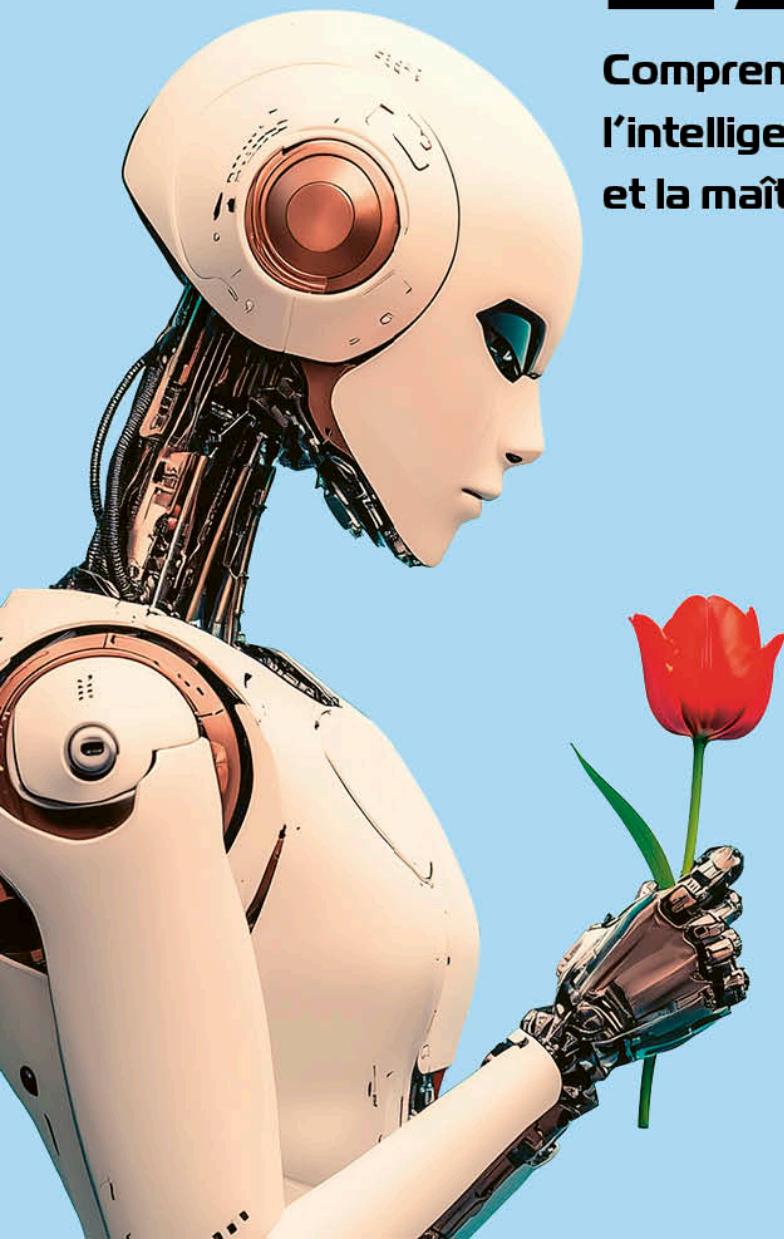


Manuel de survie à l'

IA

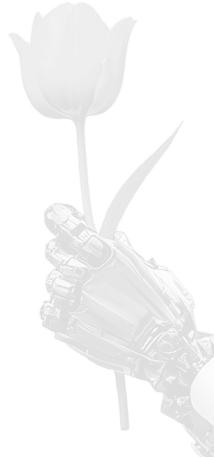
**Comprendre
l'intelligence artificielle
et la maîtriser au quotidien**



**Édouard Fillias
Alexandre Villeneuve**

ellipses

The logo consists of the word "ellipses" in a lowercase serif font, enclosed within three concentric ellipses of decreasing size.



Chapitre 1.1

Les origines et les différentes formes de l'IA

L'essence et l'origine de l'IA

Des légendaires statues parlantes de l'ancienne Égypte, aux robots de la littérature de science-fiction, en passant par les automates du dieu forgeron de l'Antiquité grecque, il semble que les êtres artificiels accompagnent notre imaginaire depuis la plus haute antiquité. Mais l'IA moderne n'a évidemment pas grand-chose en commun avec ces mythes, pas plus qu'elle ne repose sur des cerveaux « positroniques » ou toute autre imitation supposée du corps humain. Elle se fonde sur des algorithmes et du code informatique, fruits d'une histoire déjà longue, riche de ses percées scientifiques et de ses déceptions.

Tout commence en 1950 avec Alan Turing, un mathématicien britannique dont le nom est indissociable des origines de l'informatique. En 1950, quelques années avant sa mort tragique, Turing propose une idée radicale, le test qui finira par porter son nom. Le principe en est simple, mais révolutionnaire. Il s'agit de déterminer si une machine peut penser. Si une machine pouvait converser de manière indiscernable d'un humain, alors, selon Turing, elle pourrait être considérée comme « intelligente ». Cette proposition audacieuse a planté la première graine de l'IA, soulevant une question qui continue de stimuler les chercheurs : une machine peut-elle vraiment simuler l'intelligence humaine dans sa totalité ?

Mais c'est en 1956 que le terme « intelligence artificielle » est forgé, lors d'une conférence historique à Dartmouth College. Des pionniers comme John McCarthy, Marvin Minsky ou Herbert Simon y jettent les bases de ce nouveau champ de recherche. L'enthousiasme est à son comble : on pense alors pouvoir créer rapidement des machines aussi intelligentes que l'homme.

Six ans plus tard, en 1956, une conférence historique se tient à Dartmouth College, aux États-Unis. Laquelle est souvent considérée comme l'acte de naissance officiel de l'intelligence artificielle en tant que domaine de recherche. Les participants y formulent l'hypothèse qu'il est possible de créer des machines capables de simuler tout aspect de l'intelligence humaine. Ils y posent également les fondements théoriques et méthodologiques de l'IA, explorant des concepts qui resteront au cœur de la discipline pendant des décennies.

En 1959, apparaît le terme de *machine learning*, défini par son pionnier Arthur Samuel comme le « champ d'étude qui donne aux ordinateurs la capacité d'apprendre sans être explicitement programmés à apprendre ».

L'un des premiers résultats concrets de cette exploration est ELIZA, un programme développé par Joseph Weizenbaum en 1965. ELIZA est un chatbot primitif conçu pour simuler une conversation avec un thérapeute. Il repose sur des règles simples de traitement du langage naturel, mais parvient à donner l'illusion d'une véritable compréhension. Le programme, malgré sa simplicité, soulève des questions profondes sur l'anthropomorphisme et la facilité avec laquelle les humains peuvent être trompés par des machines. Weizenbaum lui-même est surpris par l'attachement émotionnel que certains utilisateurs développent envers ELIZA, un phénomène qui illustre la puissance de la suggestion et de l'interaction humaine avec les machines.

Les années 1970 voient l'apparition des systèmes experts, qui sont parmi les premières applications commerciales de l'IA. Ces systèmes sont capables de prendre des décisions dans des domaines spécifiques en se basant sur un ensemble de règles définies par des experts humains. Ils sont largement utilisés dans le domaine médical et industriel, offrant des diagnostics ou des recommandations basées sur des bases de données complexes. Par exemple, le système MYCIN, développé à l'université de Stanford, est conçu pour diagnostiquer des infections bactériennes et recommander des traitements. Toutefois, ces systèmes sont coûteux à développer et difficiles à

maintenir, ce qui limite leur adoption. Ils nécessitent surtout une connaissance technique approfondie, et la moindre erreur dans la programmation des règles peut conduire à des conclusions erronées. Un outil difficile à manier.

L'euphorie dans le domaine retombe ensuite brutalement à la fin des années 1980. Les promesses mirobolantes ne se concrétisent pas, entraînant une chute des financements. C'est le premier « hiver de l'IA », qui dure jusqu'au milieu des années 1990. Les systèmes experts montrent leurs limites, et les chercheurs peinent à surmonter les défis techniques et pratiques. Les attentes étaient élevées, mais la technologie de l'époque n'était pas encore prête pour les applications promises, trop ambitieuses pour leur époque. Les limitations en puissance de calcul, en capacités de stockage et en algorithmes efficaces, freinent considérablement les progrès. L'enthousiasme pour l'IA diminue, c'est l'heure de la désillusion.

Un tournant décisif survient en 1997, lorsque l'ordinateur Deep Blue d'IBM bat le champion du monde d'échecs Garry Kasparov. Cet exploit symbolise les capacités croissantes des machines à réaliser des tâches complexes. Il ne s'agit pas encore d'intelligence au sens humain du terme, mais c'est une démonstration spectaculaire de la puissance de calcul et de la capacité d'analyse des ordinateurs. *Deep Blue* utilise une force brute de calcul pour évaluer des millions de positions possibles par seconde, combinée à des algorithmes sophistiqués de sélection de mouvements. Cette victoire emblématique suscite un regain d'intérêt pour l'IA, montrant que les machines peuvent exceller dans des domaines autrefois réservés aux humains.

Les années 2000 voient l'émergence du *deep learning* (apprentissage profond), une méthode d'apprentissage automatique inspirée du fonctionnement du cerveau humain. Cette approche repose sur des réseaux de neurones artificiels capables d'apprendre à partir de vastes quantités de données. Les progrès du *deep learning* ouvrent de nouvelles perspectives pour l'IA, permettant des avancées significatives dans des domaines tels que la reconnaissance d'image, la traduction automatique et bien d'autres. Des entreprises comme Google, Facebook et Microsoft investissent massivement dans cette technologie, créant des modèles de plus en plus sophistiqués et performants.

La décennie suivante, les années 2010, est souvent qualifiée d'âge d'or du *deep learning*. Des assistants vocaux comme Siri d'Apple deviennent courants, et des programmes comme AlphaGo de Google DeepMind repoussent les limites de ce que l'IA peut accomplir. En 2016, AlphaGo bat le champion du monde de go, un jeu beaucoup plus complexe que les échecs, démontrant la supériorité des algorithmes d'apprentissage profond dans certains domaines. Cette victoire est d'autant plus impressionnante que le go est un jeu basé davantage sur l'intuition et la stratégie que sur la force brute de calcul. AlphaGo utilise des réseaux de neurones profonds pour apprendre à jouer en analysant des millions de parties et en jouant contre lui-même.

L'IA générative, qui émerge dans les années 2020, marque une nouvelle étape. Les modèles dits transformeurs, comme BERT de Google, révolutionnent le traitement du langage naturel. Ces modèles permettent de générer du texte, des images et même des vidéos de manière incroyablement réaliste. En 2023, l'explosion de ChatGPT, développé par OpenAI illustre le nouveau potentiel de ces technologies. ChatGPT est capable de tenir des conversations presque humaines, d'écrire des textes cohérents et de répondre à une multitude de questions, ouvrant des perspectives infinies pour l'IA. Les applications à court terme ne manquent pas, allant de l'assistance à l'écriture créative à l'automatisation du service client, en passant par l'aide à la prise de décision.

Après un certain nombre de triomphes, et bien des revers, il semble que les chercheurs soient finalement parvenus à surmonter les défis hier insurmontables. L'IA, bien que toujours en évolution, est aujourd'hui une réalité tangible, intégrée dans de nombreux aspects de notre vie quotidienne. Les héros de cette épopée, de Turing à McCarthy, de Claude Shannon à Weizenbaum, ont tous contribué à façonner cette technologie qui commence à transformer notre monde.

En France, l'intérêt pour l'IA s'est également intensifié au fil des décennies. Des chercheurs comme Yann LeCun, pionnier du *deep learning*, ont contribué à placer la France sur la carte mondiale de l'IA. En 2018, le gouvernement français a lancé une stratégie nationale pour l'intelligence artificielle, visant à positionner le pays comme un leader dans ce domaine. Cette initiative comprend des investissements dans la recherche, la formation et le soutien aux startups innovantes. Des centres de recherche et des

laboratoires, comme l'INRIA (Institut national de recherche en informatique et en automatique) et le CNRS (centre national de la recherche), jouent un rôle majeur dans le développement de technologies IA avancées.

Reproduire l'intelligence humaine, créer des machines capables de résoudre des problèmes complexes et d'interagir de manière naturelle avec les humains est un Graal propre à motiver les chercheurs et les développeurs du monde entier. C'est aussi l'objet d'une compétition acharnée entre les nations et les géants des technologies.

Sommes-nous aujourd'hui à l'aube d'une nouvelle révolution industrielle, comparable à l'invention de l'électricité ? Ou les limites de ces systèmes vont-elles à nouveau freiner les espoirs ? Une chose est sûre : l'histoire de l'IA nous enseigne la prudence face aux promesses mirobolantes. Car derrière les prouesses les plus spectaculaires se cachent souvent des réalités plus prosaïques qu'on ne l'imagine.

Les grandes technologies derrière l'IA

L'intelligence artificielle s'est développée pour englober une variété de technologies, chacune ayant des caractéristiques propres et des applications spécifiques. Pour saisir les capacités et les limites de l'IA à notre époque, il est utile de comprendre ces différentes catégories et leurs contextes. L'IA n'est pas monolithique, et derrière cet acronyme se cachent plusieurs ensembles bien distincts.

L'une des premières formes d'IA développées est le système expert. Ces systèmes, apparus dans les années 1970, se basent sur des règles préétablies pour simuler le comportement d'un expert humain dans un domaine spécifique. Un exemple emblématique de système expert est MYCIN, développé pour aider au diagnostic de maladies infectieuses. MYCIN posait des questions précises et, en fonction des réponses, proposait des diagnostics et des traitements. Ces systèmes fonctionnent grâce à une vaste base de connaissances structurée par des experts humains. Imaginez un arbre de décision simple, si « symptôme a » et « symptôme b » alors « contrôler si ce n'est pas maladie c ». Cependant, leur rigidité constitue une limitation majeure : on leur apprend laborieusement à raisonner dans une situation

donnée, et ils ne peuvent pas s'adapter sans intervention humaine pour mettre à jour leurs règles. Cette approche reste utilisée pour certains usages spécifiques où les règles sont bien définies, en finance notamment.

Le *machine learning*, ou apprentissage automatique, marque une avancée significative par rapport aux systèmes experts. Contrairement à ces derniers, le *machine learning* permet aux machines d'apprendre à partir de données sans être explicitement programmées pour chaque tâche. Par exemple, les filtres anti-spam utilisent le *machine learning* pour analyser des milliers d'exemples de courriels afin d'identifier des motifs communs aux spams d'un côté et aux courriels légitimes de l'autre, ajustant leurs algorithmes en conséquence. Netflix, qui recommande des films et des séries à ses utilisateurs en fonction de leurs historiques de visionnage et des préférences de spectateurs similaires, est un autre exemple de l'application du *machine learning*. Le *machine learning* est également utilisé dans la détection de fraudes, où il analyse les transactions pour identifier des comportements suspects.

Le *deep learning* est une sous-catégorie du *machine learning* qui utilise une structure simulant un réseau de neurones du cerveau humain. Ces réseaux de neurones artificiels sont capables de découvrir des structures complexes dans de vastes ensembles de données sans intervention humaine. Cet apprentissage autonome limite le contrôle humain, c'est pourquoi on parle souvent de « boîte noire ».

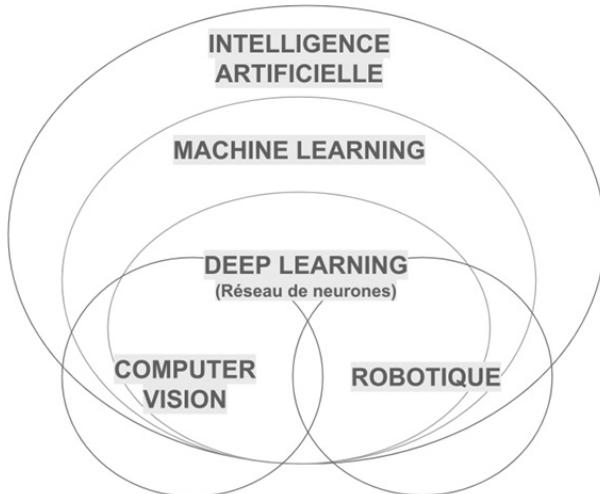
C'est un fonctionnement principalement statistique. On va fournir à l'IA une base d'entraînement (photos de chats, par exemple), et elle va comprendre d'elle-même les similarités statistiques entre les chats (yeux, oreilles pointues...). Une fois entraînée, elle sera capable de donner une note de similarité si on lui fournit une nouvelle photo de chat, ou même de générer une image de chat original, qui ne sera qu'une représentation moyenne des chats des données reçues initialement. Dès lors, on comprend que plus la base d'entraînement est large et de qualité, plus la réponse de l'IA sera précise. Si, dans la base d'entraînement de chats, il y avait une proportion trop importante de chiens, l'IA risquerait de confondre les chats et les chiens.

C'est la base de tous les assistants comme Siri d'Apple ou Alexa d'Amazon, qui comprennent et répondent à des requêtes vocales en utilisant des modèles de *deep learning* entraînés sur des milliers d'heures de paroles

humaines. C'est aussi la base de toutes ces nouvelles IA qui font le buzz comme ChatGPT. Le modèle de langage de ce dernier a été entraîné sur des données variées, notamment Wikipedia dans toutes les langues, des articles de presse, des livres et romans, mais aussi des données dans des registres de langage moins soutenus, comme les forums ou les réseaux sociaux. Ainsi, il peut répondre dans toutes les langues, à toutes les questions, et avec tous les niveaux de langage. Si ses réponses sont souvent très convaincantes, l'IA ne comprend pas réellement vos commandes ou questions ; elle génère uniquement ce qui est statistiquement le plus probable. C'est pourquoi plus votre question sera générale, plus la réponse de l'IA sera pertinente. En revanche, pour une question plus pointue (lorsqu'elle n'a pas assez de données par exemple), l'IA risque d'inventer une réponse irréaliste, un phénomène appelé hallucination.

La computer vision, ou vision par ordinateur, permet aux machines de comprendre et d'interpréter le contenu visuel du monde qui les entoure. Un exemple concret est celui des voitures autonomes, qui utilisent des caméras et des algorithmes de computer vision pour détecter et analyser les éléments de la route, tels que les panneaux de signalisation, les piétons et les autres véhicules. Un autre domaine d'application est la reconnaissance faciale, utilisée pour des fonctionnalités telles que le déverrouillage de smartphones ou la surveillance de sécurité.

La robotique combine plusieurs aspects de l'IA pour créer des machines capables d'interagir physiquement avec leur environnement. Les robots industriels, par exemple, sont utilisés depuis des décennies sur les chaînes d'assemblage de l'industrie automobile. Équipés de capteurs et de systèmes d'IA, ces robots peuvent manipuler des objets avec précision. Plus récemment, des robots plus sophistiqués, comme ceux développés par Boston Dynamics, peuvent se déplacer de manière autonome, s'adapter à des terrains variés et accomplir des tâches complexes. En France, des startups comme Exotec innovent dans le domaine de la robotique logistique, créant des robots capables de déplacer et de trier des produits dans des entrepôts de manière efficace et autonome. Cette technologie est également utilisée dans la médecine, avec des robots chirurgicaux capables d'effectuer des opérations délicates avec une précision inégalée.



Les nouvelles IA à tout faire

Depuis quelques années, l'intelligence artificielle a fait un bond en avant grâce au *deep learning*, l'une des percées les plus marquantes de cette avancée technologique est l'émergence de l'IA générative (avec sa figure de proue ChatGPT), capable de créer des contenus de qualité croissante, qu'il s'agisse de textes, d'images, de sons ou même de vidéos.

L'IA générative repose principalement sur 2 grands types de modèles : les modèles de langages et les modèles d'images et de sons. Chaque modèle a ses spécificités et ses applications, mais tous partagent le même objectif : générer du contenu de qualité et de manière autonome.

Les modèles de langage sont des algorithmes appelés LLM (*Large Language Models*). Ces modèles sont entraînés sur dénormes quantités de textes, ce qui leur permet de comprendre et de produire du texte de manière étonnamment cohérente. Par exemple, GPT (modèle de ChatGPT), l'un des modèles les plus avancés, a été entraîné sur des centaines de milliards de mots provenant de livres, d'articles, de sites web, etc. Grâce à cette immense base de données, il peut rédiger des articles, répondre à des questions, générer du code informatique, et même créer des dialogues fictifs avec une fluidité capable de tromper un lecteur humain.