

DOC EN POCHE
ENTREZ DANS L'ACTU

RODOLPHE GELIN
OLIVIER GUILHEM

L'intelligence artificielle en 30 questions

La Documentation
française



Parlons IA

Sommaire

Le point sur 7

Questions-réponses 21

1. Qu'est-ce que l'intelligence artificielle ? 22
2. Quels sont les différents types d'intelligence artificielle ? 24
3. Quelle différence entre l'IA et l'informatique ? 26
4. Qu'est-ce qu'un réseau de neurones en informatique ? 28
5. Quelles différences entre IA, algorithme et logiciel ? 30
6. Comment fonctionne l'apprentissage d'un réseau de neurones ? 32
7. Quelles sont les différentes techniques d'apprentissage ?... 34
8. Quel est le processus de création d'une IA ? 36
9. Quelles ont été les innovations majeures qui ont accéléré l'essor de l'IA ? 38
10. Quels sont les enjeux liés aux données utilisées par l'IA ? .. 40
11. Quelles sont les limites actuelles de l'IA ? 42
12. L'IA peut-elle commettre des erreurs ? 44
13. Comment rendre une IA plus fiable ? 46
14. Comment fonctionnent les IA génératives ? 48
15. Est-ce que ChatGPT est une révolution en IA ? 50
16. L'IA apprend-elle toute seule ? 52
17. L'IA a-t-elle des émotions ? 54
18. L'IA a-t-elle une personnalité ? 56
19. À quoi servent les intelligences artificielles ? 58
20. Quels sont les liens entre IA et robotique ? 60
21. Quels sont les dangers liés à l'IA ? 62
22. Quels sont les enjeux sociétaux liés à l'IA ? 64
23. L'IA peut-elle servir à créer des armes autonomes ? 66

24. L'IA peut-elle être source de fausses informations ?	68
25. L'IA est-elle plus intelligente que l'homme ?	70
26. L'IA a-t-elle un impact sur l'emploi ?	72
27. L'IA est-elle éthique ?	74
28. Quel est l'impact énergétique de l'IA ?	76
29. L'IA peut-elle être auteure d'une œuvre artistique protégeable ?	78
30. L'IA est-elle juridiquement encadrée ?	80

Pour aller plus loin	83
Les chiffres-clés	84
Les mots-clés	86
Bibliothèque	90

Le point sur

Entre le moment où nous écrivons ce livre et celui où vous le lirez, les géants de la « Tech » auront certainement présenté de nouveaux systèmes d'IA qui ridiculiseront le GPT 4o, révélé par OpenAI en mai 2024. Pourtant, cette IA, capable de discuter avec une voix très expressive de tous les sujets et d'interpréter des vidéos relevant, il y a peu, de la science-fiction. Si les techniques évoluent à grande vitesse, les principes que nous présentons ici resteront des repères utiles pour mieux comprendre cette formidable technologie.

Une technologie de pointe à la portée de tous

Ces dernières années, l'intelligence artificielle était dans tous les esprits en grande partie à cause des fantasmes qu'entretenaient des œuvres de fiction autour de la prise du pouvoir par les machines mais aussi parce que, dans l'industrie et la recherche, les progrès du *machine learning*, des « réseaux de neurones » et du *deep learning*, donnaient accès à des résultats jusque-là inespérés dans divers domaines : l'exploration spatiale, l'interprétation des images médicales, le déchiffrement de manuscrits anciens, l'amélioration de la précision des prévisions météorologiques. Toutes ces applications semblaient concerner des systèmes complexes pour lesquels l'informatique permettait au cerveau humain de résoudre des problèmes qui dépassaient ses capacités. L'IA était un outil sophistiqué, manipulé par des spécialistes pour traiter des problèmes très pointus.

* Les termes marqués par un astérisque sont définis dans *Les mots-clés* page 86.

La révélation de ChatGPT, en novembre 2022, a changé la nature de l'engouement pour cette technologie. En réalisant que l'IA « parlait » comme nous, comprenait (quasiment) comme nous, le grand public a pris conscience des capacités de l'IA. Curieusement, c'est quand l'IA s'est mise à faire ce que chacun sait faire (comprendre des requêtes exprimées en langage naturel et les exécuter) que le plus grand nombre a évalué le potentiel de cette technologie pourtant quasiment octogénaire. Du fait de sa grande facilité d'utilisation et de la capacité de chacun à juger de la pertinence de ses réponses, ChatGPT a permis à l'IA d'acquérir la notoriété des autres grandes innovations technologiques de notre histoire : le feu, la roue, l'électricité, Internet.

Pourtant, avant même l'arrivée de ChatGPT, l'intelligence artificielle était déjà omniprésente dans notre vie mais, à l'instar de l'électricité, nous ne la manipulons pas en tant que telle mais accédions à des services qui eux-mêmes l'utilisaient. Les moteurs de recherche sur Internet, les filtres antispams de nos boîtes courriel, les systèmes de recommandation de produits sur les sites de vente en ligne, les logiciels de navigation dans nos véhicules, la reconnaissance vocale qui permet de dicter des SMS à haute voix, la reconnaissance de visage qui déverrouille nos smartphones... tous ces services du quotidien, avec lesquels nous vivons depuis des années, reposent sur des algorithmes d'IA. ChatGPT n'est donc qu'une nouvelle expression de l'IA mais elle est mise à notre disposition sans véritable usage. *A priori*, elle ne fait qu'écrire des textes en réponse à des demandes. Chacun peut utiliser cette capacité brute à en faire ce qu'il veut : écrire du code informatique, résumer des articles, composer des chansons, comparer des textes... C'est un peu la même différence qu'entre donner un poisson à quelqu'un ou lui enseigner la pêche.

Les saisons de l'intelligence artificielle

L'IA est née dans les années 1950 (*cf. question 2*) et les principes qu'elle se proposait de développer étaient prometteurs. Les militaires attendaient notamment des applications essentielles comme la traduction automatique. Mais les techniques de l'époque ne permettaient pas d'atteindre les résultats espérés. Les financements se sont taris et, entre 1974 et 1980, les développements dans le domaine de l'IA se sont quasiment arrêtés. Ce fut le premier hiver de l'IA. Au début des années 1980, l'apparition des systèmes experts relance l'intérêt pour l'intelligence artificielle, notamment dans les entreprises. Si cette technique est toujours utilisée, elle ne se cantonne qu'à résoudre des problèmes relativement simples et ses limites inhérentes vont provoquer une nouvelle pause dans son essor. C'est son second hiver, qui durera de 1987 à 1993. Avec l'arrivée d'Internet, la multiplication des données disponibles et l'augmentation des puissances de calcul des ordinateurs, des techniques d'IA basées sur les données, qui avaient été imaginées dans les années 1950 mais n'avaient pas pu être mises en œuvre, vont pouvoir se déployer. L'ère du *machine learning* a commencé et l'IA connaît un été ininterrompu depuis.

Un des premiers événements qui a fait réaliser au grand public l'avènement de l'intelligence artificielle est la défaite de Garry Kasparov, champion du monde d'échecs, lors de son match contre l'ordinateur Deep Blue, d'IBM en 1997. Le jeu d'échecs est communément reconnu comme requérant une grande intelligence. Le fait qu'un ordinateur puisse battre un champion d'échecs était le signal que, au moins pour cet aspect de l'intelligence, la machine était plus performante que l'homme. Sentant bien que cette conclusion avait quelque chose de mortifère pour

l'avenir de l'humanité, une partie des personnes bien informées affirmèrent que le jeu d'échecs était finalement assez simple et que la machine n'avait pas eu beaucoup de mérite à gagner. Ils ajoutaient qu'avec un jeu vraiment compliqué et stratégique, comme le go, cela n'arriverait jamais. En 2017, le programme AlphaGo, de DeepMind, devient champion du monde de go, battant les meilleurs joueurs du monde, en jouant des coups improbables et même déconseillés dans les écoles de go. La conclusion de cette escapade de l'IA dans le monde des jeux de société est qu'il vaut mieux éviter de dire que quelque chose est impossible à l'IA.

De l'IA spécialisée à l'IA multimodale et polyvalente voire générale ?

Une limitation rassurante souvent avancée est celle de la spécialisation de l'IA. Une IA entraînée à reconnaître des images de chatons ne saurait reconnaître des papillons. Une IA spécialisée dans la vision ne saurait être compétente dans le traitement du langage. Cette présomption fut longtemps exacte. Elle l'est moins aujourd'hui. L'apparition de très larges modèles d'IA utilisant des centaines de milliards de paramètres, entraînés de nombreuses fois sur des jeux de données colossaux, rend l'IA plus robuste et polyvalente. L'IA est aussi devenue multimodale (combinant texte, image, audio) et interopérable, ce qui décroïsonne ses applications et ses possibles.

1 | Qu'est-ce que l'intelligence artificielle ?

Un concept valise

Le terme d'« intelligence artificielle » a été proposé en 1956 par John McCarthy du MIT. Son collègue Marvin Minsky expliquait que c'était « 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 ». L'intelligence artificielle est devenue depuis un concept valise qui recouvre énormément de choses. Certaines définitions très englobantes couvrent quasiment tous les logiciels alors que d'autres sont plus pointues.

Une modélisation comme aide à la décision

L'objectif est que les programmes informatiques fassent autre chose que des opérations sur des nombres (multiplications, additions, soustractions...) et qu'ils traitent et modélisent également des concepts comme le fait le cerveau humain. Dans la mémoire d'un ordinateur, ces manipulations permettent d'évaluer une situation, de prendre des décisions, et de proposer un plan d'action afin de résoudre un problème.

Des techniques pour interpréter, comprendre, traiter

En termes d'usage, on peut retenir que les techniques d'IA sont aujourd'hui les seules à pouvoir interpréter le contenu d'une image, comprendre un texte et traiter les explosions combinatoires. Même si elles peuvent faire bien autre chose, ce sont sur ces trois sujets qu'elles règnent désormais sans partage.



L'explosion combinatoire : l'exemple du jeu d'échecs

Lorsque le nombre de cas possibles à étudier devient difficile à dénombrer, on parle d'explosion combinatoire. Par exemple, au jeu d'échecs, le premier joueur a 20 coups possibles, le second joueur en a 20 également. Après le premier tour, il existe ainsi 400 configurations d'échiquier possibles. À la fin du deuxième tour, il y en aura environ 400×400 , soit 160 000, et ainsi de suite. Un programme d'échecs va donc gérer cette multitude de cas possibles avant de prendre une décision.



Régler des problèmes complexes : l'exemple de la cuisine

La différence entre un logiciel « conventionnel » et un logiciel d'IA est liée à la nature des problèmes gérés. Si l'on compare un programme classique à une recette de cuisine, le logiciel donne les instructions pour passer des données d'entrée (les ingrédients) à une donnée de sortie (le plat prêt à être mangé). C'est utile mais le plus compliqué est de répondre à la question « qu'est-ce que je vais faire à manger ce soir ? » C'est pour ce type de problème qu'il est possible de se faire aider par une intelligence artificielle : on peut lui demander de comparer tous les ingrédients disponibles à toutes les recettes connues dans les livres de cuisine pour en extraire celles qui sont faisables avec ce que l'on a. Ce qui n'est pas loin de relever de l'explosion combinatoire.



Une définition officielle

Le *Journal officiel* définit l'IA comme le « champ interdisciplinaire théorique et pratique qui a pour objet la compréhension de mécanismes de la cognition et de la réflexion, et leur imitation par un dispositif matériel et logiciel, à des fins d'assistance ou de substitution à des activités humaines ».

7 | Quelles sont les différentes techniques d'apprentissage ?

L'apprentissage supervisé

L'apprentissage est dit supervisé car la réponse attendue pour chaque question posée est connue. Il y a donc un moyen de s'assurer que le réseau donne ou non la réponse correcte. L'opération qui consiste à associer la bonne réponse à la question posée est appelée l'annotation. C'est une partie très lourde de ce type d'apprentissage car il faut un grand nombre de données annotées, généralement par des humains.

L'apprentissage non supervisé

Les données annotées sont parfois difficiles à collecter, notamment quand il existe peu d'exemples de ce qu'on cherche ou qu'on ne sait pas ce que l'on cherche. L'application la plus répandue en est la constitution de groupes d'éléments similaires dans une population. L'apprentissage non supervisé consiste à identifier les caractéristiques qui permettent de constituer des groupes homogènes. À l'issue de l'apprentissage, les critères pour être associé à un groupe ou à un autre sont établis.

L'apprentissage semi-supervisé

Ce type d'apprentissage utilise des données pour lesquelles les données annotées sont très simples à construire. En effet, au lieu de chercher la réponse à une question donnée, on va ici construire la question à partir d'une réponse connue. L'apprentissage sera supervisé, puisque la question et la réponse sont connues, mais la phase d'annotation des données d'entrée, indispensable à l'apprentissage supervisé, sera économisée.



Reconnaître des animaux

Pour apprendre à une IA à identifier les animaux représentés sur une photo, il faut lui donner des milliers de photos d'animaux annotées. Pour la phase d'apprentissage, si l'entrée du système est une photo du cheval, la sortie est le mot « cheval ». Les données utilisées pour l'apprentissage sont des milliers de photos de chevaux associées au mot « cheval », des milliers de photos de chiens associées au mot « chien » et ainsi de suite. Après cet apprentissage, l'IA sera capable de reconnaître tout animal vu pendant l'apprentissage.



La catégorisation des clients

Chaque consommateur fréquentant les boutiques en ligne a pu expérimenter l'efficacité de l'apprentissage non supervisé. Lorsque, à l'issue d'un achat, des recommandations pour d'autres produits nous sont faites, c'est parce que nous avons été identifiés comme faisant partie d'un groupe de clients ayant les mêmes appétences pour certains articles. L'IA identifie, sans supervision d'un expert du marketing, un groupe de clients qui se ressemblent parce qu'ils ont acheté le dernier Astérix, un disque de Stromae et un grille-pain.



Boucher les trous dans les images

Sur les téléphones portables, une application permet de retirer un baigneur inconnu d'une belle photo de coucher de soleil sur une plage. L'IA remplit le vide avec le bout de plage, de mer et de ciel rougi. L'apprentissage d'un tel système est fait de façon semi-supervisée. On retire une partie d'une image de paysage. L'image « percée » est utilisée comme entrée et l'image intacte est utilisée comme vérité terrain. Entraînée avec des millions d'exemples comme celui-ci, l'IA sera capable de boucher les trous dans toutes les images.

20 | Quels sont les liens entre IA et robotique ?

Le robot pour compenser les faiblesses de l'IA

L'une des plus grandes déficiences de l'IA est son rapport au monde. L'utilité et la valeur ajoutée de l'IA ne se font ressentir que par ses interactions avec son environnement. Dans le même temps, elle est dépourvue d'un corps qui lui faciliterait ces interactions. Le robot est donc le vecteur idéal par lequel elle peut agir et évoluer dans le réel. La mobilité, la capacité de détecter un obstacle, à prendre ou à déplacer des objets sont des expériences lui permettant une meilleure compréhension de l'univers qui l'entoure, un apprentissage que seul un robot peut lui offrir. Cette confrontation tangible lui permet ainsi de collecter ce dont elle a le plus besoin pour évoluer et s'améliorer, la donnée.

L'IA pour compenser les faiblesses du robot

Il est toujours possible d'embarquer dans un robot de « simples » logiciels. Le robot ainsi programmé pourrait, sans IA, se mouvoir d'un point A à un point B, monter un escalier ou bien prendre une vis et la déposer dans une boîte. Les choses se compliquent si un obstacle survient durant l'action. Le programmeur peut toujours tenter d'anticiper la survenance de cet obstacle mais jusqu'à quel point ? Le salut face à une telle complexité ? L'intelligence artificielle. Sa capacité en temps réel à traiter un nombre incommensurable de données et à réagir en conséquence en fait l'outil parfait. Connectée à l'ensemble des capteurs du robot (température, mouvement...), l'IA est capable d'extraire les données d'une situation particulière, les traiter et apporter une compréhension fine et efficace au robot.



Exemples de robots dotés d'IA

- Nao et Pepper (conçus par la société française Aldebaran) sont des robots humanoïdes, respectivement de 58 et 120 cm de haut, capables de détecter les visages et de converser avec les hommes.
- L'aspirateur intelligent Roomba (société iRobot) navigue et cartographie son environnement.
- Bras robotisé : afin d'aider les robots à détecter des objets, apprendre puis répéter les mouvements humains.
- Robots de livraison ou de logistique : pour la navigation en extérieur (robots de Starship Technologies) ou en intérieur dans des commerces (Plato de la société Aldebaran) ou dans des entrepôts (robots d'Amazon Robotics).



Exemple de capteurs utilisés en robotique et en IA

L'IA peut identifier avec une bonne efficacité les objets mais aussi les hommes. À la détection d'un visage humain, le robot est capable d'engager une ou plusieurs actions. Cela peut être un simple mouvement de la main pour le saluer ou l'inviter à se présenter. Dans ce dernier cas, une IA conversationnelle prend le relais. Celle-ci recourra aux microphones (capteurs) du robot pour interpréter ce qui lui est dit.

La liste des cas d'usage où robot et IA agissent de concert est infinie : de l'industrie où des robots dotés d'IA sont utilisés dans le contrôle qualité de pièces détachées, à la logistique où des robots naviguent dans d'immenses entrepôts en passant par la médecine avec des robots chirurgiens.

La combinaison de ces champions technologiques devrait permettre de nombreuses évolutions en matière de véhicules autonomes, de drones, de jouets intelligents, ou de robots d'exploration spatiale...