-page-altman.html>), consulté le 17 décembre 2023).
50. « OpenAI : l'association de recherche ouverte d'Elon Musk devient une entreprise fermée (<https://www.lesechos.fr/tech-medias/intelligence-artificielle/openai-lassociation-de-recherche-ouverte-delon-musk-devient-une-entreprise-fermee-999865>) », *Les Échos*, 12 mars 2019 (consulté le 17 décembre 2023).
51. Nicolas P. Rougier, « Votre chat est plus intelligent qu'une IA (<https://www.slate.fr/story/183020/tech-intelligence-artificielle-ia-chat-etre-humain-marche-reconnaissance-objet-experience-sensible>) », sur *Slate*, 18 octobre 2019 (consulté le 19 octobre 2019).
52. Stephane Nachez, « Yann LeCun, Geoffrey Hinton et Yoshua Bengio reçoivent le prix Turing (<https://www.actuaia.com/actualite/yann-lecun-geoffrey-hinton-et-yoshua-bengio-recoivent-le-prix-turing/>) », sur *ActuaIA* (consulté le 17 décembre 2023).
53. (en) « Fathers of the Deep Learning revolution receive 2018 ACM A.M. Turing Award (<https://www.acm.org/media-center/2019/march/turing-award-2018>) », sur *acm.org* (consulté le 17 décembre 2023).
54. (en) Laurie Clarke, « When AI can make art – what does it mean for creativity? », *The Observer*, 12 novembre 2022 (ISSN 0029-7712 (<https://portal.issn.org/resource/issn/0029-7712>), lire en ligne (<https://www.theguardian.com/technology/2022/nov/12/when-ai-can-make-art-what-does-it-mean-for-creativity-dall-e-midjourney>), consulté le 17 décembre 2023).
55. (en) Grace Kay, « The history of ChatGPT creator OpenAI, which Elon Musk helped found before parting ways and criticizing (<https://www.businessinsider.com/history-of-openai-company-chatgpt-elon-musk-founded-2022-12>) », sur *Business Insider* (consulté le 17 décembre 2023).
56. Liam Tung, « ChatGPT devient l'application à la croissance la plus rapide de tous les temps (<https://www.zdnet.fr/actualites/chatgpt-devient-l-application-a-la-croissance-la-plus-rapide-de-tous-les-temps-39953618.htm>) », sur *ZDNet France*, 4 février 2023 (consulté le 17 décembre 2023).
57. Benoît Mathieu, « Voici les enjeux de la course à l'intelligence artificielle », *L'Écho*, 3 juin 2023 (lire en ligne (<https://www.lecho.be/dossiers/intelligence-artificielle/voici-les-enjeux-de-la-course-a-l-intelligence-artificielle/10472118.html>)).
58. Alexandre Piquard, « L'intelligence artificielle serait aussi dangereuse que « les pandémies ou la guerre nucléaire », selon des leaders du secteur », *Le Monde*, 30 mai 2023 (lire en ligne (https://www.lemonde.fr/economie/article/2023/05/30/l-intelligence-artificielle-serait-aussi-dangereuse-que-les-pandemies-ou-la-guerre-nucleaire-selon-des-leaders-du-secteur_6175471_3234.html), consulté le 17 décembre 2023).
59. (en) Beatrice Nolan, « Here's what we know so far about Google's Gemini (<https://www.businessinsider.com/google-gemini-explainer-ai-model-2023-9>) », sur *Business Insider* (consulté le 17 décembre 2023).
60. « Avec GPT-4o, OpenAI ouvre les vannes du multimodal (<https://www.lemondeinformatique.fr/actualites/lire-avec-gpt-4o-openai-ouvre-les-vannes-du-multimodal-93713.html>) », sur *Le Monde Informatique*, 14 mai 2024 (consulté le 30 octobre 2024).
61. Pitpitt, « Intelligence artificielle générale (https://datafranca.org/wiki/Intelligence_artificielle_g%C3%A9n%C3%A9rale) », sur *DataFranca* (consulté le 11 avril 2023).
62. (en-US) « OpenAI Charter (<https://openai.com/charter>) », sur *OpenAI* (consulté le 11 avril 2023).
63. (en) Philip Boucher pour le Scientific Foresight Unit (STOA) au parlement Européen, « How artificial intelligence works (<https://www.europarl.europa.eu/at-your-service/files/be-heard/religious-and-non-confessional-dialogue/events/en-20190319-how-artificial-intelligence-works.pdf>) », sur *Parlement européen*, mars 2019 (consulté le 6 juillet 2020), p. 9.
64. Sébastien Bubeck, Varun Chandrasekaran, Ronen Eldan et Johannes Gehrke, « Sparks of Artificial General Intelligence: Early experiments with GPT-4 », *arXiv*, 27 mars 2023 ([arXiv abs/2303.12712](https://arxiv.org/abs/2303.12712) (<https://arxiv.org/abs/2303.12712>)).
65. Jaesa, « GPT-4 présente des étincelles d'intelligence artificielle générale (<https://iatranshumanisme.com/2023/03/23/gpt-4-presente-des-etincelles-dintelligence-artificielle-generale/>) », sur *Intelligence artificielle et transhumanisme*, 23 mars 2023 (consulté le 11 avril 2023).
66. (en-US) « 2022 Expert Survey on Progress in AI (<https://aiimpacts.org/2022-expert-survey-on-progress-in-ai/>) », sur *AI Impacts*, 4 août 2022 (consulté le 11 avril 2023).
67. Max Roser, « AI timelines: What do experts in artificial intelligence expect for the future? », *Our World in Data*, 29 octobre 2023 (lire en ligne (<https://ourworldindata.org/ai-timelines>), consulté le 21 novembre 2023).
68. « Superintelligence (<https://datafranca.org/wiki/Superintelligence>) », sur *datafranca* (consulté le 18 décembre 2023).
69. (en) Nick Bostrom, *Superintelligence: paths, dangers, strategies*, Oxford University Press, 2017 (ISBN 978-0-19-967811-2), « Sources of advantage for digital intelligence ».
70. Antoine Crochet-Damais, « Test de Turing : définition, principe, cas réussis, date... (<https://www.journaldunet.fr/intelligence-artificielle/guide-de-l-intelligence-artificielle/1501899-test-de-turing/>) », sur *Le Journal du Net*, 12 juillet 2022 (consulté le 18 décembre 2023).

71. (en) Olga Blinova, « Artificial intelligence: No humor, no coffee (<https://investforesight.com/artificial-intelligence-no-humor-no-coffee/>) », sur *Investforesight*, 18 novembre 2018 (consulté le 10 août 2020).
72. (en) Disha Misal, « 5 Ways To Test Whether AGI Has Truly Arrived (<https://analyticsindiamag.com/5-ways-to-test-whether-agi-has-truly-arrived/>) », sur *Analytics India Magazine*, 31 décembre 2018 (consulté le 10 août 2020).
73. (en) Nils John Nilsson, « Human-Level Artificial Intelligence? Be Serious! », *AI Magazine*, Université Stanford, 2005 (lire en ligne (<https://ai.stanford.edu/~nilsson/OnlinePubs-Nils/General%20Essays/AIMag26-04-HLAI.pdf>) [PDF]).
74. « Le prix Turing récompense trois pionniers de l'intelligence artificielle (IA) », *Le Monde*, 27 mars 2019 (lire en ligne (https://www.lemonde.fr/pixels/article/2019/03/27/le-prix-turing-recompense-trois-pionniers-de-l-intelligence-artificielle_5441937_4408996.html), consulté le 16 août 2022).
75. (en) Rodney Brooks, « How Claude Shannon Helped Kick-start Machine Learning (<https://spectrum.ieee.org/claude-shannon-information-theory>) », sur *spectrum.ieee.org*, 25 janvier 2022 (consulté le 17 mars 2024).
76. (en) « The 100 Most Influential People in AI 2023 (<https://time.com/collection/time100-ai/>) », sur *Time* (consulté le 6 novembre 2023).
77. « Déduction, analyse, organisation... Qu'est-ce que l'Intelligence artificielle ? (https://www.franceinfo.fr/replay-radio/le-billet-vert/deduction-analyse-organisation-qu-est-ce-que-l-intelligence-artificielle_7038272.html) », sur *France Info*, 10 février 2025 (consulté le 16 mai 2025).
78. Corentin Durand, « Saurez-vous entendre la différence entre Bach et l'IA musicale DeepBach ? (<https://www.numerama.com/pop-culture/219959-est-ce-du-bach-ou-une-intelligence-artificielle-se-faisant-passer-pour-le-genie-du-baroque.html>) », sur *Numerama.com*, 26 décembre 2016.
79. (en-US) « Polyphonic music generation in the style of Bach (<http://www.flow-machines.com/deepbach-polyphonic-music-generation-bach-chorales/>) », sur *flow-machines.com* (consulté le 2 juin 2018).
80. « Intelligence artificielle et santé (<https://www.inserm.fr/dossier/intelligence-artificielle-et-sante/>) », sur *Inserm* (consulté le 3 juin 2024).
81. « Intelligence artificielle : quel avenir pour les grands modèles de langage ? », *Le Monde*, 10 janvier 2024 (lire en ligne (https://www.lemonde.fr/sciences/article/2024/01/10/intelligence-artificielle-quel-avenir-pour-les-grands-modeles-de-langage_6210069_1650684.html), consulté le 3 juin 2024).
82. « Bientôt un hedge fund contrôlé par intelligence artificielle ? » (<http://bfmbusiness.bfmtv.com/bourse/bientot-un-hedge-fund-controle-par-intelligence-artificielle-869581.html>), BFM TV, 17 mars 2015.
83. « Trading haute fréquence (<https://www.lafinancepourtous.com/decryptages/marches-financiers/produits-financiers/trading-haute-frequence/>) », sur *La finance pour tous* (consulté le 14 août 2020).
84. Eric Autellet et Alexandre Papaemmanuel, « L'intelligence artificielle transforme les paradigmes traditionnels de la guerre », *Le Monde*, 31 janvier 2024 (lire en ligne (https://www.lemonde.fr/idees/article/2024/01/31/l-intelligence-artificielle-transforme-les-paradigmes-traditionnels-de-la-guerre_6214012_3232.html), consulté le 6 février 2024).
85. Cyril Bras, « L'Impact transformateur de l'intelligence artificielle sur la cybersécurité (<https://www.journaldunet.com/cybersecurite/1526089-l-impact-transformateur-de-l-intelligence-artificielle-sur-la-cybersecurite/>) », sur *Le Journal du Net*, 7 novembre 2023 (consulté le 8 février 2024).
86. (en) Robert Solano, « ChatGPT in DOD acquisitions », *USAASC*, 31 mai 2023 (lire en ligne (<https://asc.army.mil/web/news-chatgpt-in-dod-acquisitions/>)).
87. (en) « Project Maven to Deploy Computer Algorithms to War Zone by Year's End (<https://www.defense.gov/News/News-Stories/Article/Article/1254719/project-maven-to-deploy-computer-algorithms-to-war-zone-by-years-end/>) », sur *Département de la Défense des États-Unis*, 21 juillet 2017 (consulté le 8 février 2024).
88. Nathalie Guibert, « Les défis militaires de l'intelligence artificielle », *Le Monde*, 16 octobre 2018 (lire en ligne (https://www.lemonde.fr/idees/article/2018/10/16/les-defis-militaires-de-l-intelligence-artificielle_5369924_3232.html), consulté le 8 février 2024).
89. Marie-Claude Benoit, « Robots tueurs : Human Rights Watch appelle les états ayant signé la résolution de l'ONU à l'action (<https://www.actuia.com/actualite/robots-tueurs-human-rights-watch-appelle-les-etats-ayant-signe-la-resolution-de-lonu-a-laction/>) », sur *Actuia*, 9 janvier 2024 (consulté le 8 février 2024).
90. Clément Poursain, « Les drones ukrainiens pilotés par l'IA pourraient tuer sans qu'on leur en donne l'ordre (<https://korii.slate.fr/tech/drones-ukrainiens-intelligence-artificielle-tuer-sans-ordre-pilote-robots-tueurs-armes-autonomes-droit-guerre-russie>) », sur *Korii*, 16 octobre 2023 (consulté le 8 février 2024).
91. (en) Sam Biddle, « OpenAI Quietly Deletes Ban on Using ChatGPT for "Military and Warfare" (<https://theintercept.com/2024/01/12/open-ai-military-ban-chatgpt/>) », sur *The Intercept*, 12 janvier 2024 (consulté le 7 février 2024).
92. Task Force IA, *L'intelligence artificielle au service de la défense*, 2019 (lire en ligne (<https://www.defense.gouv.fr/sites/default/files/aid/20200108-NP-Rapport%20de%20la%20Task%20Force%20IA%20Septembre.pdf>) [PDF]).
93. « L'intelligence artificielle et le monde de la défense (<https://www.entreprises.gouv.fr/fr/numerique/enjeux/l-intelligence-artificielle-et-monde-de-la-defense>) », sur *entreprises.gouv.fr* (consulté le 6 février 2024).
94. Alice Vitard, « Le Comité d'éthique de la défense approuve sous conditions les armes létales autonomes », *L'Usine digitale*, 5 mai 2021 (lire en ligne (<https://www.usine-digitale.fr/article/le-comite-d-ethique-de-la-defense-approuve-sous-conditions-les-armes-letales-autonomes.N1090269>), consulté le 8 février 2024).

95. ministère de l'Europe et des Affaires étrangères, « Systèmes d'armes létales autonomes, quelle est l'action de la France ? (<https://www.diplomatie.gouv.fr/fr/politique-etrangere-de-la-france/securite-desarmement-et-non-proliferation/desarmement-et-non-proliferation/systemes-d-armes-letales-autonomes-quelle-est-l-action-de-la-france/>) », sur *diplomatie.gouv.fr*, ministère de l'Europe et des Affaires étrangères (consulté le 6 février 2024).
96. « Le ministère des Armées déploie son IA générative sécurisée et souveraine : GenIA.intradef (https://www.actuia.com/actualite/le-ministere-des-armees-deploie-son-ia-generative-securisee-et-souveraine-genial-intradef/?mc_cid=ea4cee7118&mc_eid=1498821349) », sur *ActuiA* (consulté le 17 février 2025).
97. (en-US) Amjad Iraqi, « 'Lavender': The AI machine directing Israel's bombing spree in Gaza (<https://www.972mag.com/lavender-ai-israeli-army-gaza/>) », sur *+972 Magazine*, 3 avril 2024 (consulté le 6 avril 2024).
98. (en-GB) Harry Davies, Bethan McKernan et Dan Sabbagh, « 'The Gospel': how Israel uses AI to select bombing targets in Gaza », *The Guardian*, 1^{er} décembre 2023 (lire en ligne (<https://www.theguardian.com/world/2023/dec/01/the-gospel-how-israel-uses-ai-to-select-bombing-targets>), consulté le 4 décembre 2023).
99. « Le recours à l'intelligence artificielle par Israël augmente-t-il le nombre de Gazaouis tués ? (<https://www.radiofrance.fr/franceculture/podcasts/la-revue-de-presse-internationale/la-revue-de-presse-internationale-emission-du-lundi-30-decembre-2024-5713612>) », sur *France Culture*, 30 décembre 2024 (consulté le 1^{er} janvier 2025).
100. (en) « Israel Loosened Its Rules to Bomb Hamas Fighters, Killing Many More Civilians (<https://www.nytimes.com/2024/12/26/world/middleeast/israel-hamas-gaza-bombing.html>) », sur *The New York Times*, 26 décembre 2024 (consulté le 1^{er} janvier 2025).
101. « L'intelligence artificielle, as du diagnostic médical », *Le Monde*, 3 octobre 2018 (lire en ligne (https://www.lemonde.fr/sciences/article/2018/10/03/l-intelligence-artificielle-as-du-diagnostic_5363627_1650684.html), consulté le 27 juin 2020).
102. « Maladies de l'œil : l'intelligence artificielle meilleure que les médecins ? », *Les Échos*, 14 août 2018 (lire en ligne (<https://www.lesechos.fr/intelligence-artificielle/veille-technologique/0302115279120-maladies-de-loeil-lintelligence-artificielle-meilleure-que-les-medecins-2197970.php>)).
103. Nathan Mann, « Une intelligence artificielle de Google dénoue le problème cinquantenaire du repliement des molécules », *L'Usine nouvelle*, 30 novembre 2020 (lire en ligne (<https://www.usinenouvelle.com/editorial/une-intelligence-artificielle-de-google-denoue-le-probleme-cinquantenaire-du-repliement-des-molecules.N1034924>), consulté le 28 janvier 2024).
104. François Manens, « Google vient-il de résoudre un problème vieux de 50 ans en biologie ? (<https://www.numerama.com/sciences/672896-google-vient-il-de-resoudre-un-probleme-vieux-de-50-ans-en-biologie.html>) », sur *Numerama*, 1^{er} décembre 2020 (consulté le 28 janvier 2024).
105. Sandrine Cabut, Alexandre Piquard et Martin Untersinger, « Controversé et retardé, le Health Data Hub veut pourtant faire ses preuves dans la santé », *Le Monde*, 26 janvier 2022 (lire en ligne (https://www.lemonde.fr/economie/article/2022/01/26/controversé-et-retardé-le-health-data-hub-veut-pourtant-faire-ses-preuves-dans-la-santé_6111091_3234.html), consulté le 28 janvier 2024).
106. (en) Hilary Brueck, « The newest version of ChatGPT passed the US medical licensing exam with flying colors — and diagnosed a 1 in 100,000 condition in seconds (<https://www.businessinsider.com/chatgpt-passes-medical-exam-diagnoses-rare-condition-2023-4>) », sur *Business Insider*, 6 avril 2023 (consulté le 28 janvier 2024).
107. (en) Kelly N. DuBois, « DEEP MEDICINE: How Artificial Intelligence Can Make Healthcare Human Again (<https://go.gale.com/ps/i.do?id=GALE%7CA600162655&sid=googleScholar&v=2.1&it=r&linkaccess=abs&issn=08922675&p=AONE&sw=w&userGroupName=anon%7E61539f80&aty=open-web-entry>) », septembre 2019.
108. Scott Mayer McKinney, Marcin Sieniek, « International evaluation of an AI system for breast cancer screening (<https://www.nature.com/articles/s41586-019-1799-6>) », 14 octobre 2020.
109. Jean-François Lemoine, « Révolution dans le développement de médicaments grâce à l'IA (<https://www.pourquoidocteur.fr/Articles/Question-d-actu/48829-Revolution-developpement-medicaments-grace-l-IA>) ».
110. (en) Andre Esteva, « Dermatologist-level classification of skin cancer with deep neural networks », *Nature*, 29 juin 2017 (lire en ligne (<https://www.nature.com/articles/nature21056>)).
111. Alexandre Boero, « La police britannique travaille sur une IA qui sera capable de devancer votre crime (<https://www.clubic.com/technologies-d-avenir/intelligence-artificielle/actualite-848035-police-britannique-travaille-ia-devancer-crime.html>) », sur *Clubic*, 2 décembre 2018.
112. (en) Stephen Buranyi, « The rise of the racist robots (<https://www.theguardian.com/inequality/2017/aug/08/rise-of-the-racist-robots-how-ai-is-learning-all-our-worst-impulses>) », *The Guardian*, 9 août 2017.
113. Romaric Saint-Aubert, « Quand l'IA devient l'arme du crime (<https://www.business2community.com/fr/actu/quand-l-ia-devient-l-arme-du-crime>) » , sur *Business2Community.com*, 1^{er} août 2023.
114. « Justice Prédictive : de l'idée à la réalité (<https://www.justice-predictive.com/index.php/2-non-categorise/24-justice-predictive-de-l-idee-a-la-realite>) », sur *justice-predictive.com* (consulté le 27 janvier 2019).
115. Harold Grand, « En Estonie, une intelligence artificielle va rendre des décisions de justice (<https://www.lefigaro.fr/secteur/high-tech/en-estonie-une-intelligence-artificielle-va-rendre-des-decisions-de-justice-20190401>) », sur *Le Figaro*, 1^{er} avril 2019 (consulté le 29 juin 2020).
116. « Les avocats face à la révolution numérique. S'adapter pour rebondir (<https://www.atlantico.fr/decryptage/2707979/les-avocats-face-a-la-revolution-numerique-s-adapter-pour-rebondir-thierry-wickers-pierre-aidan>) », sur *Atlantico.fr* (consulté le 10 août 2020).

- L17. Vincent Bouquet, « Supply chain : l'intelligence artificielle, une promesse de levier d'amélioration (<https://business.lesechos.fr/directions-financieres/comptabilite-et-gestion/logistique/0602147719753-supply-chain-l-intelligence-artificielle-une-promesse-de-levier-d-amelioration-332796.php#Xtor=AD-6000>) », sur *Les Echos Executives*, 4 novembre 2019 (consulté le 29 juin 2020).
- L18. « Urbanloop – smart urban mobility (<http://urbanloop.univ-lorraine.fr/>) », sur *urbanloop.univ-lorraine.fr* (consulté le 29 juin 2020).
- L19. « Google Maps roule à l'IA pour optimiser la navigation (<https://www.journaldunet.com/solutions/dsi/1445415-google-maps-roule-a-l-ia-pour-optimiser-la-navigation>) », sur *journaldunet.com* (consulté le 10 juillet 2020).
- L20. Stéphanie Condon, « Vidéo : comment l'intelligence artificielle a contribué à améliorer Google Maps (<https://www.zdnet.fr/actualites/video-comment-l-intelligence-artificielle-a-contribue-a-ameliorer-google-maps-39898773.htm>) », sur *ZDNet France* (consulté le 10 juillet 2020).
- L21. Larry Dignan, « Facebook associe son service Map With AI au projet OpenStreetMap (<https://www.zdnet.fr/actualites/facebook-associe-son-service-map-with-ai-au-projet-openstreetmap-39888195.htm>) », sur *ZDNet France* (consulté le 10 juillet 2020).
- L22. Acatech 2011.
- L23. « Les enjeux de l'intelligence artificielle dans la robotique », *Journal Innovations et Technologies*, n° 210, juillet/août 2018, p. I-VI (lire en ligne (<http://www.coboteam.fr/wp-content/uploads/jitec210-enjeux-ia-robotique-1.pdf>) [PDF], consulté le 3 février 2019).
- L24. Amélie Cordier, *Introduction à l'intelligence artificielle développementale* (cours), Laboratoire d'informatique en image et systèmes d'information, 41 p. (lire en ligne (<https://perso.liris.cnrs.fr/amelie.cordier/wiki/lib/exe/fetch.php?media=ens:iad1.pdf>) [PDF]).
- L25. Michel Bret, Edmond Couchot et Patrice Besnard (restauration), « Les Pissenlits (<http://www.archives-video.univ-paris8.fr/video.php?recordID=232>) », vidéo de l'installation interactive, durée 1 min 8 s, sur *Université Paris-VIII*, 2006 (consulté le 21 juin 2018).
- L26. *Living art : L'art numérique*, CNRS Éditions (présentation en ligne (<http://www.cnrseditions.fr/art/6207-living-art-flor-ent-aziosmanoff.html>)).
- L27. « Obvious, les Français derrière la première peinture d'une IA vendue aux enchères (<https://www.france24.com/fr/20181019-peinture-christies-obvious-algorithme-ia-edmond-belamy-vente-encheres>) », sur *France 24*, 19 octobre 2018.
- L28. « Une peinture, réalisée par une Intelligence artificielle, mise aux enchères (<https://www.objetconnecte.net/peinture-intelligence-artificielle-encheres/>) », sur *objetconnecte.net* (consulté le 28 novembre 2018).
- L29. Floriane Leclerc, « Que sont les GANs, ces systèmes de machine learning à l'origine des deepfakes ? », *L'Usine Digitale*, 20 novembre 2019 (lire en ligne (<https://www.usine-digitale.fr/article/que-sont-les-gans-ces-systemes-de-machine-learning-a-l-origine-des-deepfakes.N905544>), consulté le 11 février 2024).
- L30. « Arts : l'autre Terre de Grégory Chatonsky au Palais de Tokyo », *Le Monde*, 27 juin 2019 (lire en ligne (https://www.lemonde.fr/culture/article/2019/06/27/arts-l-autre-terre-de-gregory-chatonsky-au-palais-de-tokyo_5482045_3246.html), consulté le 25 août 2022).
- L31. « OpenAI's DALL-E creates plausible images of literally anything you ask it to (<https://techcrunch.com/2021/01/05/openais-dall-e-creates-plausible-images-of-literally-anything-you-ask-it-to/>) », sur *TechCrunch*, 6 janvier 2021 (consulté le 20 avril 2022).
- L32. « ChatGPT et Midjourney font évoluer vos pratiques professionnelles ? Racontez-nous », *Le Monde*, 8 avril 2023 (lire en ligne (https://www.lemonde.fr/pixels/appel-temoignages/2023/04/08/chatgpt-et-midjourney-font-evoluer-vos-pratiques-professionnelles-racontez-nous_6168771_4408996.html), consulté le 3 juin 2023).
- L33. Victor Vasseur, « Les fausses images de Macron en éboueur et de Trump en prison montrent l'incroyable potentiel de l'IA (<https://www.radiofrance.fr/franceinter/les-fausses-images-de-macron-en-eboueur-et-de-trump-en-prison-montrent-l-incroyable-potentiel-de-l-ia-2140206>) », sur *France Inter*, 22 mars 2023 (consulté le 3 juin 2023).
- L34. « Cette fausse photo du pape François en doudoune blanche est devenue virale, voici son histoire (<https://www.ouest-france.fr/leditiondusoir/2023-03-27/cette-fausse-photo-du-pape-francois-en-doudoune-blanche-est-devenue-virale-voici-son-histoire-562b46f9-a655-4744-84e9-c2b2559ed288>) », sur *Ouest-France.fr*, 27 mars 2023 (consulté le 3 juin 2023).
- L35. (en) Jordan Hart, The AI program used to generate fake viral images of Pope Francis and Donald Trump just suspended free trials — but professional photographers say they're still concerned (<https://www.businessinsider.com/midjourney-suspends-free-trials-fake-images-pope-trump-viral-photographers-2023-3>), Business Insider, 30 mars 2023, *Business Insider*.
- L36. Clémentine Mercier, « Intelligence artificielle : «Le droit d'auteur protège une création précise, mais pas une manière de créer» (https://www.liberation.fr/culture/intelligence-artificielle-le-droit-dauteur-protège-une-creation-précise-mais-pas-une-maniere-de-cree-20221231_W6A00KULC5HMHDEUCRL76HEY4A/) », sur *Libération*, 31 décembre 2022 (consulté le 19 novembre 2023).
- L37. Photos créées par des IA : une bascule vertigineuse et dangereuse (https://www.liberation.fr/politique/photos-crees-par-des-ia-une-bascule-vertigineuse-et-dangereuse-20230331_G3ZTL7HYTVEENMQ5DBJ7VT3FJQ/), Jonathan Bouchet-Petersen, 31 mars 2023, *Libération* : « Emmanuel Macron ramassant les poubelles, le Pape en doudoune blanche type Bibendum, Donald Trump se débattant au milieu de policiers venus l'arrêter... Ces images ont en commun d'avoir été générées par des intelligences artificielles. »

- L38. Stéphanie Lemoine, « Grégory Chatonsky : « L'intelligence artificielle est une nouvelle façon de naviguer dans la culture humaine » (<https://www.lejournaldesarts.fr/creation/gregory-chatonsky-lintelligence-artificielle-est-une-nouvelle-facon-de-naviguer-dans-la>) », sur *Le Journal Des Arts*, 11 août 2023 (consulté le 11 février 2024).
- L39. « Au Japon, une autrice lauréate d'un prix littéraire dit avoir écrit en partie son roman avec ChatGPT (https://www.iberation.fr/culture/livres/au-japon-une-autrice-laureate-dun-prix-litteraire-dit-avoir-ecrit-en-partie-son-roman-avec-chatgpt-20240118_3LA4XHUGKBFAJCFKGNZ74UZPAY/) », sur *Libération*, 18 janvier 2024 (consulté le 11 février 2024).
- L40. « Avec Sora, OpenAI impressionne dans la génération de vidéo (<https://www.lemondeinformatique.fr/actualites/lire-avec-sora-openai-impressionne-dans-la-generation-de-video-92986.html>) », sur *Le Monde informatique*, 16 février 2024 (consulté le 17 février 2024).
- L41. (en) Vandana Nair, « AI-Music Platform Race Accelerates with Udio (<https://analyticsindiamag.com/ai-music-platform-race-accelerates-with-udio/>) », sur *Analytics India Magazine*, 11 avril 2024 (consulté le 15 avril 2024).
- L42. « Comment l'intelligence artificielle va-t-elle changer notre vie quotidienne ? (<http://www.robots-et-compagnie.com/intelligence-artificielle-changer-notre-vie-quotidienne/>) », sur *Robots & Cie*, 30 novembre 2016 (consulté le 27 janvier 2019).
- L43. « Quand l'IA aide les développeurs à programmer (<https://experiences.microsoft.fr/technique/intelligence-artificielle-ia-technique/ia-developpeurs-programmer/>) », sur *Microsoft* (consulté le 27 janvier 2019).
- L44. Ivan Capecchi, « Médias, faites GAFA à l'IA ! », *L'ADN.eu*, 11 mai 2017 (lire en ligne (<http://www.ladn.eu/nouveaux-usages/etude-marketing/medias-faites-gafa-a-lia/>)).
- L45. « La première animatrice télé virtuelle fait ses débuts en Corée du Sud (<https://www.20minutes.fr/high-tech/2912063-20201119-la-premiere-animatrice-tele-virtuelle-fait-ses-debuts-en-coree-du-sud>) », sur *20minutes.fr* (consulté le 23 novembre 2020).
- L46. (en-US) « The A.I. Chair (<https://www.gessato.com/artificial-intelligence-chair/>) », sur *Gessato*, 17 avril 2019 (consulté le 27 février 2020).
- L47. « Bill Gates (<http://www.universalis.fr/encyclopedie/gates-william-h-iii-dit-bill/>) », sur *Encyclopædia Universalis* (consulté le 16 octobre 2016).
- L48. « Intelligence artificielle, les défis actuels et l'action d'Inria (<https://www.inria.fr/actualite/actualites-inria/livre-blanc-sur-l-intelligence-artificielle>) », sur *INRIA*, 2016 (consulté le 15 décembre 2016).
- L49. « « Intelligence artificielle, intelligence humaine : la double énigme », de Daniel Andler : chimériques machines qui pensent », *Le Monde*, 4 juin 2023 (lire en ligne (https://www.lemonde.fr/livres/article/2023/06/04/intelligence-artificielle-intelligence-humaine-la-double-énigme-de-daniel-andler-chimeriques-machines-qui-pensent_6176109_3260.html)) , consulté le 12 juin 2023).
- L50. (en) Irving John Good, « *Speculations Concerning the First Ultraintelligent Machine* », *Advances in Computers*, vol. 6, 1966, p. 31-88.
- L51. (en) Vernor Vinge, « The coming technological singularity: How to survive in the post-human era », NASA, 1^{er} décembre 1993 (lire en ligne (<https://ntrs.nasa.gov/citations/19940022856>)), consulté le 3 mars 2024).
- L52. Ray Kurzweil (trad. Adeline Mesmin), *Humanité 2.0: la bible du changement*, M21, 2007 (ISBN 978-2-916260-04-4).
- L53. (en) Raffi Khatchadourian, « The Doomsday Invention (<https://www.newyorker.com/magazine/2015/11/23/doomsday-invention-artificial-intelligence-nick-bostrom>) », sur *The New Yorker*, 16 novembre 2015 (consulté le 3 mars 2024).
- L54. « Sommet sur les risques de l'intelligence artificielle : ce que l'on sait du programme (<https://www.ouest-france.fr/high-tech/intelligence-artificielle/sommet-sur-les-risques-de-lintelligence-artificielle-ce-que-lon-sait-du-programme-dc3c8434-78a7-11ee-b7ba-95ef9cb6f5e8>) », sur *Ouest-France*, 1^{er} novembre 2023 (consulté le 4 mars 2024).
- L55. Rémi Sussan, « Stephen Hawking craint l'intelligence artificielle. Et s'il avait tort ? (<https://www.nouvelobs.com/rue89/rue89-connexions-dangereuses/20150206.RUE7731/stephen-hawking-craint-l-intelligence-artificielle-et-s-il-avait-tort.html>) », sur *nouvelobs.com*, 6 février 2015 (consulté le 15 avril 2023).
- L56. « Selon Stephen Hawking, l'intelligence artificielle est un danger pour l'humanité (<http://www.webdeveloppementdurable.com/selon-stephen-hawking-lintelligence-artificielle-danger-pour-lhumanite/>) », sur *webdeveloppementdurable.com* (consulté le 28 février 2016).
- L57. Alexandre Piquard, « Intelligence Artificielle Web Summit (https://www.lemonde.fr/economie/article/2017/11/10/l-intelligence-artificielle-star-inquietante-du-web-summit-a-lisbonne_5213087_3234.html) », *Le Monde*, 10 novembre 2017.
- L58. L'intelligence artificielle inquiète les grands cerveaux contemporains (<http://www.france24.com/fr/20150202-bill-gates-ia-risque-intelligence-artificielle-stephen-hawking-humanite-elon-musk-robot-algorithme>), France 24, 2 février 2015.
- L59. « L'intelligence artificielle pourrait mettre 50% de l'humanité au chômage (https://web.archive.org/web/20170502114530/http://www.lexpress.fr/actualite/sciences/l-intelligence-artificielle-pourrait-mettre-50-de-l-humanite-au-chomage_1763475.html) », sur *L'Express*, 13 février 2016 (version du 2 mai 2017 sur *Internet Archive*).
- L60. (en) Joe McKendrick, « Let's Make Artificial Intelligence 'Boring' Again (<https://www.forbes.com/sites/joemckendrick/2018/11/13/lets-make-artificial-intelligence-boring-again/>) », sur *Forbes* (consulté le 24 février 2023).

- L61. (en) Nick Bostrom, « What happens when our computers get smarter than we are? (https://www.ted.com/talks/nick_bostrom_what_happens_when_our_computers_get_smarter_than_we_are/transcript) », 27 avril 2015 (consulté le 15 avril 2023) : « We should not be confident in our ability to keep a superintelligent genie locked up in its bottle forever. Sooner or later, it will out. I believe that the answer here is to figure out how to create superintelligent A.I. such that even if—when—it escapes, it is still safe because it is fundamentally on our side because it shares our values. »
- L62. DR, « L'intelligence artificielle devient de plus en plus incompréhensible (<https://www.parismatch.com/actu/science/s/lintelligence-artificielle-devient-de-plus-en-plus-incomprehensible-218444>) », sur *Paris Match*, 8 novembre 2022.
- L63. « L'IA pourrait poser un « risque d'extinction » pour l'humanité, affirment 350 experts (<https://www.lesechos.fr/tech-medias/intelligence-artificielle/ia-pourrait-poser-un-risque-dextinction-pour-lhumanite-affirment-350-experts-1947735>) », sur *Les Échos*, 30 mai 2023 (consulté le 4 novembre 2023).
- L64. Kevin Roose, « Vie numérique: L'IA constitue un « risque d'extinction » pour l'humanité », *La Presse (traduit depuis The New York Times)*, 30 mai 2023 (lire en ligne (<https://www.lapresse.ca/affaires/techno/2023-05-30/vie-numerique/l-ia-constitue-un-risque-d-extinction-pour-l-humanite.php>), consulté le 4 novembre 2023).
- L65. Lewis Mumford, *Le Mythe de la machine*, deux volumes, 1967-1970; tard., Fayard, 1974.
- L66. Jacques Ellul, *La Technique ou l'Enjeu du siècle*, 1954; 3^e édition : Economica 2008.
- L67. Günther Anders, *L'obsolescence de l'homme*, tome 1, 1956; tome 2, 1980; tard. fr. 2002 et 2011.
- L68. *Pièces et Main d'Œuvre* Site officiel (<http://www.piecesetmaindoeuvre.com>).
- L69. *Technologos* Site officiel (<http://technologos.fr>).
- L70. François Jarrige, *Techno-critiques. Du refus des machines à la contestation des technosciences*, La Découverte, 2014.
- L71. « Intelligence artificielle : les risques d'une utilisation malveillante (<https://www.futura-20sciences.com/tech/actualites/intelligence-artificielle-intelligence-artificielle-risques-utilisation-malveillante-70247/>) », *Futura*, 21 février 2018 (consulté le 21 février 2018).
- L72. Damien Licata Caruso, « L'intelligence artificielle générative, ce monstre énergivore qui inquiète (<https://www.leparisien.fr/high-tech/lintelligence-artificielle-generative-ce-monstre-energivore-qui-inquiete-04-07-2024-K53TRWJBE5GQ3EHQUYHOM5OZWE.php>) », sur *leparisien.fr*, 4 juillet 2024 (consulté le 3 septembre 2024).
- L73. « La climatisation et le boom de l'IA vont-ils faire dérailler les engagements climatiques des États ? (<https://theconversation.com/la-climatisation-et-le-boom-de-lia-vont-ils-faire-derailer-les-engagements-climatiques-des-etats-238456>) », *The conversation*, 2024 (consulté le 23 novembre 2024).
- L74. Le nucléaire, une planche de salut énergétique pour l'IA (<https://www.lemondeinformatique.fr/actualites/lire-le-nucleaire-une-planche-de-salut-energetique-pour-l-ia-94255.html>), *Le Monde informatique*, 10 juillet 2024.
- L75. Les géants du numérique se convertissent au nucléaire pour étancher les besoins énergétiques toujours plus importants de l'IA (https://www.lemonde.fr/economie/article/2024/09/23/les-geants-du-numerique-se-convertissent-au-nucleaire_6329378_3234.html), *le monde.fr*, 23 septembre 2024
- L76. (en) « The open-source AI boom is built on Big Tech's handouts. How long will it last? (<https://www.technologyreview.com/2023/05/12/1072950/open-source-ai-google-openai-eleuther-meta/>) », sur *MIT Technology Review* (consulté le 13 avril 2024).
- L77. (en) Kyle Wiggers, « Google open sources tools to support AI model development (<https://techcrunch.com/2024/04/09/google-open-sources-tools-to-support-ai-model-development/>) », sur *TechCrunch*, 9 avril 2024 (consulté le 13 avril 2024).
- L78. « La majorité des efforts de Mozilla pour donner de l'élan au mouvement sont orientés vers le développement d'une intelligence artificielle (IA) digne de confiance. (<https://foundation.mozilla.org/fr/internet-health/trustworthy-artificial-intelligence/>) », *La santé d'Internet*, Mozilla Foundation, 2023 (consulté le 22 avril 2023).
- L79. Pierre Fontaine, « Comment la fondation Mozilla veut créer une IA de confiance pour nous sauver des géants de la tech (<https://www.01net.com/actualites/comment-la-fondation-mozilla-veut-creer-une-ia-de-confiance-contre-les-geants-de-la-tech.html>) », sur *01net*, 23 mars 2023 (consulté le 21 avril 2023).
- L80. (en) « AI startup Mistral launches a 281GB AI model to rival OpenAI, Meta, and Google (<https://www.zdnet.com/article/ai-startup-mistral-launches-a-281gb-ai-model-to-rival-openai-meta-and-google/>) », sur *ZDNet*, 10 avril 2024 (consulté le 13 avril 2024).
- L81. (en) Alex McFarland et Antoine Tardif, « 5 Best Open Source LLMs (April 2024) (<https://www.unite.ai/best-open-source-llms/>) », sur *Unite.ai*, 1^{er} avril 2024 (consulté le 13 avril 2024).
- L82. « L'équilibre délicat entre sécurité et innovation dans l'IA : « bannir les modèles "open weights" serait un désastre » selon un chercheur (<https://intelligence-artificielle.developpez.com/actu/356012/L-equilibre-delicat-entre-securite-et-innovation-dans-l-ia-bannir-les-modeles-open-weights-serait-un-desastre-selon-un-chercheur-l-administration-biden-envisage-de-bloquer-l-acces-a-ces-modeles-afin-d-eviter-les-abus/>) », sur *Developpez.com*, 3 avril 2024 (consulté le 13 avril 2024).
- L83. (en) Matt Marshall, « How enterprises are using open source LLMs: 16 examples », *VentureBeat*, 29 janvier 2024 (lire en ligne (<https://venturebeat.com/ai/how-enterprises-are-using-open-source-llms-16-examples/>)).
- L84. (en) Kelsey Piper, « Should we make our most powerful AI models open source to all? (<https://www.vox.com/future-perfect/2024/2/2/24058484/open-source-artificial-intelligence-ai-risk-meta-llama-2-chatgpt-openai-deepfake>) », sur *Vox*, 2 février 2024 (consulté le 13 avril 2024).

185. « Intelligence artificielle : la transformation en cours des métiers peut-elle favoriser l'émergence d'un revenu universel ? (https://www.francetvinfo.fr/replay-radio/aujourd-hui-c-est-demain/intelligence-artificielle-la-transformation-en-cours-des-metiers-peut-elle-favoriser-l-emergence-d-un-revenu-universel_6047222.html) », sur *France Info*, 21 septembre 2023 (consulté le 8 février 2025).
186. Sylvain Biget, « Les emplois qui seront les plus touchés ou remplacés par les IA génératives comme ChatGPT, selon Goldman-Sachs (<https://www.futura-sciences.com/tech/actualites/intelligence-artificielle-emplois-seront-plus-touches-remplaces-ia-generatives-comme-chatgpt-selon-goldman-sachs-104373/>) », sur *futura-sciences.com*, 1^{er} avril 2023 (consulté le 5 août 2024).
187. « L'intelligence artificielle en veut-elle à votre emploi? (<https://www.lesaffaires.com/secteurs/techno/lintelligence-artificielle-en-veut-elle-a-votre-emploi-2/>) », sur *lesaffaires.com*, 4 août 2023 (consulté le 5 août 2024).
188. Thomas Jestin, « L'IA va-t-elle provoquer un chômage de masse ? Creuser les inégalités ? Renforcer les GAFA ? (<https://www.journaldunet.com/intelligence-artificielle/1523055-l-ia-va-t-elle-provoquer-un-chomage-de-masse-creuser-les-inegalites-renforcer-les-gafa-voici-pourquoi-nous-avons-toutes-les-chances-d-eviter-le-pire/>) », sur *Journal du Net*, 3 octobre 2023 (consulté le 5 août 2024).
189. François Desnoyers, « « Un monde sans travail » : quand le chômage technologique arrivera... », *Le Monde*, 26 janvier 2023 (lire en ligne (https://www.lemonde.fr/emploi/article/2023/01/26/un-monde-sans-travail-quand-le-chomage-technologique-arrivera_6159363_1698637.html), consulté le 5 août 2024).
190. « Long story | Sam Altman et le revenu universel de base (<https://www.forbes.fr/business/long-story-sam-altman-et-le-revenu-universel-de-base/>) », sur *Forbes France*, 30 juillet 2024 (consulté le 8 février 2025).
191. Langlois, « L'éthique en matière d'intelligence artificielle : les biais discriminatoires (<https://langlois.ca/ressources/ethique-en-matiere-dintelligence-artificielle-les-biais-discriminatoires/>) », sur *Langlois* (consulté le 5 janvier 2025).
192. Nathan Noiry, « Des biais de représentativité en intelligence artificielle (<https://www.lemonde.fr/blog/binaire/2021/08/31/des-biais-de-representativite-en-intelligence-artificielle/>) », *Le Monde* (consulté le 5 janvier 2025).
193. Harari 2024, p. 397-400.
194. Harari 2024.
195. Éric Sadin, « L'asservissement par l'Intelligence Artificielle ? (<https://www.youtube.com/watch?v=VzeOnBRzDik&t=3315s>) » [vidéo], sur *Thinkerview*, 8 novembre 2018 (consulté le 4 janvier 2015).
196. Éric Sadin, « Le devenir légume de l'humanité ? (<https://www.youtube.com/watch?v=P1xA8xSXmqg>) » [vidéo], sur *Thinkerview*, 19 octobre 2023 (consulté le 4 janvier 2025).
197. (en) Nicholas Carr, *The Shallows: What the Internet Is Doing to Our Brains*.
198. « Les impacts sociaux de l'IA (<https://ecoinfo.cnrs.fr/2022/03/15/impacts-sociaux-de-lia/>) », sur *EcoInfo* (consulté le 5 janvier 2025).
199. « Les 20 menaces les plus dangereuses de l'intelligence artificielle (<https://www.futura-sciences.com/tech/question-s-ponses/intelligence-artificielle-20-menaces-plus-dangereuses-intelligence-artificielle-14343/>) » [vidéo], sur *Futura* (consulté le 5 janvier 2025).
200. Yuval Noah Harari, *21 leçons pour le xx^e siècle*.
201. « Myanmar. Les systèmes de Facebook ont promu la violence contre les Rohingyas – Meta doit des réparations. (<https://www.amnesty.org/fr/latest/news/2022/09/myanmar-facebook-systems-promoted-violence-against-rohingya-a-meta-owes-reparations-new-report/>) », sur *Amnesty International*, 29 septembre 2022 (consulté le 6 janvier 2025).
202. (en) Michael Kan, « GPT-4 Was Able To Hire and Deceive A Human Worker Into Completing a Task (<https://www.pcmag.com/news/gpt-4-was-able-to-hire-and-deceive-a-human-worker-into-completing-a-task>) », sur *PCMag*.
203. « GPT-4 a réussi à tromper un humain, l'incitant à résoudre un CAPTCHA après avoir prétendu être aveugle (<https://intelligence-artificielle.developpez.com/actu/342611/GPT-4-a-reussi-a-tromper-un-humain-l-incident-a-resoudre-un-CAPTCHA-apres-avoir-pretendu-etre-aveugle-OpenAI-a-presente-cette-version-de-ChatGPT-comme-etant-meilleure-et-plus-precise/>) », sur *Developpez.com*, 19 mars 2023 (consulté le 5 janvier 2025).
204. « Laurent Alexandre : "L'IA va accroître les inégalités intellectuelles et la fragmentation sociale" (<https://www.lenouveleconomiste.fr/entretien-avec-laurent-alexandre-107968/>) », *Le Nouvel Économiste*, 12 décembre 2023 (consulté le 15 mai 2025).
205. (en) « Appliquer les Principes de l'OCDE sur l'IA : progrès et perspectives (<https://oecd.ai/en/mcm-fr>) », sur *oecd.ai* (consulté le 3 novembre 2023).
206. « Création du partenariat mondial pour l'intelligence artificielle : secrétariat au sein de l'OCDE à Paris (<https://www.actuia.com/actualite/parteneriat-global-sur-lintelligence-artificielle-lodge-en-hebergera-le-secretariat/>) », sur *Actuia* (consulté le 3 novembre 2023).
207. (en) « AI for Good Global Summit 2018 (<https://www.itu.int/en/ITU-T/AI/2018/Pages/default.aspx>) », sur *Union internationale des télécommunications* (consulté le 3 novembre 2023).
208. Ausma Bernot et Fan Yang, « Qui gouvernera l'IA ? La course des nations pour réguler l'intelligence artificielle (<https://theconversation.com/qui-gouvernera-lia-la-course-des-nations-pour-reguler-lintelligence-artificielle-217320>) », sur *The Conversation*, 12 novembre 2023 (consulté le 20 janvier 2024).

209. Francis Donnat, « L'intelligence artificielle, une menace pour la vie privée ? », *Pouvoirs*, Éditions du Seuil, n° 170, septembre 2019, p. 95-103 (ISBN 978-2-02-140678-8, ISSN 0152-0768 (<https://portal.issn.org/resource/issn/0152-0768>), DOI 10.3917/pouv.170.0095 (<https://dx.doi.org/10.3917/pouv.170.0095>)) , lire en ligne (<https://www.cairn.info/revue-pouvoirs-2019-3-page-95.htm>)).
210. Lucie Lequier, « Tout comprendre à l'IA Act, le texte européen de réglementation de l'intelligence artificielle (<https://www.numerama.com/tech/1371250-tout-comprendre-a-lia-act-le-texte-europeen-de-reglementation-de-lintelligenc-e-artificielle.html>) », sur *Numerama*, 11 mai 2023 (consulté le 3 novembre 2023).
211. (en) Dom Galeon, « The United Arab Emirates is the first country in the world to hire a minister for artificial intelligence (<https://www.businessinsider.com/world-first-ai-minister-uae-2017-12?IR=T>) », sur *businessinsider.com*, 16 décembre 2017.
212. (en) House of Commons Digital, Culture, Media and Sport Committee, *Disinformation and 'fake news': Final Report ; Eighth Report of Session 2017–19 ; Report, together with formal minutes relating to the report*, 2019 (lire en ligne (<https://publications.parliament.uk/pa/cm201719/cmselect/cmcumeds/1791/1791.pdf>) [PDF]) ; rapport commandé par *the House of Commons*, Ref HC 1791, publié le 18 février 2019, imprimé le 14 février 2019 pour le Gouvernement britannique. Voir notamment paragraphe 149.
213. (en) « Investigation into data analytics for political purposes (<https://ico.org.uk/action-weve-taken/investigation-into-data-analytics-for-political-purposes/>) », sur *ico.org.uk*, 6 octobre 2020 (consulté le 19 mars 2021).
214. « "Sans Cambridge Analytica, il n'y aurait pas eu de Brexit", affirme le lanceur d'alerte Christopher Wylie (https://www.francetvinfo.fr/monde/europe/la-grande-bretagne-et-l-ue/sans-cambridge-analytica-il-n-y-aurait-pas-eu-de-brexit-affirme-le-lanceur-d-alerte-christopher-wylie_2677946.html) », sur *francetvinfo.fr*, 28 mars 2018 (consulté le 28 mars 2018).
215. Morgane Tual, « Intelligence artificielle : les géants du Web lancent un partenariat sur l'éthique », *Le Monde*, 28 septembre 2016 (ISSN 1950-6244 (<https://portal.issn.org/resource/issn/1950-6244>), lire en ligne (https://www.lemonde.fr/pixels/article/2016/09/28/intelligence-artificielle-les-geants-du-web-lancent-un-partenariat-sur-l-ethique_5005123_4408996.html)).
216. « Intelligence artificielle : Google DeepMind se dote d'une unité de recherche sur l'éthique », *Le Monde*, 4 octobre 2017 (ISSN 1950-6244 (<https://portal.issn.org/resource/issn/1950-6244>), lire en ligne (https://www.lemonde.fr/pixels/article/2017/10/04/intelligence-artificielle-google-deepmind-se-dote-d-une-unite-de-recherche-sur-l-ethique_5196189_4408996.html)).
217. (en) « Asilomar AI Principles (<https://futureoflife.org/open-letter/ai-principles/>) », sur *Future of Life Institute* (consulté le 5 janvier 2025).
218. Hatem Bakri, « Intelligence Artificielle. Quels Impacts sur la Société ? (<https://mbamci.com/2017/02/intelligence-artificielle-impacts/>) », sur *MBA MCI* (consulté le 5 janvier 2025).
219. (en-US) « Lethal Autonomous Weapons Pledge (<https://futureoflife.org/lethal-autonomous-weapons-pledge/>) », sur *futureoflife.org* (consulté le 13 août 2018).
220. « Des milliers d'experts en intelligence artificielle s'engagent à ne pas participer à la création d'armes », *Le Monde*, 18 juillet 2018 (lire en ligne (https://abonnes.lemonde.fr/pixels/article/2018/07/18/des-milliers-d-experts-en-intelligence-artificielle-s-engagent-a-ne-pas-participer-a-la-creation-d-armes_5333024_4408996.html)).
221. L'UNESCO rejoint le Partenariat mondial sur l'Intelligence Artificielle comme observateur (<https://www.unesco.org/fr/articles/lunesco-rejoint-le-partenariat-mondial-sur-lintelligence-artificielle-comme-observateur>), UNESCO, 2022.
222. Publications Office of the European Union, « Intelligence artificielle Une approche européenne axée sur l'excellence et la confiance, COM/2020/65 final (<https://op.europa.eu/fr/publication-detail/-/publication/ac957f13-53c6-11ea-aece-01aa75ed71a1>) », sur *Office des publications de l'Union européenne*, 19 février 2020 (consulté le 5 novembre 2023).
223. « Réglementation de l'Intelligence artificielle en Europe (<https://france.devoteam.com/paroles-dexperts/reglementation-de-lintelligence-artificielle-en-europe-vers-un-rqpd-de-lia/>) », sur *Devoteam France* (consulté le 5 novembre 2023).
224. (en) « Une Europe adaptée à l'ère du numérique : La Commission propose de nouvelles règles et actions en faveur de l'excellence et de la confiance dans l'intelligence artificielle (https://ec.europa.eu/commission/presscorner/detail/fr/ip_21_1682) », sur *Commission européenne*, 21 avril 2021 (consulté le 21 avril 2023).
225. « Premier Forum mondial sur l'éthique de l'IA à Prague, un an après l'adoption de la Recommandation de l'UNESCO (<https://www.unesco.org/fr/articles/premier-forum-mondial-sur-lethique-de-lia-prague-un-apres-ladoption-de-la-recommandation-de-lunesco>) », sur UNESCO, 9 décembre 2022.
226. « Intelligence artificielle : l'UNESCO appelle à mettre en œuvre sans délai le cadre éthique mondial (<https://news.un.org/fr/story/2023/03/1133847>) », sur *ONU Info*, 31 mars 2023 (consulté le 19 novembre 2023).
227. « Recommandation sur l'éthique de l'intelligence artificielle (https://unesdoc.unesco.org/ark:/48223/pf0000381137_fre) », sur *unesco.org*, 2022 (consulté le 27 septembre 2024).
228. « Éthique: l'Unesco appelle les États à appliquer sa recommandation sur l'intelligence artificielle (<https://web.archive.org/web/20230330234422/https://www.techniques-ingenieur.fr/actualite/breves-afp/ethique-lunesco-appelle-les-etats-a-appliquer-sa-recommandation-sur-lintelligence-artificielle-121945/>) », sur *Éditions techniques de l'ingénieur* (version du 30 mars 2023 sur *Internet Archive*).

229. « François met en garde contre les conséquences pour la paix de l'IA (<https://www.vaticannews.va/fr/pape/news/2023-12/francois-met-en-garde-contre-les-risques-de-l-ai-pour-la-paix.html>) », sur *vaticannews.fr* (consulté le 22 mai 2005).
230. Sylvain Naillat, « Faut-il un droit des robots ? (<https://www.latribune.fr/opinions/tribunes/faut-il-un-droit-des-robots-697651.html>) », sur *La Tribune*, 28 avril 2017 (consulté le 4 mars 2024).
231. « Règles de droit civil sur la robotique. Résolution du Parlement européen du 16 février 2017 contenant des recommandations à la Commission concernant des règles de droit civil sur la robotique (2015/2103(INL)) (<http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//NONSGML+TA+P8-TA-2017-0051+0+DOC+PDF+V0//FR>) ».
232. (en) Mark Coeckelbergh et Henrik Skaug Sætra, « Climate change and the political pathways of AI: The technocracy-democracy dilemma in light of artificial intelligence and human agency », *Technology in Society*, vol. 75, novembre 2023 (DOI 10.1016/j.techsoc.2023.102406 (<https://dx.doi.org/10.1016/j.techsoc.2023.102406>) ).
233. « Quelles différences entre l'IA symbolique et l'apprentissage automatique ? (<https://www.actuia.com/decouvrir/quelles-differences-entre-lia-symbolique-et-lapprentissage-automatique/>) », sur *Actuia*, 5 février 2023 (consulté le 17 mars 2024).
234. (en) John Eikenberry, *GNU/Linux AI & Alife HOWTO*, 2013 (lire en ligne (<http://www.ibiblio.org/pub/Linux/docs/HOWTO/other-formats/pdf/AI-Alife-HOWTO.pdf>) [PDF]).
235. André Le Garff, *Dictionnaire de l'informatique*, 1975, 576 p.
236. Voir la définition utilisée par John Searle dans son expérience de la chambre chinoise.
237. Michael David Mitchell, « Robotique et neurosciences détectent le siège de la conscience de soi (<http://actu.epfl.ch/news/robotique-et-neurosciences-detectent-le-siege-de-l/>) », *Actualités*, École polytechnique fédérale de Lausanne, 2 mai 2011.
238. (en) Hilary Putnam, « The Best of All Possible Brains? (<https://archive.nytimes.com/www.nytimes.com/books/97/04/27/nnp/17540.html>) », sur *archive.nytimes.com*, 20 novembre 1994 (consulté le 18 novembre 2023).
239. (en) « Discovery of quantum vibrations in 'microtubules' inside brain neurons supports controversial theory of consciousness (<https://www.sciencedaily.com/releases/2014/01/140116085105.htm>) », sur *ScienceDaily* (consulté le 6 juillet 2020).
240. « Le mystère de la "chambre chinoise" », *Le Monde*, 10 janvier 2013 (lire en ligne (https://www.lemonde.fr/science/article/2013/01/10/le-mystere-de-la-chambre-chinoise_1815124_1650684.html), consulté le 6 juillet 2020).
241. (en) Dreyfus, Hubert L., *What computers can't do : a critique of artificial reason*, Harper & Row, 1972 (ISBN 0-06-011082-1 et 978-0-06-011082-6, OCLC 257056 (<https://worldcat.org/fr/title/257056>)), p. 106.
242. (en) George Musser, « How AI Knows Things No One Told It (<https://www.scientificamerican.com/article/how-ai-knows-things-no-one-told-it/>) », sur *Scientific American*, 1^{er} septembre 2023 (consulté le 3 décembre 2023).
243. Jean-Gabriel Ganascia, « Génération automatique de textes, plagiat et intégrité scientifique (<https://www.larecherche.fr/ethique-chronique-%C3%A9thique/g%C3%A9n%C3%A9ration-automatique-de-textes-plagiat-et-int%C3%A9grit%C3%A9-scientifique>) », *La Recherche* (consulté le 3 décembre 2023).
244. (en) Shyam Nandan Upadhyay, « Text is a Projection of the World: OpenAI's Sutskever (<https://analyticsindiamag.com/text-is-a-projection-of-the-world-says-openais-sutskever/>) », sur *Analytics India Magazine*, 24 mars 2023 (consulté le 3 décembre 2023).
245. (en-US) « The Challenges Of Building AI Apps (<https://social.techcrunch.com/2015/10/15/machine-learning-its-the-hard-problems-that-are-valuable/>) », sur *TechCrunch* (consulté le 21 juillet 2020).
246. Abdel Belaïd, « La reconnaissance automatique de l'écriture (<https://www.pourlascience.fr/sd/informatique/la-reconnaissance-automatique-de-lecriture-4509.php>) », *Pour la science* (consulté le 21 juillet 2020).
247. Veronica Smink, « Les trois stades de l'intelligence artificielle : dans lequel nous nous trouvons et pourquoi beaucoup pensent que le troisième pourrait être fatal ? (<https://www.bbc.com/afrique/articles/cpw6lj4pqd0o>) », sur *BBC News Afrique*, 31 mai 2023 (consulté le 20 décembre 2023).
248. Martin Gibert, *Faire la morale aux robots : une introduction à l'éthique des algorithmes*, 2020, 95 p. (ISBN 978-2-89759-516-6 et 2-89759-516-7, OCLC 1146545386 (<https://worldcat.org/fr/title/1146545386>)).
249. « Intelligence : qu'est-ce que c'est ? (<https://www.futura-sciences.com/sante/definitions/corps-humain-intelligence-13498/>) », sur *Futura* (consulté le 17 mars 2024).
250. Michel Lévy-Provençal, « Se préparer à l'intelligence artificielle générale (<https://www.lesechos.fr/idees-debats/sciences-prospective/se-preparer-a-lintelligence-artificielle-generale-2026946>) », sur *Les Échos*, 6 novembre 2023 (consulté le 17 mars 2024).
251. Jean Rohmer, « Comprendre l'intelligence artificielle symbolique (<https://theconversation.com/comprendre-lintelligence-artificielle-symbolique-104033>) », sur *The Conversation*, 19 novembre 2018 (consulté le 2 novembre 2023).
252. « Quand l'écrivain de science-fiction Isaac Asimov prédisait le futur (<https://www.franceculture.fr/litterature/quand-l-ecrivain-de-science-fiction-isaac-asimov-predisait-le-futur>) », sur *France Culture*, 2 janvier 2020 (consulté le 7 juillet 2020).
253. (en) Georgia Panteli, « From puppet to cyborg: posthuman and postmodern retellings of the Pinocchio myth (http://discovery.ucl.ac.uk/id/eprint/1528658/13/Panteli_Georgia_PhD%20Thesis.pdf) » [PDF], sur *discovery.ucl.ac.uk*.

254. « Plumbing Stanley Kubrick (<https://web.archive.org/web/20080703134444/http://www.ianwatson.info/kubrick.htm>) », sur *ianwatson.info* via *web.archive.org*, 3 juillet 2008 (consulté le 7 juillet 2020).
255. Site présentant les épisodes de la série Derrick (<https://inspecteurderrick.com/>)
256. Site fernsehserien.de; résumé (en allemand) de l'épisode . "108. Dr. Römer und der Mann des Jahres" (<https://www.fernsehserien.de/derrick/folgen/108-dr-roemer-und-der-mann-des-jahres-64874>).
257. « Westworld (http://www.allocine.fr/series/fichserie_gen_cserie=16930.html) », sur *Allociné.fr* (consulté le 22 décembre 2016).
258. Article de Dylan L. McClain, *New York Times*, 16 janvier 2017, et article « Hans J. Berliner (1929-2017), Grand Maître d'échecs et programmeur de génie », Denis Rozier, *Le Courrier des Échecs* n° 639, mars 2017.
259. Thibault Neveu, « Intelligence artificielle : démonstration en direct des nouvelles performances de DeepMind (<http://www.actuia.com/actualite/intelligence-artificielle-demonstration-en-direct-des-nouvelles-performances-de-deep-mind/>) », sur *actuia.com*, 24 janvier 2019.
260. « Les jeux et l'intelligence artificielle (https://www.u-picardie.fr/~furst/docs/jeux_ia.pdf) », sur *u-picardie.fr* (consulté le 27 septembre 2016).
261. (en) David Levy et Monroe Newborn, « *More Chess and Computers: The Microcomputer Revolution, The Challenge Match* », *Computer Science Press*, Potomac (Maryland) et Batsford, Londres, 1980 (ISBN 0-914894-07-2).
262. (en) « *For First Time, a Chess Computer Outwits Grandmaster in Tournament* » (<https://www.nytimes.com/1988/09/26/nyregion/for-first-time-a-chess-computer-outwits-grandmaster-in-tournament.html>), Harold C. Schonberg, *The New York Times.com*, 26 septembre 1988.
263. (en) « *Learning, Deep or not Deep, de MENACE à AlphaGo* » (<http://www.alliot.fr/COURS/JEUX/learning.pdf>) [PDF], sur *alliot.fr*.
264. (en) David Silver, Thomas Hubert, Julian Schrittwieser, Ioannis Antonoglou, Matthew Lai et al., « *Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm* » (<https://arxiv.org/abs/1712.01815>) », 5 décembre 2017..
265. (en) Sarah Knapton et Leon Watson, « *Entire human chess knowledge learned and surpassed by DeepMind's AlphaZero in four hours* » (<https://www.telegraph.co.uk/science/2017/12/06/entire-human-chess-knowledge-learned-surpassed-deepminds-alphazero/>) », sur *telegraph.co.uk*, 6 décembre 2017.
266. « Jeu de go : comment savoir si les programmes d'intelligence artificielle sont vraiment... intelligents » (https://www.lemonde.fr/pixels/article/2016/03/09/comment-savoir-si-les-programmes-d-intelligence-artificielle-sont-vraiment-intelligents_4879177_4408996.html), *Le Monde*, 9 mars 2016.
267. (en) David Silver, Julian Schrittwieser, Karen Simonyan, Ioannis Antonoglou, Aja Huang, Arthur Guez, Thomas Hubert, Lucas Baker, Matthew Lai, Adrian Bolton, Yutian Chen, Timothy Lillicrap, Hui Fan, Laurent Sifre, George van den Driessche, Thore Graepel et Demis Hassabis, « *Mastering the game of Go without human knowledge* », *Nature*, vol. 550, n° 7676, 19 octobre 2017, p. 354-359 (ISSN 0028-0836 (<https://portal.issn.org/resource/issn/0028-0836>), DOI 10.1038/nature24270 (<https://dx.doi.org/10.1038/nature24270>) ).
268. « Jeu de go : Le champion du monde de nouveau battu par l'ordinateur de Google » (<http://tempsreel.nouvelobs.com/culture/20160310.AFP9477/jeu-de-go-le-champion-du-monde-de-nouveau-battu-par-l-ordinateur-de-google.html>), *L'Obs*, 10 mars 2016.
269. (en) J. Rehmeier, N. Fox et R. Rico, « *Ante up, human: The adventures of Polaris the poker-playing robot* », *Wired*, vol. 16, n° 12, décembre 2008, p. 86–191.
270. (en) Michael Bowling, Neil Burch, Michael Johanson et Oskari Tammelin, « Heads-Up Limit Hold'em Poker Is Solved », *Comm. of the ACM*, vol. 60, n° 11, 2017, p. 81.
271. « Comment une intelligence artificielle ridiculise les meilleurs joueurs de poker » (<https://www.numerama.com/tech/27764-comment-une-intelligence-artificielle-ridiculise-les-meilleurs-joueurs-de-poker.html>) », sur *Numerama*, 27 janvier 2017.
272. « L'IA de la start-up française Nukkai imbattable au bridge » (<https://www.lemondeinformatique.fr/actualites/lire-l-ia-d-e-la-start-up-francaise-nukkai-imbattable-au-bridge-86251.html>) », *Le Monde informatique*, 26 mars 2022 (consulté le 28 janvier 2024).

Voir aussi



Une catégorie est consacrée à ce sujet :
Intelligence artificielle.

Sur les autres projets Wikimedia :



intelligence artificielle, sur le Wiktionnaire



Département: Intelligence artificielle,
sur Wikiversity

Articles connexes

Aspects juridiques

- Législation sur l'intelligence artificielle, aussi appelée *AI Act*
- Digital Services Act, ou loi sur les services numériques de l'Union européenne

Notions générales

- Agent intelligent
- Agent logiciel
- Agent virtuel
- Algorithme
- Algorithme génétique
- Alignement des intelligences artificielles
- Applications de l'intelligence artificielle
- Automation
- Bio-informatique
- Cerveau artificiel
- Cyborg
- Déclaration de Montréal pour un développement responsable de l'intelligence artificielle
- Sommet pour l'action sur l'intelligence artificielle
- Deepfake
- Effet IA
- Éthique de l'intelligence artificielle
- Explosion d'intelligence
- Histoire de l'intelligence artificielle
- Informatique de la personnalité (Personality computing)
- Informatique quantique
- Intelligence artificielle de remédiation
- Intelligence artificielle générale
- Interactions homme-machine
- Philosophie de l'intelligence artificielle
- Point de vue d'Hubert Dreyfus sur l'intelligence artificielle
- Principaux projets et réalisations en intelligence artificielle
- Progrès
- Progrès technique
- Réseau de neurones artificiels
- Singularité technologique
- Singularitarisme
- Système expert
- Téléchargement de l'esprit
- Test de Turing
- Vie artificielle

Notions techniques

- Agent conversationnel
- Apprentissage automatique
- Apprentissage par renforcement
- Apprentissage profond (*Deep learning*)

- [Architecture cognitive](#)
- [Diagnostic](#)
- [Exploration de données](#)
- [Forêt d'arbres décisionnels](#)
- [Inférence bayésienne](#)
- [Intelligence artificielle amicale](#)
- [Intelligence artificielle distribuée](#)
- [Intelligence artificielle faible](#)
- [Intelligence artificielle dans la fiction](#)
- [Logique floue](#)
- [Machine à vecteurs de support](#)
- [Métaheuristiques](#)
- [Planification](#)
- [Problème de satisfaction de contraintes](#)
- [Programmation génétique](#)
- [Programmation par contraintes](#)
- [Raisonnement par cas](#)
- [Réseaux de neurones](#)
- [Système multi-agents](#)
- [Théorème de Cox-Jaynes](#)

Chercheurs en intelligence artificielle (espace anglophone)

- [Edward Feigenbaum](#)
- [Irving John Good](#)
- [Douglas Engelbart](#)
- [Douglas Hofstadter](#)
- [Douglas Lenat](#)
- [John McCarthy](#)
- [Marvin Lee Minsky](#)
- [Allen Newell](#)
- [Nils Nilsson](#)
- [Seymour Papert](#)
- [Rosalind Picard](#)
- [Roger Schank](#)
- [Herbert Simon](#)
- [Ray Solomonoff](#)
- [Gerald Jay Sussman](#)
- [Alan Turing](#)
- [Joseph Weizenbaum](#)

Chercheurs en intelligence artificielle (espace francophone)

- [Bubacarr Bah](#)
- [Hugues Bersini](#)
- [Alain Colmerauer](#)
- [Jean-Paul Delahaye](#)
- [Rose Dieng-Kuntz](#)
- [Yann Le Cun](#)
- [François Pachet](#)
- [Jacques Pitrat](#)
- [Gérard Sabah](#)

Laboratoires et entreprises en intelligence artificielle

- [OpenAI](#), société développant [ChatGPT](#) et [DALL-E](#).
- [Google DeepMind](#)
- [Anthropic](#)

- Meta AI
- Inflection AI

Bibliographie

 : document utilisé comme source pour la rédaction de cet article.

Aspects techniques

- Stuart J. Russell et Peter Norvig (trad. de l'anglais), *Intelligence artificielle* [« Artificial Intelligence: A Modern Approach »], Paris, Pearson Education France, 2010, 3^e éd., 1198 p. (ISBN 978-2-7440-7455-4, [présentation en ligne \(https://books.google.fr/books?id=DWTIFWSGxJMC\)](https://books.google.fr/books?id=DWTIFWSGxJMC)).
- Alan Turing, Jean-Yves Girard, *La Machine de Turing*, Éditions du Seuil, 1995 [détail de l'édition], *Les Ordinateurs et l'Intelligence*, p. 133–174.
- Claire Rémy, *L'Intelligence artificielle*, Paris, Éditions Dunod, 1994, 158 p. (ISBN 2-10-002258-X).
- Jean-Marc Alliot et Thomas Schiex, *Intelligence artificielle et informatique théorique*, Toulouse, CEPADUES, 2002, 543 p. (ISBN 2-85428-578-6).
- (en) Michael R. Genesereth et Nils J. Nilsson, *Logical Foundations of Artificial Intelligence*, Los Altos, Californie, États-Unis, Morgan Kaufmann, 1987, 405 p. [détail de l'édition] (ISBN 0-934613-31-1).
- Jean-Louis Laurière, *Intelligence Artificielle*, Eyrolles, 1986.
- Jean-Paul Delahaye, *Outils logiques pour l'intelligence artificielle*, Eyrolles, 1985 [détail des éditions] (ISBN 978-2212084122).
- Jean-Paul Haton, Marie-Christine Haton, *L'Intelligence Artificielle*, Paris, Que sais-je ?, 1990, 127 p. (ISBN 2-13-043164-X).
- (en) Kate Crawford, *Atlas of AI : Power, Politics, and the Planetary Costs of Artificial Intelligence*, Yale University Press, 2021.
- (en) Jacques Pitrat Artificial Beings, the conscience of a conscious machine, *Artificial Beings, the conscience of a conscious machine*, ISTE-Wiley, 2019 (ISBN 9781848211018).

Aspects prospectifs

- *Intelligence artificielle : État de l'art et perspectives pour la France*, Direction générale des entreprises, février 2019, pdf (ISBN 978-2-11-152634-1, ISSN 2491-0058 (<https://portal.issn.org/resource/issn/2491-0058>), résumé (<https://www.entreprises.gouv.fr/la-dge/publications/intelligence-artificielle-etat-de-lart-et-perspectives-pour-la-france>)).
- « Jusqu'où ira l'intelligence artificielle ? », *Pour la science*, hors-série n° 115, mai-juin 2022, p. 4-119.
- Thibault Prévost, *Les prophètes de l'IA - Pourquoi la Silicon Valley nous vend l'apocalypse*, Montréal (Québec), Lux, 2025, 216 p. (ISBN 9782898331602).

Aspects philosophiques

- Marie David et Cédric Sauviat, *Intelligence artificielle, la nouvelle barbarie*, Éditions du Rocher, 2019, 313 p.
- Laurent Alexandre, *La Guerre des intelligences. Intelligence artificielle versus intelligence humaine*, Paris, Éditions Jean-Claude Lattès, 2017, 250 p. (ISBN 978-2-7096-6084-6, [lire en ligne \(https://books.google.com/books?id=3qg0DwAAQBAJ\)](https://books.google.com/books?id=3qg0DwAAQBAJ)).
- Gilbert Boss, *Les machines à penser : L'homme et l'ordinateur*, Zurich, Éditions du Grand midi, 1987, 202 p. (ISBN 2-88093-105-3).
- Jacques Bolo, *Philosophie contre intelligence artificielle*, Paris, Lingua Franca, 1996, 375 p. (ISBN 2-912059-00-3).
- Alan Ross Anderson (trad. de l'anglais), *Pensée et machine*, Seyssel, Éditions Champ Vallon, 1983 (réimpr. 1993), 150 p. (ISBN 2-903528-28-4, [lire en ligne \(https://books.google.com/books?id=Aab_kiv_DwMC&printsec=frontcover\)](https://books.google.com/books?id=Aab_kiv_DwMC&printsec=frontcover)).
- Jean Sallantin et Jean-Jacques Szczeciniarz, *Le Concept de preuve à la lumière de l'intelligence artificielle*, Paris, Presses universitaires de France, 1999, 370 p. (ISBN 2-13-050104-4).
- Jean-Gabriel Ganascia, *L'Âme-machine, les enjeux de l'intelligence artificielle*, Paris, Éditions du Seuil, Collection Science Ouverte, 1990, 284 p. (ISBN 2-02-011470-4).

- Nick Bostrom, *Superintelligence : Paths, Dangers, Strategies*, 2014, 328 p. (ISBN 978-0-19-967811-2, lire en ligne (http://books.google.com/books?id=7_H8AwAAQBAJ&printsec=frontcover)).

Aspects économiques

- (en) Daron Acemoğlu « Equality Debate: Power and Progress, with Daron Acemoglu » (8 novembre 2023) (lire en ligne (<https://www.youtube.com/watch?v=XBQEjfmW2w>))  — World Inequality Lab, (visioconférence, 6 novembre 2023).
- (en) Daron Acemoğlu et Simon Johnson, *Power and Progress : Our Thousand-Year Struggle Over Technology and Prosperity*, Basic Books, 2023, 560 p. (ISBN 978-1399804462).
- (en) Daron Acemoğlu, « The Simple Macroeconomics of AI », *Economic Policy*, Massachusetts Institute of Technology, 5 avril 2024 (lire en ligne (<https://economics.mit.edu/sites/default/files/2024-04/The%20Simple%20Macroeconomics%20of%20AI.pdf>))  [PDF]).
- Abhijit Banerjee et Esther Duflo (trad. Christophe Jacquet), chap. 7 « Le piano mécanique », dans *Économie utile pour des temps difficiles*, Éditions du Seuil, 2020.
- (en) Joseph Briggs et Devesh Kodnani, « The Potentially Large Effects of Artificial Intelligence on Economic Growth », *Global Economics Analyst*, Goldman Sachs, 26 mars 2023 (lire en ligne (<https://www.gspublishing.com/content/research/en/reports/2023/03/27/d64e052b-0f6e-45d7-967b-d7be35fabd16.html>)) .
- (en) Mauro Cazzaniga, Florence Jaumotte, Longji Li, Giovanni Melina, Augustus Panton, Carlo Pizzinelli, Emma Rockall et Marina Tavares, « Gen-AI: Artificial Intelligence and the Future of Work (<https://www.imf.org/en/Publications/Staff-Discussion-Notes/Issues/2024/01/14/Gen-AI-Artificial-Intelligence-and-the-Future-of-Work-542379>) »  [PDF], sur *Fonds monétaire international*, 14 janvier 2024.
- (en) Michael Chui, James Manyika et Mehdi Miremadi, « Four fundamentals of workplace automation (<https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/four-fundamentals-of-workplace-automation>) » , sur *McKinsey*, 1^{er} novembre 2015.
- Kristalina Georgieva, « L'IA transformera l'économie mondiale. Faisons en sorte que l'humanité y soit gagnante », *Fonds monétaire international*, 14 janvier 2024 (lire en ligne (<https://www.imf.org/fr/Blogs/Articles/2024/01/14/ai-will-transform-the-global-economy-lets-make-sure-it-benefits-humanity>)) .
- (en) Joel Mokyr, « Secular stagnation? Not in your life », *VoxEU Column*, Centre for Economic Policy Research, 11 août 2014 (lire en ligne (<https://cepr.org/voxeu/columns/secular-stagnation-not-your-life>)) .

Fondements cognitifs, psychologiques et biologiques

- Hervé Chaudet et Liliane Pellegrin, *Intelligence artificielle et psychologie cognitive*, Paris, Éditions Dunod, 1998, 179 p. (ISBN 2-10-002989-4).

Aspects linguistiques

- Gérard Sabah, *L'Intelligence artificielle et le langage, Représentations des connaissances, Processus de compréhension*, vol. 1, Hermès, 1989, 768 p. (ISBN 2-86601-134-1).
- Gérard Sabah, *L'Intelligence artificielle et le langage, Représentations des connaissances, Processus de compréhension*, vol. 2, Paris, Hermès, 1990, 768 p. (ISBN 2-86601-187-2).
- Gérard Sabah, *Compréhension des langues et interaction (Traité IC2, Série Cognition et Traitement de l'Information)*, Paris, Hermès science: Lavoisier, 2006, 400 p. (ISBN 2-7462-1256-0).
- (en) Krzysztof Wołk, *Machine learning in translation corpora processing*, Boca Raton, FL, Taylor & Francis, 2019, 264 p. (ISBN 978-0-367-18673-9)

Histoire

- Daniel Crevier et Nathalie Bukcek (trad. de l'anglais), *À la recherche de l'intelligence artificielle*, Paris, Flammarion, 1997, 438 p. (ISBN 2-08-081428-1) (traduction de (en) *The Tumultuous history of the search for artificial intelligence*).
- Nicolas Neysen, *Métamorphoses numériques : Explorations sur l'intelligence artificielle, le métavers, la blockchain et le marketing d'influence*, Presses universitaires de Liège, 2025
- Yuval Noah Harari, *Nexus. Une brève histoire des réseaux d'information de l'âge de pierre à l'IA.*, Albin Michel, 2024, 567 p. (ISBN 978-2-226-49488-7). 

Vulgarisation

- Gérard Tisseau et Jacques Pitrat, *Intelligence artificielle : problèmes et méthodes*, Paris, Presses universitaires de France, 1996, 255 p. (ISBN 2-13-047429-2).

- Jack Challoner (trad. de l'anglais), *L'Intelligence artificielle : Un guide d'initiation au futur de l'informatique et de la robotique*, Paris, Pearson Education, 2003, 72 p. (ISBN 2-7440-1600-4).
- Hugues Bersini, *De l'intelligence humaine à l'intelligence artificielle*, Paris, Ellipse, 2006, 192 p. (ISBN 2-7298-2813-3).
- Jean-Gabriel Ganascia, *L'Intelligence artificielle*, Paris, Éditions du Cavalier bleu, coll. « Idées reçues », 2007, 127 p. (ISBN 978-2-84670-165-5).
- Howard Selina (illustrations) et Henry Brighton (texte) (trad. de l'anglais), *L'Intelligence artificielle en images*, Les Ulis, EDP Sciences, coll. « Aperçu », 2015, 176 p. (ISBN 978-2-7598-1772-6).
- Marion MONTAIGNE (dessin) et Jean-Noël Lafargue (scénario), *L'Intelligence artificielle : fantasmes et réalités*, Bruxelles, Le Lombard, coll. « La petite bédéthèque des savoirs », 2016, 72 p. (ISBN 978-2-8036-3638-9).
- Arnold Zéphir (scénariste), FibreTigre (scénariste) et Héloïse Chochois, *Intelligences Artificielles : Miroirs de nos vies*, Delcourt, coll. « Octopus », 6 mars 2019 (ISBN 978-2413013143)
- [vidéo] « L'IA et l'efficacité (<https://savoir.media/lia-et-moi/clip/lia-et-lefficacite>) » sur *Savoir média*, 2021, 30 min (consulté le 31 mai 2022)

Politique, relations internationales

- Daniel Ventre, *Intelligence artificielle, cybersécurité et cyberdéfense*, Londres, ISTE, 2020, 246 p. (ISBN 978-1-78405-679-7 et 978-1-78406-679-6, SUDOC 248481614 (<https://www.sudoc.fr/248481614>), présentation en ligne (<https://www.istegroup.com/fr/produit/intelligence-artificielle-cybersecurite-et-cyberdefense/>)).

Liens externes

- « IA militaire : drones de guerre » (<https://www.radiofrance.fr/franceculture/podcasts/la-science-cqfd/intelligence-artificielle-et-armee-5194965>), *La Science, CQFD*, France Culture, 25 juin 2024.
- « Intelligence artificielle (<https://interstices.info/domaine/intelligence-artificielle/>) », publications en ligne, *Interstices*, INRIA.
- (en) European Association for Artificial Intelligence (EurAI) (<https://www.eurai.org>) (Association européenne pour l'intelligence artificielle)
- Association française pour l'intelligence artificielle (<http://www.afia.asso.fr>) (AfIA).
- GDR-Radia (<https://gdr-radia.cnrs.fr/>), groupement de recherche du CNRS sur les aspects formels et algorithmiques de l'intelligence artificielle.
- Dossier sur l'Intelligence artificielle (http://savoirs.ens.fr/focus_detail.php?id=52), savoirs.ens.fr (conférences), École normale supérieure.
- Sébastien Konieczny, « L'intelligence artificielle, menace ou avancée ? » (https://www.huffingtonpost.fr/sebastien-konieczny/intelligence-artificielle-google_b_9408180.html), *Huffington Post*, 9 mars 2016
- « Intelligence artificielle : en mon âme et conscience » (<https://www.radiofrance.fr/franceculture/podcasts/la-science-cqfd/conscience-artificielle-4375043>), *La Science, CQFD*, France Culture, 1^{er} juin 2023.
- William Audureau, « Tout comprendre à l'intelligence artificielle, cette technologie source de nombreux malentendus (https://www.lemonde.fr/les-decodeurs/article/2024/04/20/tout-comprendre-a-l-intelligence-artificielle-cette-technologie-source-de-nombreux-malentendus_6228954_4355770.html) », sur *Les Décodeurs*, Le Monde, 20 avril 2024

Bases de données et dictionnaires

-
- Ressources relatives à la recherche : Internet Encyclopedia of Philosophy (<https://www.iep.utm.edu/art-intel/>) · JSTOR (<https://www.jstor.org/topic/artificial-intelligence>) · *Stanford Encyclopedia of Philosophy* (<https://plato.stanford.edu/entries/artificial-intelligence/>)
- Ressource relative à la santé : Medical Subject Headings (<https://meshb.nlm.nih.gov/record/ui?ui=D001185>)
- Ressource relative à l'audiovisuel : France 24 (<https://www.france24.com/fr/tag/intelligence-artificielle/>)
- Notices dans des dictionnaires ou encyclopédies généralistes : *Britannica* (<https://www.britannica.com/technology/artificial-intelligence>) · *Brockhaus* (<https://brockhaus.de/ecs/enzy/article/kunstliche-intelligenz>) · *Den Store Danske Encyklopædi* (https://denstoredanske.lex.dk/kunstig_intelligens/) · *Enciclopedia De Agostini* (<http://www.sapere.it/enciclopedia/intellig%C3%A8nza%2Bartificiale.html>) · *Internetowa encyklopedia PWN* (<https://encyklopedia.pwn.pl/haslo/3983490>) · *Larousse* (https://www.larousse.fr/encyclopedia/divers/intelligence_artificielle/187257) · *Nationalencyklopedin* (<https://www.ne.se/uppslagsverk/encyklopedi/lång/artificiell-intelligens>) · *Store norske leksikon* (https://snl.no/kunstig_intelligens) · *Treccani* (<http://www.treccani.it/enciclopedia/intelligenza-artificiale>)

- Notices d'autorité : BnF (<https://catalogue.bnf.fr/ark:/12148/cb11932084t>) (données (<https://data.bnf.fr/ark:/12148/cb11932084t>)) • LCCN (<http://id.loc.gov/authorities/sh85008180>) • GND (<http://d-nb.info/gnd/4033447-8>) • Japon (<https://id.ndl.go.jp/auth/ndlna/00574798>) • Espagne (http://catalogo.bne.es/uhtbin/authoritybrowse.cgi?action=display&authority_id=XX4659822) • Israël (http://olduli.nli.org.il/F/?func=find-b&local_base=NLX10&find_code=UID&request=987007294969105171) • Tchéquie (https://aleph.nkp.cz/F/?func=find-c&local_base=aut&ccl_term=ica=ph116536) • Lettonie (https://kopkatalogs.lv/F/?func=direct&local_base=lncl10&doc_number=000050010)
-