AltGR

GUIDE DE SURVIE EN MILIEU HOSTILE

intelligence artificielle
+ deep learning
+ data science
+ chatGPT
+ etc

pensé pour les CDO, CTO, CAIO, CDAO

à quoi sert ce guide ?

avoir les idées claires

L'ensemble des acteurs privés et publics rêve d'adopter le modèle opérationnel des géants du numérique : déléguer la prise de décision à des algorithmes qui exploitent systématiquement d'immenses bases de données. La décennie 2010-2020 voit le monde économique succomber à cette mode et s'exprimer uniquement par mots-clés : *tech*, *Big Data*, *data mining*, *data science*, *deep learning* et maintenant les *large languages models* et l'inévitable **ChatGPT**. Pourtant, ces technologies sont largement incomprises. Elles doivent prouver leur pertinence dans un contexte de production, sans toutefois accaparer les budgets consacrés à l'innovation.

Ecrit par des techniciens de la donnée qui évoluent dans le monde de la R&D avec près de 100 projets réalisés depuis 2015, ce guide donne les éléments nécessaires pour contextualiser ces technologies et comprendre leurs usages.

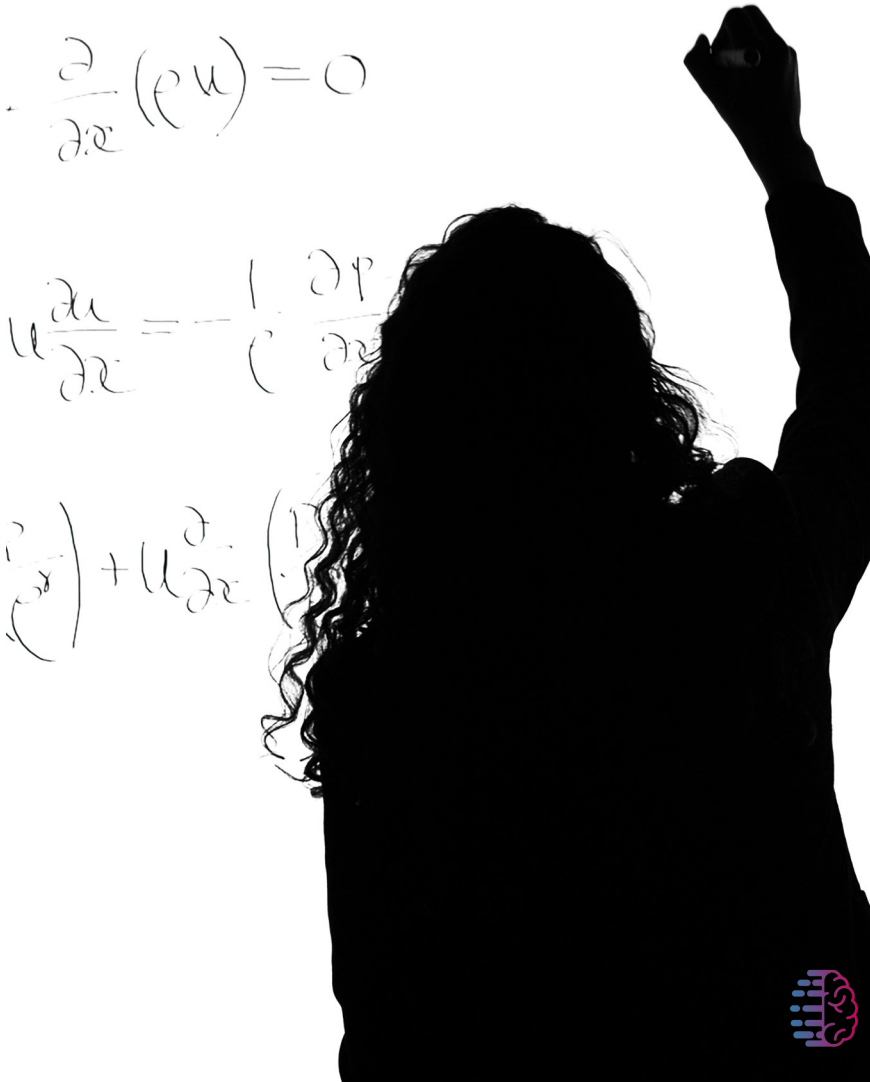
suivre les innovations

De nouveaux rôles stratégiques émergent : les *chief AI officer* (CAIO) et *chief data and AI officer* (CDAO) renforcent aujourd'hui les *chief data officer* (CDO). Ces personnes clés ont la lourde tâche de guider les organisations dans l'utilisation de ces nouvelles technologies liées à l'IA. Elles nécessitent une compréhension fine de domaines aussi variés que les mathématiques, l'informatique et les sciences cognitives.

Quels sont les impacts à attendre sur la stratégie et la gouvernance ? Quelles sont les technologies à suivre ? Les enjeux sont importants.

« Tout problème pour lequel aucune solution algorithmique n'est connue relève a priori de l'intelligence artificielle. »

Jean-Louis Laurière (1985)



notre but ?

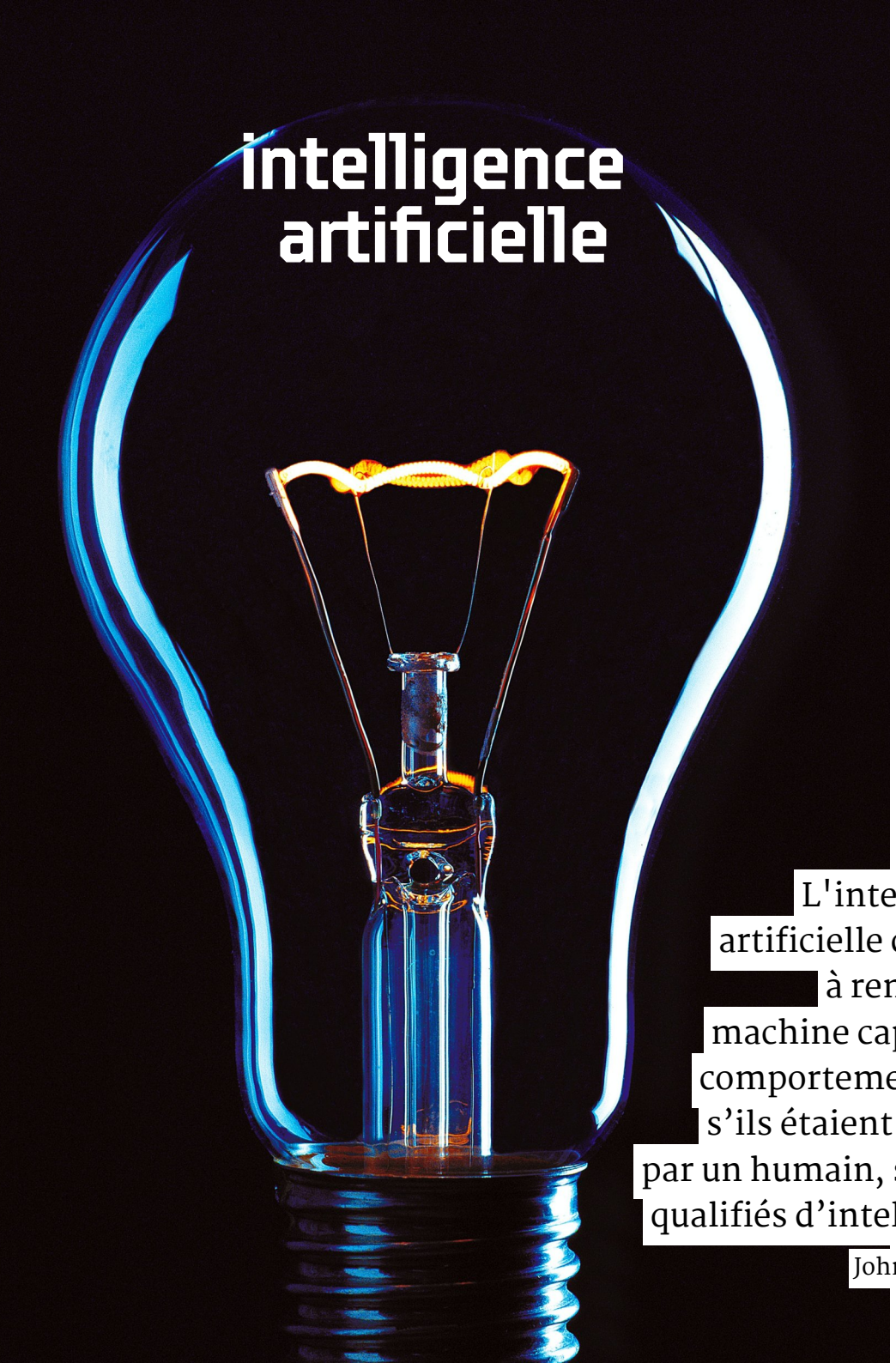
comprendre les technologiées liées à l'intelligence artificielle

L'histoire moderne de l'intelligence artificielle débute en 1956 (**p.6**). Depuis bientôt 70 ans, scientifiques et ingénieurs conçoivent des programmes “intelligents” (**p.7**), dont certains font appel aux algorithmes d'apprentissage automatique (**p.9**). Les réseaux de neurones artificiels (**p.12**) sous la forme du *deep learning* (**p.13**), vivent un printemps remarquable depuis 2016. De nombreux systèmes d'IA spécialisées sont créés en application de cette technologie, dont l'agent conversationnel **ChatGPT** (**p.15**).

> « La science des données [p.17] est une discipline émergente qui a évolué d’une approche essentiellement statistique,dans les années 2010, à un domaine mélangeant différentes branches des mathématiques et de l’informatique, dont l’intelligence artificielle. »
>
> Stéphane Mallat, Collège de France

AItGR

intelligence artificielle



> L'intelligence artificielle consiste à rendre une machine capable de comportements qui, s'ils étaient réalisés par un humain, seraient qualifiés d'intelligents.
>
> John McCarthy

L'intelligence artificielle fonctionne par cycles.
Le public se passionne au printemps,
et se refroidit vite quand l'hiver arrive.
Nous vivons aujourd'hui
le printemps du *deep learning*.

un domaine complexe

L'IA est d'abord un mythe. Notre imaginaire est peuplé de machines pensantes issues de la mythologie, de la littérature ou du cinéma. Cet héritage culturel est une source de nombreux fantasmes.

L'IA est un domaine de recherche liés aux sciences cognitives. Les scientifiques essaient de capturer l'intelligence humaine ou animale pour la reproduire à l'aide de machines et de programmes. Le domaine est vaste : logique, résolution de problèmes, robotique, algorithmique, traitement de l'information. Les scientifiques progressent avec difficulté et patience depuis les années 1950, de manière irrégulière et par vagues successives. La modélisation d'une forme d'IA générale qui résoudrait tous les problèmes reste toutefois un mystère.

L'IA est enfin un ensemble de technologies. Les ingénieurs construisent des outils d'aide à la décision, d'exploration ou d'automatisation. Certaines réalisations liées à l'informatique sont bien connues du grand public : **DeepBlue**, **Google Search**, **AlphaGo**, **ChatGPT**, **Dall·e**, **MidJourney**, **IBM Watson**.

Depuis 2010, une famille d'algorithmes d'apprentissage porte l'essentiel des progrès : les réseaux de neurones profonds.

les pionniers de l'IA

Les gens qui faisaient ces prédictions n'étaient pas cinglés. Ils essayaient simplement de préparer le public à des choses qui étaient plausibles à l'époque.

Patrick H. Winston (1984)

« On tentera de trouver comment faire en sorte que les machines utilisent le langage, forment des abstractions et des concepts, résolvent des types de problèmes aujourd'hui réservés aux humains et s'améliorent elles-mêmes. »

John McCarthy (1955), prix Turing 1971

« D'ici peu, nous pourrions apprendre à faire travailler ces programmes sur l'amélioration de leurs propres capacités.

Une fois un certain seuil franchi, cela pourrait conduire à une spirale d'accélération et il pourrait être difficile de mettre au point des garde-fous fiables pour la freiner. »

Marvin Minksy (1968), prix Turing 1969

« Mon but n'est pas de vous surprendre ou de vous choquer.

Nous disposons aujourd'hui de machines qui pensent, apprennent et créent. De plus, leurs capacités sur ces sujets vont augmenter rapidement. Dans un horizon proche, l'éventail des problèmes qu'elles pourront traiter sera proche de ce que l'esprit humain peut faire. »

Herbert A. Simon (1958), prix Turing 1975

AItGR

l'âge d'or : 1956-1970

Les ordinateurs sont des machines nouvelles, dédiées au traitement automatique de l'information. L'IA se fixe pour but de rendre ces machines, programmes ou robots, intelligents.

La première approche est symbolique. Comment modéliser et reproduire le raisonnement logique et la pensée humaine ? Pour manier ces concepts très abstraits, l'intelligence est divisée en un ensemble de fonctions, pour lesquelles une solution logicielle est construite, selon un paradigme réductionniste (*divide & conquer*). On parle habituellement de GOFAI, pour "*good old fashioned artificial intelligence*".

Le domaine fait de beaux progrès dans la thématique de la résolution de problèmes. Les avancées théoriques et pratiques suivent le développement de l'ensemble des outils informatiques. De nombreux systèmes sont créés (**SHRDLU**, **ELIZA**, **GPS**, **SHAKEY**) ainsi que des langages de programmation afin d'interagir efficacement avec les ordinateurs. Certains sont encore utilisés aujourd'hui (**LISP**, **PROLOG**).

Cette période est caractérisée par un optimisme débridé et communicatif des pionniers. Il se transmet au grand public. D'importants financements suivent. L'IA générale semble accessible dans quelques dizaines d'années.

premier hiver : '70s

Les attentes irréalistes sont progressivement douchées par la réalité. L'approche symbolique se heurte à la difficile modélisation du raisonnement logique. La décomposition des problèmes en tâches simples multiplie les branchements lorsque la complexité algorithmique est défavorable. L'espace des possibles devient alors gigantesque : c'est l'explosion combinatoire, phénomène qui est au cœur de nombreux problèmes réels. Le calcul, même rapide, ne suffit pas.

Aux blocages théoriques, deux difficultés s'ajoutent : le manque de sens commun et le problème du cadre. Si l'expérience humaine du monde se construit sur des implicites, un programme ne connaît rien. Tout doit être décrit. La logique ne suffit pas.

second printemps : '80s

L'émergence des systèmes experts relance l'intérêt du public pour l'IA. Le but est de concevoir des programmes capables d'exploiter des bases de connaissances constituées sur des domaines spécifiques. Certains systèmes experts sont utilisés de manière routinière comme **DENDRAL** ou **R1/XCON**. Ils sont cependant complexes et coûteux.

second hiver : fin '80s

La communauté de l'IA est critiquée pour avoir, encore, trop promis. Deux échecs industriels marquent cet hiver.

[1] Les ordinateurs optimisés pour traiter le langage de programmation **LISP**, alors dominant de l'IA, ne trouvent pas leur marché.

[2] Le Japon se lance en 1982 dans un projet très ambitieux : créer la cinquième génération de systèmes informatiques (**FGCS**) en concevant des ordinateurs basés sur le calcul massivement parallèle et la programmation logique. Trop en avance sur son temps, ce projet est un échec marquant.

vers les agents

Fin 1990 et début 2000, le paradigme évolue. Le système organisé autour d'un modèle central est remplacé par une approche comportementale. Elle devient le nouveau paradigme, à travers le développement d'une architecture dite réactive, qui permet le fonctionnement de robots (**Homer**, **Roomba**) ou agents (**SIRI**) en définissant une série de comportements possibles et en les classant par ordre de priorité.

Courant 2010, une révolution vient frapper le petit monde de l'apprentissage automatique : le *deep learning.*

les programmes d'intelligence artificielle

Il n'existe pas de raccourci pour l'intelligence, pas d'équations de Maxwell de la pensée à découvrir.

Doug Lenat (1990)

« Il y a ce mythe stupide selon lequel l'IA aéchoué, mais l'IA est partout autour de vous à chaque seconde de la journée. Les gens ne le remarquent tout simplement pas. »

Rodney Brooks (2022)

"cerveaux intelligents"

Après la seconde guerre mondiale, une nouvelle discipline se crée : l'informatique. Les idées de Norbert Wiener, Claude Shannon, John von Neumann et Alan Turing, pour ne citer que les plus connus, se cristallisent. On s'emploie rapidement à construire des "cerveaux électroniques" et d'essayer de les rendre "intelligents". C'est une tâche ardue, car l'intelligence est mal définie.

L'IA est en tension permanente entre scientifiques et ingénieurs, entre la compréhension théorique des systèmes et leur utilité pratique. Le but commun est de construire des machines capables d'accomplir des tâches qui, à nos yeux et selon un consensus implicite, nécessitent de l'intelligence. Accomplir cette tâche implique de travailler sur des questions fondamentales de l'informatique, comme l'architecture des ordinateurs, la programmation haut niveau ou la difficile articulation entre une capacité de raisonnement symbolique, proche du raisonnement humain, et la nécessaire intégration de connaissances relatives à des environnements parfois complexes.

Presque 70 ans après sa création, ce domaine est toujours très actif et passionnant.

programmer l'intelligence

[approches symboliques]

Ces systèmes regroupent les programmes basés sur la modélisation du raisonnement logique, ou plus généralement des architectures cognitives. Ces approches relèvent majoritairement de la recherche scientifique

ex : GOFAI, Pandemonium, Society of mind, General Problem Solver, SOAR, ACT-R

[systèmes experts]

Ces systèmes répondent à des problèmes précis en exploitant des bases de connaissances assemblées par des experts humains. L'articulation logique est confiée à des règles métiers.

ex : DENDRAL. X1/CON, DeepBlue, Stockfish

[approches statistiques]

Ces systèmes exploitent de vastes bases de données de manière statistique pour en extraire une connaissance spécifique à un domaine. La connaissance apportée par l'humain est minimale. L'apprentissage automatique entre dans cette catégorie.

ex : ChatGPT, MidJourney, AlphaGo.

[calcul et optimisation]

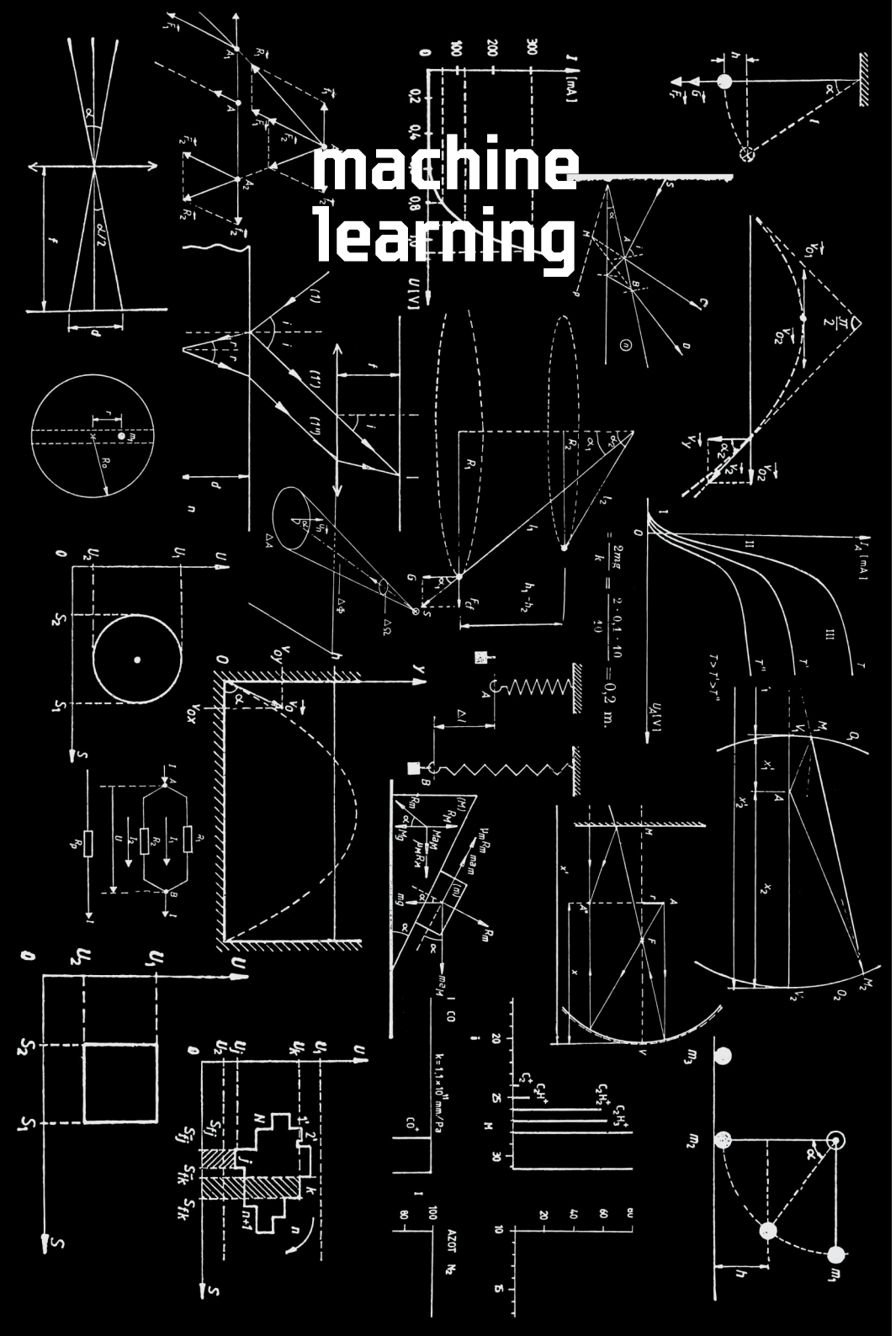
Ces systèmes ne relèvent plus de l'IA. Ces programmes se basent sur des algorithmes capables de “résoudre” des problèmes mathématiques. Il peut s'agir d’optimisation, de calcul sous contraintes, ou de problèmes avec une combinatoire défavorable. Les solutions recherchées sont optimales ou approchées. Ces outils forment des bibliothèques logicielles de calcul scientifique ou de recherche opérationnelle.

ex : BLAS, LAPACK, COIN-OR, solveurs SAT, NETLIB, numpy/scipy

Comprendre les principes de l'intelligence et les reproduire est le but des sciences cognitives. L'IA forte, ou intelligence artificielle générale (AGI), désigne une machine dotée d'une intelligence au moins égale à celle de l'homme. Graal ou cauchemar ?
L'état de l'art actuel se limite à l'IA faible. Il est constitué d'une collection de systèmes d'IA dites étroites ou spécialisées.

> Il n'existe pas d’algorithme pour l'intelligence artificielle générale.
>
> Erik Larson (2021)

AltGR



> « Programmer les ordinateurs pour qu'ils apprennent par l'expérience pour, à la fin, être dispensé d'une grande partie de l'effort requis par une programmation détaillée. »
>
> Arthur L. Samuel (1959)

une autre forme de calcul

L'apprentissage statistique est une discipline de l'intelligence artificielle, dont l'origine remonte à 1959. De nombreux algorithmes d'apprentissage automatique ont ainsi été inventés et perfectionnés au cours des soixante-dix dernières années.

Le but de l'apprentissage statistique est de construire un programme capable d'effectuer un calcul sans décrire explicitement les étapes de ce calcul, mais en exploitant une base de données dite d'apprentissage. Il faut noter qu'il n'y a pas vraiment d'apprentissage au sens humain du terme. Pour fonctionner, ces programmes sont entraînés sur de vastes bases de données.

Lors de la phase d'entraînement, des algorithmes complexes vont détecter des relations entre les données de la base et exploiter les motifs et régularités présents pour déduire des règles descriptives. Ces relations sont sauvegardées sous la forme d'un modèle.

Lors de la phase d'exploitation du modèle, le programme va calculer une réponse à partir de données qu'il ne connaît pas. On parle souvent d'apprentissage automatique, ou *machine learning*.

AItGR

les principaux types d'apprentissage automatique

> « Ces résultats suggèrent que les techniques d'apprentissage automatique peuvent produire des contrôleurs plus robustes que les contrôleurs conventionnels programmés manuellement. »
>
> S. Russell, P. Norvig (1995)

[supervisé]

Le but est d'apprendre le lien entre des variables explicatives et une variable à expliquer. Les estimateurs supervisés ajustent un modèle qui renvoie la valeur de la donnée de sortie en fonction des données d'entrée. Si la variable à expliquer est une étiquette prise parmi un petit ensemble d'étiquettes, alors on parle d'une tâche de classification. Si la variable à expliquer est un nombre pris dans un ensemble grand, voire infini, on dit qu'il s'agit d'une tâche de régression. Pour fonctionner, ces programmes sont entraînés sur de vastes bases de données.

Ce que la machine a appris dans un contexte peut difficilement être réutilisé dans un autre.

[non supervisé]

Le but est d'apprendre les relations entre des variables explicatives, sans variable à expliquer. L'objectif peut être de découvrir des groupes d'exemples similaires, de déterminer la répartition des données, de réduire la dimensionnalité pour effectuer des projections ou de visualiser les données. Le système doit cibler les données selon leurs attributs disponibles. Une notion de similarité entre deux données explicatives est la plupart du temps indispensable.

Aucun expert n'est requis. L'algorithme doit découvrir par lui-même la structure plus ou moins *cachée* des données. C'est ensuite à l'homme d'associer ou déduire du sens à cette structure.

[par renforcement]

Le but est d'apprendre lorsque l'information vient d'une interaction avec l'environnement. Le programme (appelé "agent") commence sans information, ni fonction d'utilité. Il va bâtir un modèle prédictif de son environnement pour prendre des décisions. L'évaluation des décisions se fait en analysant les retours d'information fournis par l'environnement : les bonnes décisions sont liées à une récompense, et les mauvaises par une pénalisation. En cherchant à maximiser les récompenses reçues, le modèle prédictif s'améliore et devient ainsi capable de prédire la valeur d'une décision. Cette approche est particulièrement utilisée dans un contexte d'opposition face à un adversaire (jeux).

quelques algorithmes d'apprentissage automatique

> Trouver le bon estimateur est souvent la partie la plus difficile de la résolution d'un problème d'apprentissage automatique.
>
> scikit-learn.org

[modèle linéaire]

Ces algorithmes sont utilisés pour une première intention. Ils supposent que la variable à expliquer est une combinaison linéaire des variables explicatives. Les coefficients de la combinaison minimisent une fonction d'erreur. Ils sont souvent contraints pour garantir leur stabilité et éviter le surapprentissage.

ex : moindres carrés, lasso, ridge, régression logistique, GLM, perceptron

[classifieur bayésien naïf]

La classification bayésienne est basée sur l'application du théorème de Bayes. Ce théorème permet d'affiner une loi de probabilité au fur et à mesure des observations qui en sont tirées. Le classeur suppose naïvement que les variables explicatives sont toutes indépendantes, puis va apprendre la loi de probabilité de chacune de ces variables.

[arbre de décision]

Cet algorithme construit une structure arborescente de décisions où chaque branchement pose une question simple sur une variable explicative (ex: valeur >seuil ?). Les données sont partitionnées suivant les réponses à chacun des branchements rencontrés. Un ensemble d'arbres de décision aléatoires peuvent constituer une forêt aléatoire, la décision finale étant prise après le vote des différents arbres qui la constituent.

[machine à vecteurs de support]

Ces outils sont puissants et robustes. On les appelle souvent SVM, pour *support vector machine*. L'algorithme projette les données dans un espace de dimensionnalité plus grande afin de les séparer selon des hyperplans.

[plus proches voisins]

Lors de l'apprentissage, l'algorithme partitionne les points en groupes (*clusters*). Un point inconnu est rattaché à un groupe s'il est plus proche de ses points que des points des autres groupes. La notion de proximité est adaptée à la géométrie du problème.

[gradient stochastique]

Cet algorithme (SGD) accélère le calcul des paramètres libres d'un modèle sans sacrifier la précision. Il utilise une méthode de descente du gradient modifiée, basée sur une évaluation locale raisonnable et efficace. Il peut être utilisé pour l'entraînement de différents algorithmes d'apprentissage.

réseaux de neurones artificiels

« D'un point de vue informatique, les réseaux de neurones artificiels sont un triomphe de l'intelligence artificielle. »

Margareth Boden (2016)

des performances remarquables

Un réseau de neurones artificiels (*artificial neural networks*, ANN) est un groupe interconnecté de neurones dont le modèle de calcul s'inspire du traitement de l'information réalisé par les cellules nerveuses biologiques (McCulloch & Pitts, 1943). Le réseau forme un système adaptatif dont la structure peut évoluer en fonction des informations qui y circulent. La ressemblance avec un système nerveux biologique, même très rudimentaire, est toutefois très faible.

Les ANN forment une des plus anciennes classes d'algorithmes d'apprentissage automatique. Leur histoire est chahutée : les ANN ont leurs propres printemps et hivers ! Très utilisés en apprentissage automatique depuis les années 1990, ils ont beaucoup évolué à partir de 2010 avec l'arrivée des architectures profondes (*deep neural networks*).

Autrefois monocouches, les ANN sont aujourd'hui composés de dizaines de couches, comptant chacune un grand nombre de "neurones" (10^5-10^6). Les performances obtenues sur de grandes bases de données (images, textes) par ces algorithmes de *deep learning* sont impressionnantes.

L'approche proposée par les réseaux de neurones artificiels est théoriquement intéressante à bien des égards, tout en tant extrêmement importante d'un point de vue commercial.

Margareth Boden (2016)

AItGR

deep learning
apprentissage profond

> « Nos résultats montrent qu'un grand réseau de neurones convolutif profond est capable d'obtenir des résultats records sur un ensemble de données complexe en utilisant un apprentissage purement supervisé. »
>
> Geoffrey Hinton (2012)
> prix Turing 2018

[qu'est-ce que c'est?]

L'apprentissage profond (*deep learning*) est une classe d'algorithmes d'apprentissage automatique. C'est un type de réseaux de neurones (ANN) avec une architecture multicouches. Cette approche permet aux réseaux de traiter la hiérarchie dans des jeux de données complexes, et d'en extraire un modèle qui contient une représentation des connaissances. Les réseaux plus superficiels ne sont pas capables de détecter cette structure. Dans le cas des images, un algorithme *deep learning* va détecter les contrastes, puis les bords, certaines formes, des parties d'objets et enfin les objets.

[pourquoi ça marche?]

Le mécanisme de rétropropagation du gradient (*backpropagation*, 1974 et 1986) est au cœur de la phase d'entraînement du système. Il traite le problème de l'attribution de crédit : quels éléments d'un système complexe sont responsables d'une bonne prédiction ? L'algorithme fait remonter la responsabilité de la couche de sortie vers les couches cachées, en identifiant et adaptant les unités individuelles. Les calculs matriciels lourds, très utilisés dans le *deep learning*, ont fortement bénéficié des progrès réalisés sur le matériel informatique (processeurs GPUs).

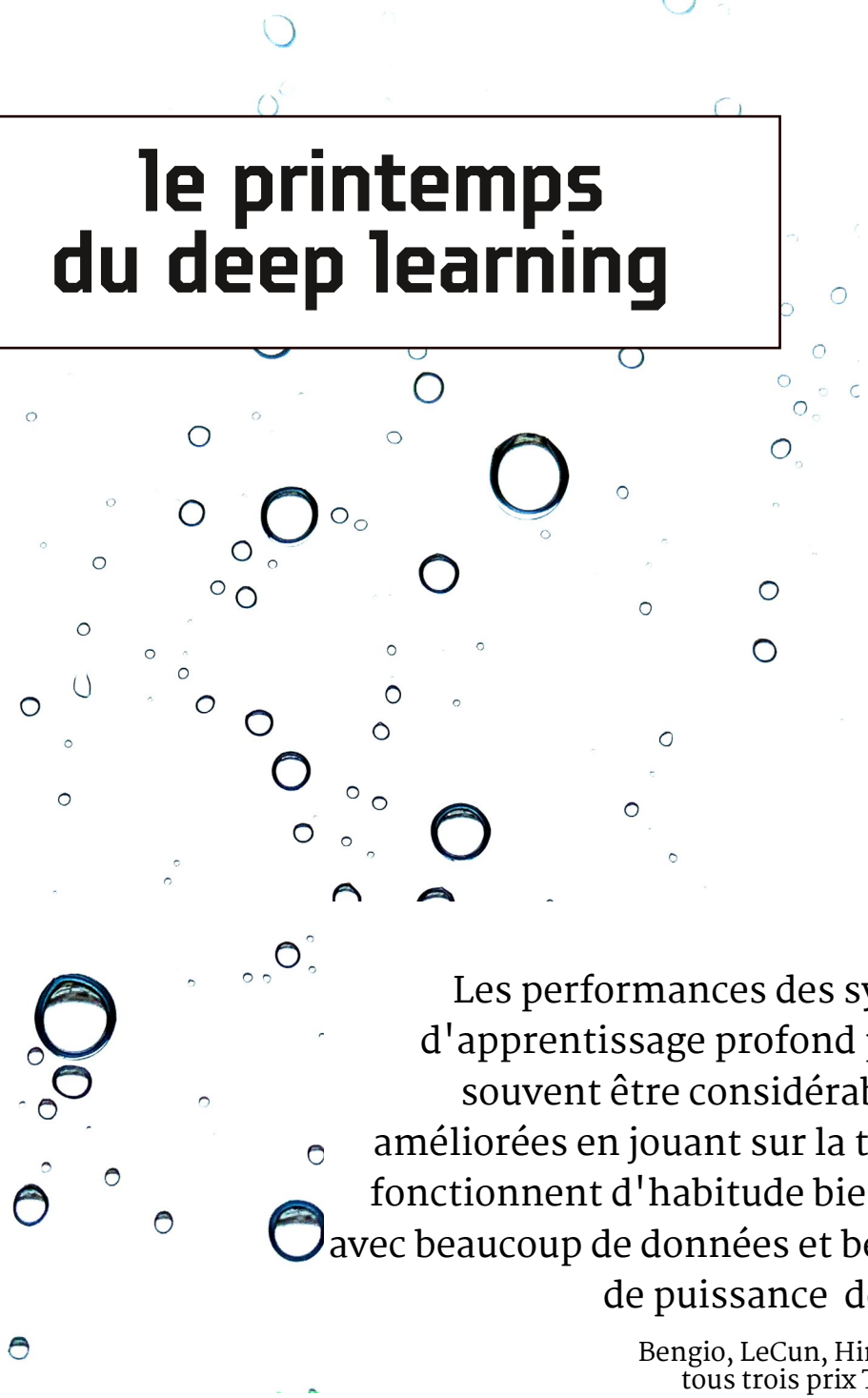
[est-ce nouveau?]

Les ANN sont contemporains d'Alan Turing et précèdent les premiers ordinateurs. Le perceptron de Rosenblatt (1958) est la première implémentation d'un ANN. Il stupéfie les journalistes car il sait reconnaître des lettres sans avoir été explicitement programmé. En 1969, cette approche, appelée alors connexionnisme, est déclarée être une impasse théorique (Minsky & Papert). Les ANN ne reviendront sur le devant de la scène que fin 1980 avec les approches cognitives dites PDP (*Parallel Distributed Processing*) avec notamment une généralisation de l'utilisation de l'architecture des réseaux multicouches.

[et demain?]

Les performances stupéfiantes du *deep learning* ont récemment suscité un grand enthousiasme et un fort battage publicitaire, que certains spécialistes jugent irresponsable De nombreux acteurs historiques du domaine ont d'ailleurs rejoint les grands acteurs du numérique. Cependant, s'il est indéniablement utile, cela ne signifie pas qu'il soit bien compris. règles de l'apprentissage multicouches restent théoriquement confuses, et les ANN se distinguent des cerveaux biologiques par d'innombrables aspects importants, dont certains ne sont pas encore connus.

le printemps du deep learning



> Les performances des systèmes d'apprentissage profond peuvent souvent être considérablement améliorées en jouant sur la taille. Ils fonctionnent d'habitude bien mieux avec beaucoup de données et beaucoup de puissance de calcul
>
> Bengio, LeCun, Hinton (2021)
> tous trois prix Turing 2018

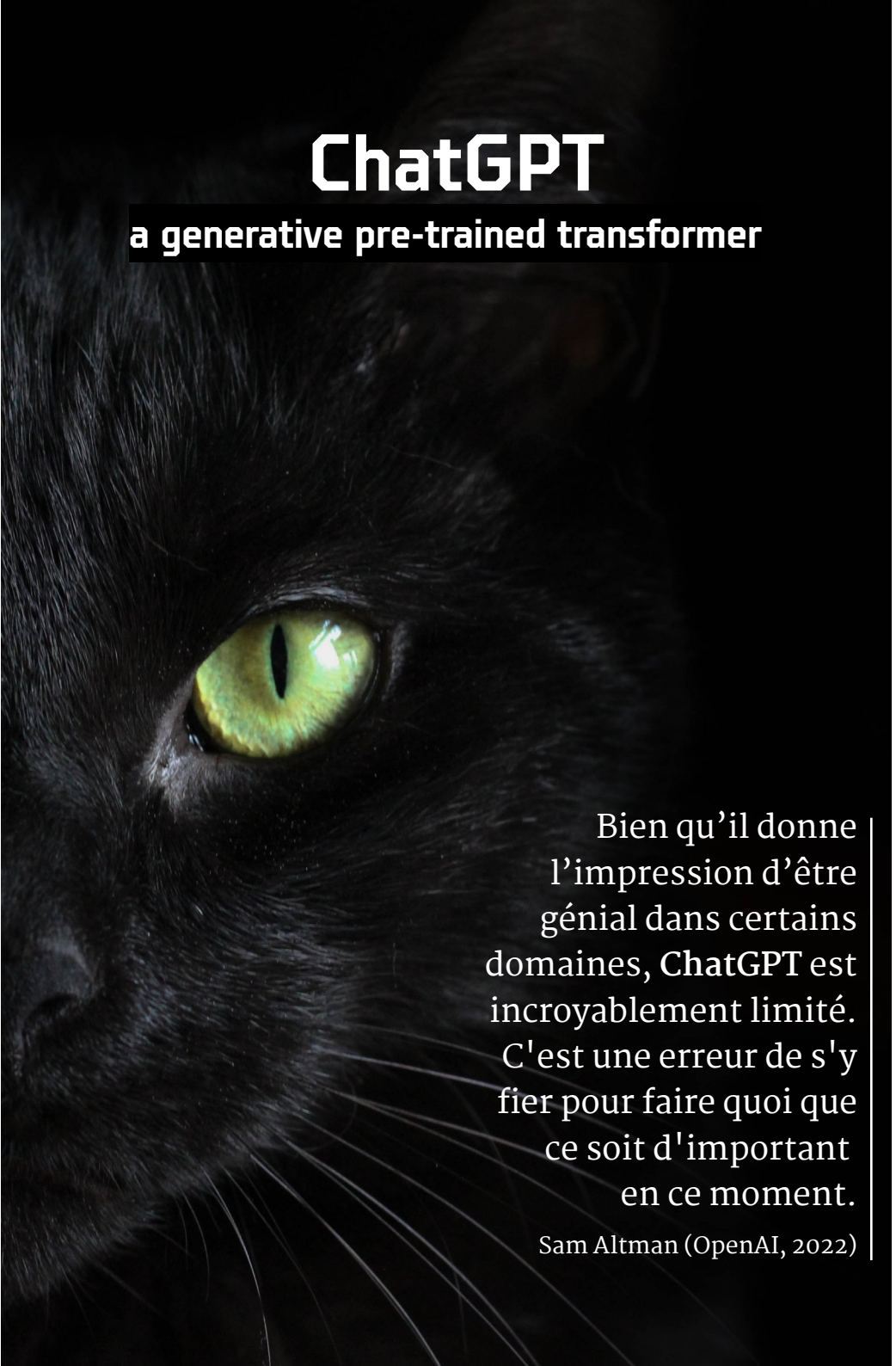
anatomie d'une bulle

Google rachète **DeepMind** en 2014 pour 650 millions de dollars. Cette petite société anglaise est fondée en 2010 par des chercheurs en IA. Avec une équipe de 25 personnes, **DeepMind** est spécialisée dans une activité assez particulière : elle apprend à des programmes de *deep learning* à jouer à des jeux vidéo, surtout des antiquités des années 1980 comme **Pong**, **Breakout** ou **Space Invaders**. Aux échecs, les programmes de **DeepMind** montrent de très belles performances face aux programmes spécialisés. Plus impressionnant encore, les champions humains du jeu de go s'inclinent à leur tour en 2016. **StarCraft 2** tombe lui aussi, en 2019. C'est le symbole du début du printemps du *deep learning* pour le grand public.

La montée en puissance du connexionniste était déjà observée par les experts du domaine. Les premiers systèmes capables d'exploiter efficacement la puissance de calcul des GPU remportent les concours de classification automatique d'images (Ciresan 2010, *AlexNet* 2012).

Les cas d'usage sont nombreux. Au-delà des performance immédiatement visibles sur le traitement d'image, les réseaux de neurones profonds sont des outils très utiles dans la détection et l'exploitation de tendances et motifs dans les bases de données massives.

La bulle est relancée régulièrement avec les arrivées des outils de génération de texte et d'image, et les agents conversationnels.



ChatGPT

a generative pre-trained transformer

Bien qu'il donne l'impression d'être génial dans certains domaines, ChatGPT est incroyablement limité. C'est une erreur de s'y fier pour faire quoi que ce soit d'important en ce moment.

Sam Altman (OpenAI, 2022)

« ChatGPT est une technologie fascinante qui a le potentiel de transformer la façon dont nous communiquons avec les machines. »

Andrew Ng

enfin, un agent conversationnel !

La conception d'agents conversationnels est classique en IA. Citons par exemple les grands anciens, dont le programme psychothérapeute **ELIZA** (Weizenbaum, 1966) et le programme d'assemblage de bloc **SHRDLU** (Winograd, 1979).

ChatGPT est un programme proposé par la société californienne **OpenAI** lancé en 2022. C'est un système d'IA spécialisée qui s'appuie sur un grand modèle de langage pour produire du texte (*large language model*, LLM) et une interface homme-machine en langage naturel appelée *prompt*.

La grande force de cet agent conversationnel tient à sa compréhension des directives humaines, souvent sous la forme de question ou d'ordre, et à sa capacité de produire des réponses d'une qualité souvent jugée satisfaisante. L'adoption par le grand public est massif, notamment pour un usage professionnel direct (site web) et indirect via une interface programmatique (API).

L'utilisation de **ChatGPT** par le grand public, début 2023, a été massive et soudaine.

intelligence artificielle générative

« Pourquoi les LLM semblent-ils meilleurs à générer du code qu'à générer du texte ? Un programme manipule des éléments qui évoluent dans un univers limité, discret, déterministe et entièrement observable.

Le monde réel n'a rien à voir avec cela. »

Yann LeCun (2023), prix Turing 2018

large language models

Un modèle de langage est un outil de traitement de langue naturelle. C'est un modèle statistique de la distribution de symboles (ex: lettres, mots, *token*) capable de prédire un mot à partir d'une séquence de mots fournie en entrée. Un *large langage model*(LLM) possède un grand nombre de paramètres internes, supérieur au milliard.

Les premières implémentation (2018) s'appuient sur des algorithmes de *deep learning*, notamment des *transformers*, entraînés sur de très grands corpus de texte (ex : **the Pile**). L'entraînement fait appel à une combinaison d'apprentissage supervisé (dont RHLF et *prompt engineering*) et non supervisé. Un processus d'ajustement (*fine tuning*) permet de les utiliser comme base pour le développement d'autres outils : on parle alors de *modèle de fondation*.

des outils versatiles

Les LLM servent principalement d'outils de génération de texte, utilisés dans les agents conversationnels (ex : **ChatGPT**, **Bard**). Certains LLM dits multimodaux sont entraînés à traiter des images. Les LLM les plus connus sont les modèles **GPT-{3,3.5,4}** d'**OpenAI** , **PaLM** de **Google AI**, **Chinchilla** de **DeepMind** et **LLaMa** de **Meta**.

D'autres systèmes, comme **MidJourney** et **Dall·E**, génèrent des images à partir de textes, en mariant les LLM à des technologies s'appuyant sur des modèles de diffusion probabiliste.

Ces outils d'IA générative font appel à des processus aléatoires pour atteindre un certain degré de créativité, mais leurs sorties résultent principalement d'un assemblage résultant d'un calcul statistique.

créativité et hallucinations

Les LLM capturent une grande partie de la syntaxe et de la sémantique du langage, ce qui leur permet de présenter des performances intéressantes en traduction, reformulation et résumé de contenu et plus généralement en classification et catégorisation de textes.

Ces programmes sont parfois appelés "perroquets stochastiques" car ils n'ont pas la compréhension sémantique des symboles qu'ils manipulent.

La créativité des IA génératives est une caractéristique inhérente aux algorithmes de génération actuels. Elle limite ainsi fortement la fiabilité des textes générés. Un LLM produit des textes avec une syntaxe correcte et une formulation très assurée : s'ils paraissent tout à fait plausibles, ils peuvent être subtilement faux. On parle d'hallucination.

data science

« Traiter des données pour valider une hypothèse ou estimer des paramètres est longtemps resté du ressort exclusif des statistiques. »

Stéphane Mallat, Collège de France

traiter les données massives

La science des données, ou *data science*, se situe au croisement de trois domaines : les mathématiques, l'informatique et l'algorithmique. C'est à la fois une discipline théorique et une pratique expérimentale. La science des données relève des statistiques et plus généralement des mathématiques appliquées. L'augmentation considérable de la masse des données a fait exploser la combinatoire des possibles. Cette *malédiction de la dimensionnalité* est une difficulté centrale de l'analyse de données. Modéliser et représenter les hiérarchies et structures cachées dans les données (image, textes, sons) n'est pas une mince affaire.

Le traitement de données massives et la démocratisation de nouveaux outils (ex: *scikit-learn* de INRIA) nécessitent d'intégrer des compétences en génie logiciel. Le rôle de *data scientist* apparaît ainsi, comme professionnel du traitement, de la manipulation et de la valorisation des données. Il maîtrise habituellement des outils informatiques comme **Python** ou **R**, ainsi que les bibliothèques logicielles spécialisées. Les algorithmes d'apprentissage statistique font partie des outils de *data science*. Ils sont pensés pour optimiser l'analyse des données à partir d'exemples. Ils sont à l'origine des résultats spectaculaires qu'on rattache aujourd'hui à l'intelligence artificielle.

aller plus loin
quelques références

illustrations

Les illustrations sont issues de **unsplash.com**.

Les photos ont été trouvées sur le net et issues des références ci-dessous.

références

Notes de cours de Stéphane Mallat au Collège de France : https://www.di.ens.fr/~mallat/CoursCollege.html

Mind as machine (2006) et *AI : Its Nature and Future* (2016) de Margaret Boden.

The Quest for Artificial Intelligence, Nils Nilsson (2010).

A Brief History of Artificial Intelligence: What It Is, Where We Are, and Where We Are Going Michael Wooldridge (2021).

The Myth of Artificial Intelligence: Why Computers Can't Think the Way We Do, Erik. J. Larson (2021).

Atlas of AI. Power, Politics, and the Planetary Costs of Artificial Intelligence, Kate Crawford (2021).

Computer: A History of the Information Machine, Campbell-Kelly, Aspray, Ensmenger et Yost (2023).

A New History of Modern Computing, Haigh et Ceruzzi (2021).

une histoire de l'IA
disponible sur
http://alt-gr.tech

AltGR

AltGR

algorithmique française

mathématiques
systèmes experts
apprentissage automatique

http://alt-gr.tech
aix-en-provence
+ lyon + paris
contact@alt-gr.tech