

L'intelligence artificielle, définitions et historique

Dossier de veille – Octobre 2024

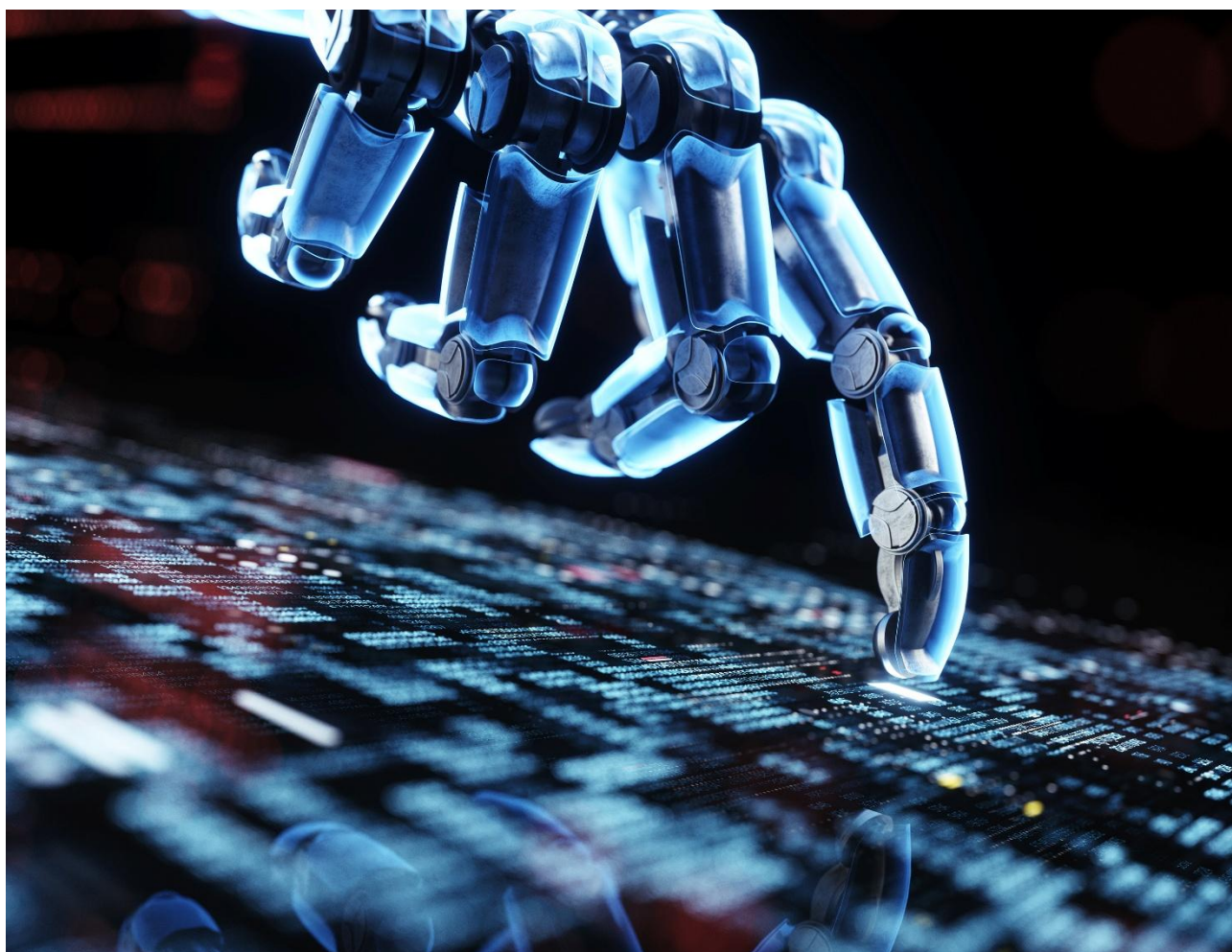


Table des matières

Introduction	3
L'IA, de quoi parle-t-on ?	4
Définition de l'intelligence artificielle	4
IA étroite, IA générale et IA super-intelligente, quelle différence ?	4
La data, clé de voute du développement des applications d'IA	5
Principaux concepts et approches en intelligence artificielle	6
Focus sur l'IA générative	7
De 1943 à nos jours, focus sur les solutions innovantes au fil des années	10
1940 – 1960 : les premiers pas de l'IA	10
1970 – 1990 : les systèmes experts	10
Depuis 2010 : le 3ème âge de l'IA	10
Une accélération depuis 2023 et grandes dates à venir	12

Introduction

Si ses prémices remontent à la création du premier réseau de neurones par Warren McCulloch et Walter Pitts en 1943, l'Intelligence Artificielle (IA) connaît un véritable tournant depuis l'avènement des grands modèles de langage (LLM) ces dernières années. Dès son lancement en novembre 2022, le chatbot ChatGPT marque l'avènement de l'IA en tant qu'outil de service, et devient le produit technologique le plus vite adopté de l'histoire avec plus d'un million d'utilisateurs en une semaine.

Une telle dynamique technologique apporte son lot de nouveaux enjeux, qu'il convient d'analyser et de comprendre pour pouvoir déployer des stratégies pertinentes au sein de nos organisations, prenant en compte les opportunités, les risques, les impacts.

Dans cette perspective, ce dossier vise à poser quelques définitions de l'intelligence artificielle et de son fonctionnement. Seront abordés également les principaux concepts et principales approches en IA. Une seconde partie de dossier retracera l'évolution des solutions d'IA depuis ses origines jusqu'à nos jours. Les solutions d'intelligence artificielle évoluant de façon extrêmement rapide, ces éléments sont à considérer comme valables à date de publication de ce dossier, à l'automne 2024.

Ce dossier portant sur les bases pour comprendre l'IA et ses notions, est complété par deux autres dossiers qui permettront d'approfondir les thématiques suivantes :

- Intelligence artificielle, cas d'usage et panorama des solutions,
- Intelligence artificielle, IA act et souveraineté numérique.

L'ensemble de ces dossiers est à retrouver [sur le site du Lab de la Sécurité sociale](#).

L'IA, de quoi parle-t-on ?

Définition de l'intelligence artificielle

L'intelligence artificielle est définie par le Parlement européen comme « la possibilité pour une machine de **reproduire des comportements humains**, tels que la perception, l'analyse, la planification et la créativité ». Elle désigne les systèmes en capacité de réaliser des tâches plus ou moins complexes via l'exécution d'algorithmes.

Ainsi, les principes clés de l'IA sont

1. La capacité d'apprentissage des machines à partir de données
2. La capacité de prendre des décisions autonomes.

Aujourd'hui, l'intelligence artificielle représente un secteur dont l'application et le poids économique ne cessent de croître. Alors qu'en 2015 le marché de l'intelligence artificielle pesait 200 millions de dollars, il devrait s'élever à près de 265 milliards de dollars en 2027. On estime par ailleurs que la productivité engendrée par cette automatisation des tâches équivaldrait au rendement de 110 à 140 millions de travailleurs à temps plein.

Qu'est-ce qu'un algorithme ?

D'après la définition donnée par le CNIL, un algorithme est la description d'une suite d'étapes permettant d'obtenir un résultat à partir d'éléments fournis en entrée.

IA étroite, IA générale et IA super-intelligente, quelle différence ?

Aujourd'hui, on distingue trois niveaux d'IA : l'IA « étroite » ou faible (ANI), l'IA « générale » ou forte (AGI) et la Super IA (ASI). Ces IA se différencient les unes des autres en fonction de leur capacité à faire et à apprendre.

- **L'IA « étroite »** : elle reproduit des éléments précis et résout des problèmes définis. Elle représente aujourd'hui **l'ensemble** des applications d'IA comme les voitures autonomes, la reconnaissance d'image, etc.
- **L'IA « générale »** : Ce niveau d'IA est théoriquement capable d'apprendre et de s'adapter à tout problème tel que le ferait un humain. Ce niveau cherche à reproduire la complexité du cerveau humain, mais reste aujourd'hui **dans le champ de la recherche**. Il n'existe pas d'application aujourd'hui.
- **L'IA « super-intelligente »** : elle désigne la forme de l'IA qui dépasse les capacités cognitives de l'humain, de manière totalement autonome. Cette forme d'IA est hypothétique et pose des questions éthiques et philosophiques importantes.

Quelles que soient les formes d'IA, ces dernières sont étroitement liées aux données (la data est le carburant des applications d'IA).

La data, clé de voute du développement des applications d'IA

Rappel : La **data** (ou données) représente des informations brutes ou des faits collectés, stockés et analysés dans le cadre d'un processus informatique.

Le **Big Data** (ou mégadonnées, grosses données, données massives) désigne quant à lui un ensemble très volumineux de données qu'aucun outil classique de gestion de base de données ou de gestion de l'information ne peut vraiment travailler.

La data est l'élément clé garantissant le bon fonctionnement, l'efficacité et la pertinence de solution d'IA : contrairement à un produit informatique dont le code est intrinsèquement statique, les données sont la matière première d'une solution d'IA et l'alimentent en continu ("**Pas d'IA sans Data**"). En effet, à partir des volumes de données, l'IA extrait les informations ou caractéristiques dont elle a besoin avant de les traiter pour aboutir au résultat escompté. Les données apparaissent ainsi comme étant une des clés du développement d'application d'IA.

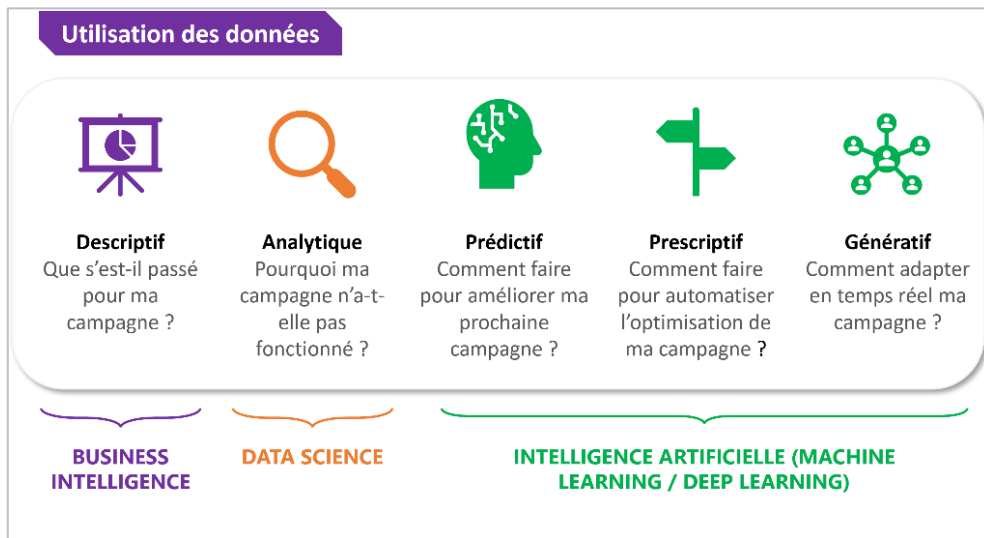
Ce rôle conféré aux données n'est pas neutre. En effet, il découle du rôle prépondérant de la data trois conséquences clés qu'il convient de prendre en compte dès les phases de définition d'application :

- Une part de l'efficacité de l'IA dépend de la **quantité** et de la **qualité des données** ;
- La pertinence d'une solution d'IA **dans le temps** dépend des données qui l'alimentent (et de leur dynamique d'évolution) et des données permettant le suivi des performances et la création de valeurs, d'où la nécessité d'anticiper **la gestion de la dynamique des données** dans sa phase de conception ;
- L'émergence de l'IA a souligné **l'importance cruciale des données** pour les entreprises, entraînant une prolifération de capteurs, de structures organisationnelles dédiées à la gestion des données, et l'émergence de nouveaux modèles économiques ;

Ainsi, la question de la disponibilité et de la qualité des données est une question à adresser dès les prémices du développement d'une solution d'IA.

L'IA se caractérise par une exploitation des données dite « de data science »

La data peut être exploitée de différentes manières en fonction des fins recherchées. On parle de **business intelligence** pour une utilisation dite descriptive, de **data science** pour une utilisation analytique, et d'**intelligence artificielle** (*machine learning / deep learning*) pour une utilisation prédictive, prescriptive ou générative.

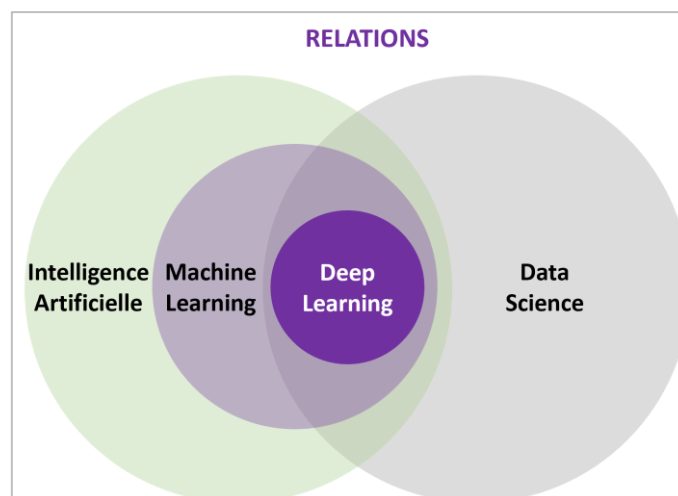


Principaux concepts et approches en intelligence artificielle

Domaines de l'IA et types d'apprentissage

L'IA se décompose en 3 domaines, chacun lié à une méthode d'apprentissage spécifique. On y trouve :

- La Data Science qui utilise des méthodes statistiques, mathématiques et informatiques pour extraire des connaissances et des informations à partir de données.
- Le **Machine learning** (apprentissage automatique) qui est une technologie d'intelligence artificielle permettant aux ordinateurs **d'apprendre sans avoir été explicitement programmés à cet effet**.
- Le **Deep learning** (apprentissage profond) est une sous-branche de l'apprentissage automatique qui implique l'utilisation de **réseaux de neurones artificiels** pour apprendre des représentations de données à partir de couches de traitement successives.



Ainsi, la data science est un domaine plus large qui englobe la manipulation, l'analyse et l'interprétation des données, tandis que le *machine learning* et le *deep learning* sont des sous-domaines qui se

concentrent spécifiquement sur l'utilisation de techniques et d'algorithmes pour apprendre à partir de données et prendre des décisions. Le *deep learning*, en particulier, se distingue par son utilisation de réseaux de neurones profonds pour apprendre des modèles complexes à partir de données brutes.

On distingue deux grands types d'algorithmes d'apprentissage : les algorithmes supervisés et non-supervisés.

Dans le cas des algorithmes supervisés, les données fournies à l'algorithme pour entretenir son apprentissage sont étiquetées par l'humain en amont, pour que la machine puisse reconnaître une donnée similaire sans étiquette par la suite.

Dans le cas des algorithmes non-supervisés, ou encore « auto-apprenants », la machine doit analyser seule l'ensemble des données pour observer de potentielles structures ou tendances. Certains algorithmes sont conçus pour que leur comportement évolue dans le temps, en fonction des données qui leur ont été fournies

Qu'est-ce qu'un biais algorithmique ?

Le **bias algorithmique**, tel un biais cognitif pour les algorithmes, se réfère à la tendance des algorithmes à **favoriser certains contenus ou personnes**, ce qui entraîne des **résultats faussés ou inexacts**. Ce phénomène peut être causé par des préjugés inconscients des développeurs, des données d'entraînement biaisées ou des biais systémiques, rendant ainsi l'algorithme partial, non-neutre voire préjudiciable.

On peut compter plusieurs types de biais : biais des data scientists et des développeurs, biais statistiques, biais économiques...

Focus sur l'IA générative

Les LLM (Large Language Model)

Un "**Large Language Model**" (LLM) est un modèle de langage conçu pour **comprendre et générer du texte de manière intelligente**. Ainsi, les Large Language Models (LLM) sont des algorithmes de *Deep learning* qui peuvent exécuter un large panel de tâches de **traitement du langage naturel (NLP)**. De ce fait, ils ont la capacité de reconnaître, traduire, prédire ou encore générer des textes ou autres contenus.

Par extension, les LLM peuvent également être utilisés pour d'autres « types » de langage, tels que des structures de protéines, des codes de logiciels ou d'autres modèles de langage ayant une structure spécifique, et ce grâce à l'utilisation de modèles de transformateur.

Un **modèle de transformateur** désigne l'architecture la plus courante d'un grand modèle de langage, et se compose d'un **encodeur et d'un décodeur** : un modèle de transformateur traite les données en convertissant les entrées en « **tokens** », puis en réalisant des équations mathématiques simultanées pour découvrir les relations qui existent entre les « tokens ».

Les décodeurs utilisent le contexte de la phrase pour faire des prédictions et deviner le prochain « token ». Chaque token a une probabilité d'être prédit.

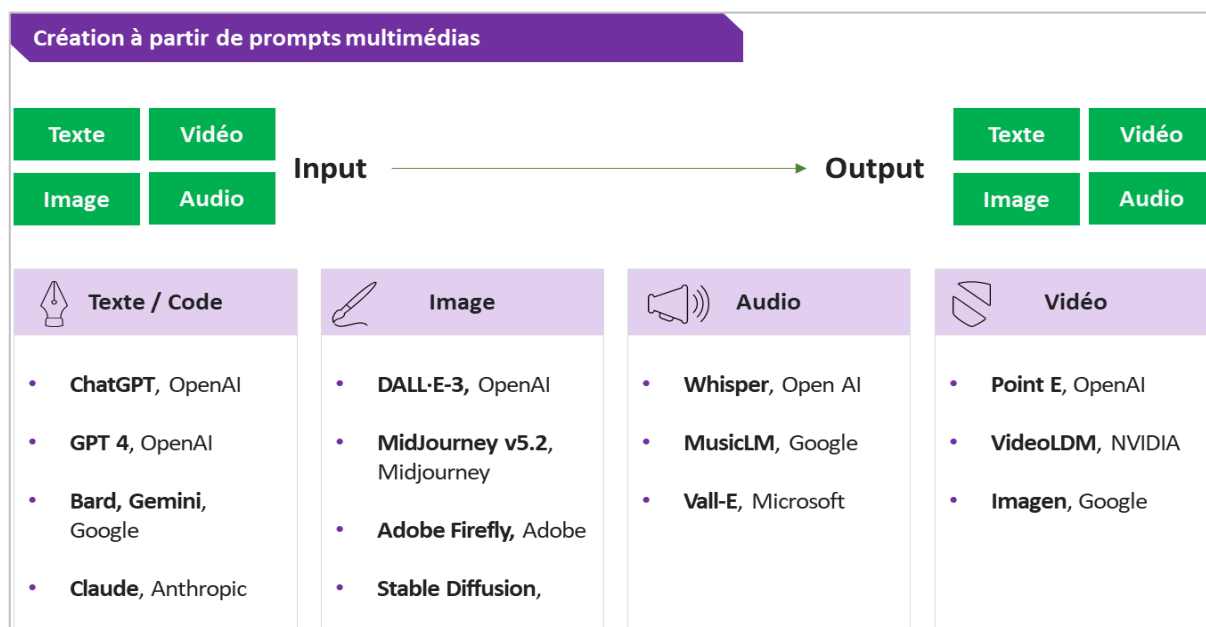
Exemple		
<p>Le « paramètre de température » introduit quant à lui un caractère aléatoire et détermine à quel point le modèle sera "créatif" avec ses réponses.</p> 	chaise	0,2
	table	0,15
	fenêtre	0,04
	lune	0,001
	...	

Le « **paramètre de température** » introduit quant à lui un caractère aléatoire et détermine à quel point le modèle sera "**créatif**" avec ses réponses.

Grace à ces différents paramètres, les LLM génèrent du texte et sont utilisés dans une large variété d'application dont la rédaction automatique de contenu, l'assistance à la rédaction, la traduction automatique, l'analyse de sentiment, la génération de dialogue, etc.

Modèles d'IA génératives, usages et limites

L'IA générative est un type de système d'intelligence artificielle capable de générer du texte, des images ou d'autres images ou d'autres médias en réponse à des prompts, d'apprendre les modèles et la structure des données d'entrée et de générer un nouveau contenu similaire.



Le LLM est exemple d'IA générative.

Aujourd'hui plusieurs applications existent et permettent de générer différents contenus.

Ces applications à travers leurs caractéristiques ont des usages et des limites qui leur sont propres :

	CHAT GPT, BARD, BING	COPILOT, DUET	LLMQ PROPRIETAIRES	LLM OPEN SOURCE
Caractéristique	Modèle accessible sous forme d'interfaces conversationnelles	Application intégrée à une suite bureautique	Modèles intégrés dans les solutions cloud propriétaire	Multitude de petits ou grands modèles généralistes ou spécialisés
Usages	Généraliste	Bureautique	Puissance d'usage	Spécialisé et souveraineté
Modèles économiques	Abonnement / Token	Licence /consommation cloud	Consommation cloud	Hébergement et compute
Limites	Modèle généraliste et réglementation	Limité au périmètre bureautique	Souveraineté	Compétences et infra dédiées

Les différentes caractéristiques des solutions et plus particulièrement leurs usages et leurs limites doivent guider le choix de la solution à utiliser / développer en réponse à un besoin spécifique.

De 1943 à nos jours, focus sur les solutions innovantes au fil des années

1940 – 1960 : les premiers pas de l'IA

Portée par la Seconde Guerre Mondiale, de nombreux scientifiques se sont intéressés à la discipline de la **cybernétique**, qui désigne la **science des communications** et de la **régulation** dans **l'être vivant et la machine**. L'objectif était simple : réussir à **dupliquer la raison humaine à l'intérieur d'une machine** et d'élaborer un modèle pour penser le cerveau.

Les premiers articles, recherches et invention à ce sujet ont pu créer les fondations de notre conception de l'IA telle qu'on la connaît aujourd'hui :

- **1943** : Warren McCulloch et Walter Pitts publie l'article « A Logical Calculus of Ideas Immanent in Nervous Activity », qui signe le premier modèle mathématique et informatique pour la création d'un réseau de neurones
- **1950** : Marvin Minsky et Dean Edmonds créent le premier ordinateur à réseau de neurones
- **1950** : John Von Neumann et Alan Turing initient la transition entre les calculateurs à la logique décimale du XIXe siècle et des machines à logique binaire
- **1950** : Alan Turing publie le Turing Test, avec pour fonction d'évaluer les IA
- **1956** : naissance du terme intelligence artificielle lors d'une conférence de John McCarthy (naissance du concept d'IA tel qu'il est connu aujourd'hui)
- **1959** : invention du terme *Machine learning* par Arthur Samuel chez IBM

1970 – 1990 : les systèmes experts

Au début des années 60, l'engouement pour l'IA s'essouffle compte-tenu des difficultés techniques pour la développer. Mais le manque de mémoire des machines est résolu à la fin des années 1970 avec l'apparition des **premiers microprocesseurs** : l'IA connaît ainsi un second souffle avec l'avènement des **systèmes experts**.

Plusieurs événements marquants dans le développement de l'IA sont à noter sur la période :

- **1972** : Création du système MYCIN (pour le diagnostic des maladies du sang et la prescription médicamenteuse) avec « moteur d'inférence », programmé pour être un miroir logique d'un raisonnement humain
- **1989** : Mise au point par Yann Lecun du premier réseau de neurones capable de reconnaître des chiffres écrits à la main, origine du *Deep learning*
- **1997** : Le système Deep Blue d'IBM bat le champion d'échec Gary Kasparov

Depuis 2010 : le 3ème âge de l'IA

A la fin des années 1990, les systèmes de programmation se trouvaient vite bloqués lors de la multiplication de règles et les machines ne parvenaient plus à tenir un raisonnement compréhensible. Par ailleurs la maintenance devenait trop coûteuse.

Deux éléments ont permis de relancer la discipline dans les années 2010 : **l'accès à des volumes massifs de données**, et le développement des **processeurs de cartes graphiques** pour accélérer le calcul des algorithmes d'apprentissage. Enfin, **l'approche inductive** ne consiste plus à coder les règles

comme pour les systèmes experts, mais bien de laisser les ordinateurs apprendre par eux-mêmes par le biais de corrélation et de classification.

Ces découvertes, permettant d'immenses gains de temps et de coûts, ont relancé les financements des laboratoires de recherches et signé la création de nouvelles technologies sans précédents :

- **2010** : création de l'entreprise DeepMind (rachetée en 2014 par Google), à l'origine de nombreux programmes d'apprentissage par renforcement profond et dont l'IA AlphaGo de DeepMind bat le meilleur joueur de go au monde en 2016
- **2012** : un réseau de neurones constitué de 16 000 Microprocesseur est capable après entraînement de reconnaître un chat sur des vidéos
- **2015** : Création d'OpenAI
- **2023** : développement de grands modèles multimodaux (capables de traiter plusieurs modalités comme le texte, les images, le son) tels que Google Gemini210 ou GPT-4211.

Une accélération depuis 2023 et grandes dates à venir

