

## IA – Glossaire

---

### **AGI (Artificial General Intelligence)**

AGI (Artificial General Intelligence) ou Intelligence Artificielle Générale, est un terme utilisé pour décrire un type d'intelligence artificielle capable d'apprendre et de raisonner de manière autonome, similaire à l'intelligence humaine. Contrairement à l'intelligence artificielle spécialisée, qui est conçue pour des tâches spécifiques, l'AGI vise à créer des systèmes capables de maîtriser une variété de compétences et de connaissances.

### **ALGORITHME**

Le terme vient du nom du mathématicien persan Muhammad ibn Musa al-Khwarizmi (vers l'an 820), à qui l'on doit l'introduction en Occident de la numération décimale.

Suite finie de règles formelles (opérations logiques, instructions) permettant d'obtenir un résultat à partir d'éléments fournis en entrée. Cette suite peut être l'objet d'un processus automatisé d'exécution et s'appuyer sur des modèles conçus par le biais d'apprentissage machine.

### **ALGORITHME GÉNÉTIQUE**

Un algorithme génétique est une méthode d'optimisation inspirée de la théorie de l'évolution naturelle. Il utilise des techniques de sélection, de croisement et de mutation pour trouver des solutions optimales ou quasi-optimales à des problèmes complexes.

### **ALIGNEMENT**

L'alignement, en intelligence artificielle, fait référence à l'adéquation entre les objectifs, les valeurs et les comportements d'un système d'IA et ceux de ses concepteurs ou utilisateurs. Un système d'IA bien aligné agit de manière à respecter et à soutenir les intentions de ses utilisateurs, tout en évitant les comportements indésirables ou nuisibles. L'alignement est crucial pour garantir que les systèmes d'IA soient utilisés de manière responsable et éthique, et pour minimiser les risques potentiels associés à leur déploiement.

### **ANNOTATION (IA)**

L'annotation est le procédé par lequel les données sont décrites manuellement afin d'être caractérisées, par exemple en attribuant à une image de chien l'étiquette correspondante. On parle aussi de vérité terrain ou groundtruth.

### **APPRENTISSAGE AUTO-SUPERVISÉ**

C'est une méthode d'apprentissage automatique. Il apprend à partir d'échantillons de données non étiquetés. Il peut être considéré comme une forme intermédiaire entre l'apprentissage supervisé et non supervisé. Il est basé sur un réseau de neurones artificiels. Le réseau de neurones apprend en deux étapes. Tout d'abord, la tâche est résolue sur la base de pseudo-étiquettes qui aident à initialiser les poids du réseau. Deuxièmement, la tâche réelle est effectuée avec un apprentissage supervisé ou non supervisé

### **APPRENTISSAGE AUTOMATIQUE**

L'apprentissage machine permet de construire un modèle mathématique à partir de données, en incluant un grand nombre de variables qui ne sont pas connues à l'avance. Les paramètres sont configurés au fur et à mesure lors d'une phase d'apprentissage, qui utilise des jeux de données d'entraînement pour trouver des liens et les classifie. Les différentes méthodes d'apprentissage machine sont choisies par les concepteurs en fonction de la nature des tâches à accomplir (regroupement, arbre de décision). Ces méthodes sont habituellement classées en 3 catégories : apprentissage supervisé par un humain, apprentissage non supervisé et apprentissage non supervisé par renforcement.

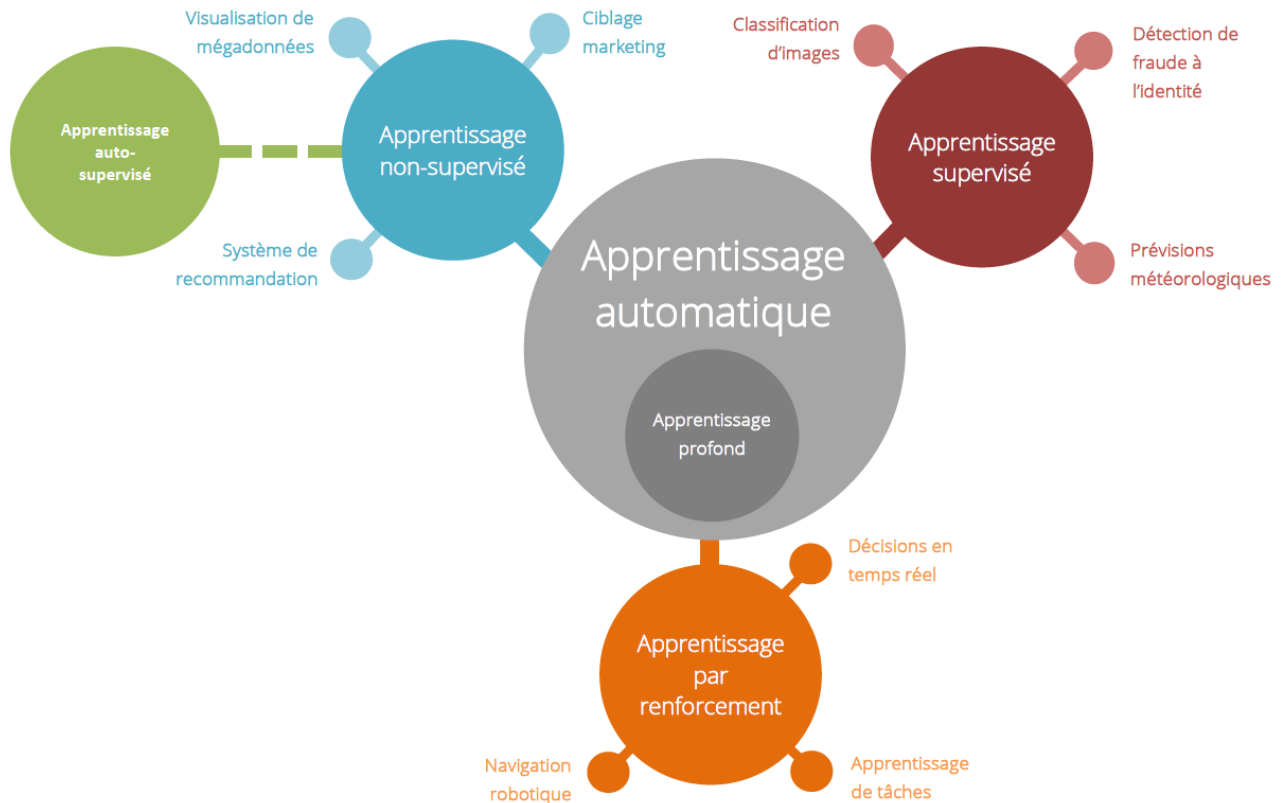
### **APPRENTISSAGE CONTINU**

L'apprentissage continu est la capacité d'un système à s'améliorer et à s'adapter à mesure qu'il intègre de nouvelles données, y compris pendant sa mise en service. Dans le cas de l'apprentissage continu, la phase d'apprentissage du système se poursuit pendant le déploiement du modèle.

### **APPRENTISSAGE FÉDÉRÉ**

L'apprentissage fédéré est basé sur la notion de *collaborative learning*, où plusieurs entités travaillent ensemble pour améliorer un modèle commun tout en préservant la confidentialité des données. Au lieu de centraliser toutes les données pour entraîner un modèle unique,

l'apprentissage fédéré utilise des modèles locaux déployés sur chaque appareil ou organisation participant au réseau. Ces modèles locaux sont ensuite mis à jour indépendamment avec les données locales, et seules les mises à jour du modèle sont partagées entre les participants



## APPRENTISSAGE NON SUPERVISÉ

L'apprentissage non supervisé est un procédé d'apprentissage automatique dans lequel l'algorithme utilise un jeu de données brutes et obtient un résultat en se fondant sur la détection de similarités entre certaines de ces données.

## APPRENTISSAGE PAR RENFORCEMENT

L'apprentissage par renforcement ou Reinforcement Learning est **une méthode de Machine Learning**. Elle consiste à entraîner des modèles d'intelligence artificielle d'une manière bien spécifique. L'agent IA doit apprendre à atteindre un objectif au sein d'un environnement incertain et potentiellement complexe. Pour y parvenir, **l'ordinateur essaie toutes les façons possibles et apprend de ses erreurs**.

**À chaque tentative, l'IA reçoit une récompense ou une punition** en fonction des actions effectuées. Elle est programmée pour maximiser sa récompense, et tentera donc de trouver la méthode le lui permettant.

## **APPRENTISSAGE PAR RENFORCEMENT ET RÉTROACTION HUMAINE – REINFORCEMENT LEARNING**

Approche d'apprentissage par renforcement qui utilise les commentaires et les évaluations d'utilisateurs humains pour guider l'apprentissage d'un modèle d'intelligence artificielle. Ce type d'apprentissage est utilisé dans les générateurs de texte fondés sur les grands modèles de langue.

## **APPRENTISSAGE PAR TRANSFERT**

En apprentissage automatique, l'apprentissage par transfert consiste à utiliser les connaissances acquises lors de l'apprentissage d'une tâche pour améliorer les performances sur une tâche analogue, généralement lorsque les données d'apprentissage sont limitées pour cette nouvelle tâche.

## **APPRENTISSAGE PROFOND (DEEP LEARNING)**

L'apprentissage profond est un procédé d'apprentissage automatique utilisant des réseaux de neurones possédant plusieurs couches de neurones cachées. Ces algorithmes possédant de très nombreux paramètres, ils demandent un nombre très important de données afin d'être entraînés.

## **APPRENTISSAGE SUPERVISÉ**

L'apprentissage supervisé est un procédé d'apprentissage automatique dans lequel l'algorithme s'entraîne à une tâche déterminée en utilisant un jeu de données assorties chacune d'une annotation indiquant le résultat attendu.

## **APPRENTISSAGE NON SUPERVISÉ -Clustering (Regroupement)**

Le clustering, ou regroupement, est une méthode d'apprentissage non supervisé qui vise à diviser un ensemble de données en groupes homogènes, appelés clusters, en fonction de leurs caractéristiques. Le clustering est utilisé dans divers domaines tels que l'analyse de marché, la segmentation des clients ou la détection d'anomalies.

## **ATTAQUE PAR EMPOISONNEMENT (DATA POISONING ATTACK)**

Les attaques par empoisonnement visent à modifier le comportement du système d'IA en introduisant des données corrompues en phase

d'entraînement (ou d'apprentissage). Elles supposent que l'attaquant soit en mesure de soumettre des données à utiliser lors de l'entraînement du système d'IA.

### **ATTAQUE PAR EXEMPLES CONTRADICTOIRES (ADVERSARIAL EXAMPLES ATTACK)**

Les attaques par exemples contradictoires visent à soumettre des entrées malicieuses ou corrompues au système d'IA en phase de production. Exemple : une image qui aurait été modifiée de façon à tromper un classifieur d'image et ainsi attribuer à une image de panda, l'étiquette de singe ...

### **ATTAQUE PAR EXFILTRATION DE MODÈLE (MODEL EVASION ATTACK)**

Les attaques par exfiltration de modèle visent à permettre le vol d'un modèle d'IA et/ou de ses paramètres et hyperparamètres. Le modèle constitue un actif de grande valeur pour un système d'IA

### **ATTAQUE PAR INFÉRENCE D'APPARTENANCE (MEMBERSHIP INFERENCE ATTACK)**

Les attaques par inférence d'appartenance visent à permettre à un attaquant d'acquérir des connaissances sur les données utilisées pour la production du modèle d'IA. En pratique, il s'agit de déterminer si des données relatives à un individu ont été utilisées lors de la phase d'entraînement

### **ATTAQUE PAR INVERSION DE MODÈLE (MODEL INVERSION ATTACK)**

Les attaques par inversion visent à reconstruire les données ayant servi pour l'apprentissage du système. En pratique, les attaques par inversion sont menées en soumettant un grand nombre d'entrées au système d'IA et en observant les sorties produites. On utilise, de façon équivalente, le terme d'attaque par inversion de donnée.

### **AUGMENTATION DE DONNÉES (IA)**

Dans le domaine de l'intelligence artificielle, le processus d'augmentation de données accroît la quantité de données d'entraînement par la

création de nouvelles données à partir des données existantes. Cette augmentation peut être réalisée par différentes opérations, par exemple, dans le cas d'images , par translation, rotation,ajout de bruit, etc

### **AUTOENCODEUR (Auto-encodeur)**

Un auto-encodeur est un type de réseau de neurones artificiels utilisé pour l'apprentissage non supervisé. Il apprend à compresser et à décompresser des données en réduisant leurs dimensions tout en préservant autant d'informations que possible. Les auto-encodeurs sont souvent utilisés pour la réduction de dimensionnalité, la détection d'anomalies et la génération de données.

### **BACKPROPAGATION (Rétropropagation)**

La Backpropagation est une méthode d'optimisation utilisée pour entraîner des réseaux de neurones. Elle consiste à calculer le gradient de la fonction d'erreur par rapport à chaque poids du réseau en propageant les erreurs de l'avant vers l'arrière. La rétropropagation est utilisée conjointement avec la descente de gradient pour ajuster les poids du réseau et minimiser la fonction d'erreur.

### **BAGGING (Bootstrap Aggregating)**

Le bagging est une méthode d'apprentissage ensembliste qui consiste à entraîner plusieurs modèles sur des sous-ensembles aléatoires de données d'entraînement, puis à combiner leurs prédictions pour obtenir une prédiction finale. Le bagging permet de réduire la variance et d'améliorer la stabilité des modèles.

### **BOOSTING (Amplification)**

Le boosting est une méthode d'apprentissage ensembliste qui consiste à entraîner plusieurs modèles de manière séquentielle, où chaque modèle tente de corriger les erreurs des modèles précédents. Les prédictions des modèles individuels sont ensuite combinées pour obtenir une prédiction finale. Le boosting permet d'améliorer la précision des modèles.

## **BULLE DE FILTRE**

Phénomène principalement observé sur les réseaux sociaux où les algorithmes de recommandation – Selon cette théorie, des algorithmes sélectionnent « discrètement » les contenus visibles par chaque internaute, en s'appuyant sur différentes données collectées sur lui.

## **CALCUL MULTIPARTITE SÉCURISÉ**

Le calcul multipartite sécurisé (en anglais, « secure multi-party computation ») est une branche de la cryptographie permettant à plusieurs entités (ou parties) de calculer conjointement une fonction sur leurs données respectives, sans que celles-ci ne soient divulguées aux autres participants et tout en étant assuré que le résultat est exact....

## **CARACTÉRISTIQUE (IA)**

Dans le domaine de l'intelligence artificielle, la caractéristique (feature en anglais) est la variable utilisée pour représenter une propriété définie d'une entité ou d'un objet. Il peut s'agir d'informations relatives à la forme, la texture, ou encore à la couleur d'une image.

## **CHABOT**

Le chatbot est une intelligence artificielle qui utilise des capacités de traitement du langage naturel pour mener une conversation. À terme, le but est que le chatbot puisse répondre aux questions comme un humain pourrait le faire. Certains diront qu'on en est encore loin.

## **CHIFFREMENT HOMOMORPHE**

Le chiffrement homomorphe est une technique de cryptographie permettant de réaliser des opérations sur des données chiffrées sans que celles-ci aient à être déchiffrées. Le résultat de ces opérations reste sous forme chiffrée et ne peut être déchiffré que par les destinataires autorisés

## **CLASSIFICATION (IA)**

La classification est une méthode de catégorisation qui consiste à attribuer une classe ou catégorie à une entrée qui lui est soumise en fonction de sa proximité à la classe en question selon des critères bien

choisis. Exemple : un système d'IA entraîné pour la reconnaissance d'images d'animaux...

### **CONFIDENTIALITÉ DIFFÉRENTIELLE**

La confidentialité différentielle désigne une propriété mathématique assurant que la présence ou l'absence d'un individu dans une base de données n'affecte pas le résultat obtenu par un processus d'anonymisation appliqué à celle-ci. Pour l'atteindre, l'ajout d'un bruit spécifique est généralement...

### **COUCHE DE NEURONES**

Organisation des neurones dans un réseau. Il n'y a pas de connexion entre les neurones d'une même couche : les connexions ne se font qu'avec les neurones de la couche suivante. Généralement, chaque neurone d'une couche est lié avec tous les neurones de la couche en aval et celle-ci uniquement

### **CRITÈRE D'ARRÊT (IA)**

Élément de contrôle de l'évolution d'un algorithme d'apprentissage automatique qui permet, s'il est atteint, d'arrêter le processus itératif. Il peut prendre différentes formes : la fonction de perte (qui permet de calculer l'erreur de prédiction) a suffisamment décru, le nombre d'itérations fixé est atteint etc

### **CROSS-VALIDATION (Validation croisée)**

La validation croisée est une technique d'évaluation des modèles d'apprentissage automatique utilisée pour estimer leur performance sur de nouvelles données. Elle consiste à diviser les données en plusieurs sous-ensembles, puis à entraîner et à valider le modèle sur chaque sous-ensemble, en prenant la moyenne des performances pour obtenir une estimation plus robuste

### **DATAMINING (analyse et fouille de données)**

Le datamining permet d'analyser un grand volume de données et d'en faire ressortir des modèles, des corrélations, des tendances.



## **DEEP LEARNING (apprentissage profond)**

Voir apprentissage automatique et neurones.

## **DÉRIVE DES DONNÉES**

Variation des données utilisées à l'étape de production par rapport aux données qui ont été utilisées pour tester et valider le modèle avant son déploiement. Plusieurs facteurs peuvent entraîner cette dérive : des modifications de processus en amont, des problèmes de qualité des données, etc

## **DÉRIVE DU MODÈLE**

La dérive du modèle est la perte d'adéquation entre un modèle et la tâche qu'il doit accomplir. Cette dérive peut résulter d'un réapprentissage du modèle ou d'une évolution de l'environnement dans lequel il s'applique (domaine d'emploi).

## **DÉSAPPRENTISSAGE MACHINE**

Le désapprentissage machine est une technique liée à l'apprentissage automatique visant à faire disparaître une information des connaissances apprises par un modèle d'IA. L'objectif de ces techniques est de supprimer une information contenue dans un modèle sans avoir à ré-entraîner ce dernier. ...

## **DOMAINE D'EMPLOI (IA)**

Dans le domaine de l'intelligence artificielle, le domaine d'emploi est la description de l'environnement et de la population visée par le procédé d'apprentissage automatique....

## **DONNÉE BRUTE (IA)**

Dans le domaine de l'intelligence artificielle, une donnée brute est une donnée n'ayant subi aucune transformation depuis son observation initiale.

## **DONNEES A CARACTERE PERSONNEL**

Informations relatives à une personne physique identifiée ou identifiable, directement ou indirectement, par référence à un ou plusieurs éléments qui lui sont propres.

Parmi celles-ci, les données sensibles au sens du règlement général sur la protection des données concernent les données à caractère personnel relatives à l'origine raciale ou ethnique, les opinions politiques, les convictions religieuses ou philosophiques, l'appartenance syndicale ainsi que les données génétiques, les données biométriques, les données concernant la santé ou concernant la vie sexuelle ou l'orientation sexuelle.

### **DONNÉE DE SORTIE (IA)**

Dans le domaine de l'intelligence artificielle, une donnée de sortie est une valeur représentant tout ou partie de l'opération effectuée par le système d'IA à partir des données d'entrée.

### **DONNÉE D'ENTRÉE (IA)**

Dans le domaine de l'intelligence artificielle, une donnée d'entrée est une donnée utilisée pour l'apprentissage automatique ou la prise de décision du système d'IA (en phase de production).

### **ÉCHANTILLON (IA)**

Dans le domaine de l'intelligence artificielle, l'échantillon est une fraction représentative d'une population ou d'un univers statistique.

### **ENSEMBLE DE TEST (IA)**

Jeu de données utilisé lors de la phase de test.

### **ENSEMBLE DE VALIDATION (IA)**

Jeu de données utilisé lors de la phase de validation

### **ENSEMBLE D'ENTRAÎNEMENT/D'APPRENTISSAGE**

Jeu de données (texte, sons, images, listes, etc.) utilisé lors de la phase d'entraînement / d'apprentissage : le système s'entraîne sur ces données pour effectuer la tâche attendue de lui.

### **ENTRAÎNEMENT (OU APPRENTISSAGE)**

L'entraînement est le processus de l'apprentissage automatique pendant lequel le système d'intelligence artificielle construit un modèle à partir de données.

## **ENVIRONNEMENT D'EXÉCUTION DE CONFIANCE**

Les Trusted Execution Environment, ou environnement d'exécution de confiance, sont des zones sécurisées et isolées des autres environnements d'exécution situés dans un processeur. Les TEE garantissent que des données confidentielles restent stockées, traitées et protégées dans un environnement de confiance

## **EMBEDDING (Incorporation)**

L'embedding, ou incorporation, est une technique couramment utilisée en intelligence artificielle, et en particulier dans le traitement du langage naturel, pour transformer des entités discrètes, comme les mots, en vecteurs de nombres continus. Ces représentations vectorielles, ou embeddings, ont la propriété de capturer les relations sémantiques entre les entités. Par exemple, dans un espace d'embedding, les mots avec des significations similaires se trouvent proches les uns des autres. Les embeddings ne sont pas limités au texte : on peut également avoir des embeddings pour des entités comme les utilisateurs et les produits dans un système de recommandation, où les embeddings capturent des aspects de leur comportement ou de leurs caractéristiques.

## **ESTIMATION BAYÉSIENNE**

L'estimation ou inférence bayésienne s'appuie sur un théorème énoncé par le mathématicien Thomas Bayes. Ce théorème donne une méthode pour calculer la probabilité d'un phénomène grâce à la connaissance de certaines informations. L'estimation bayésienne est donc la méthode qui s'appuie sur ce raisonnement...

## **ESTIMATION DE POSES**

Technique de vision par ordinateur (« computer vision ») qui permet de détecter la posture d'une personne sur une image ou une vidéo, afin d'en extraire un modèle biomécanique.

## **EXPLICABILITÉ (IA)**

Dans le domaine de l'intelligence artificielle, l'explicabilité est la capacité de mettre en relation et de rendre compréhensible les éléments pris en compte par le système d'IA pour la production d'un résultat. Il peut s'agir,

par exemple, des variables d'entrée et de leurs conséquences sur la prévision d'un score et, ainsi sur la décision

### **EXTRACTION DE CARACTÉRISTIQUES (FEATURE EXTRACTION)**

Dans le domaine de l'intelligence artificielle, l'extraction de caractéristiques est l'étape au cours de laquelle sont induites depuis des données brutes (fichier son, image, document textuel, tableau numérique, etc.) des caractéristiques (features) sur lesquelles le système d'IA doit se reposer pour effectuer la tâche pour laquelle il est programmé.

### **FEW SHOT LEARNING**

Méthode d'apprentissage automatique où un modèle est entraîné à classer des items en fonction de leur similarité à partir de très peu de données.

### **FONCTION DE PERTE OU DE COÛT (LOSS FUNCTION)**

Dans le domaine de l'intelligence artificielle, Une fonction de perte est une fonction mathématique qui quantifie la différence entre les valeurs prédites et réelles dans un modèle d'apprentissage automatique . Il mesure les performances du modèle et guide le processus d'optimisation en fournissant des informations sur son adéquation aux données.

### **FONCTION D'ACTIVATION**

Dans le domaine de l'intelligence artificielle, la fonction d'activation peut être vu comme l'équivalent du « potentiel d'activation » qu'on retrouve dans les neurones biologiques. Cette fonction détermine si un neurone artificiel doit être activé ou pas et, dans le premier cas, le degré de cette activation

### **FORÊTS ALÉATOIRES (RANDOM FORESTS)**

Les forêts aléatoires sont une méthode d'apprentissage automatique ensembliste, se basant sur de multiples arbres de décision entraînés sur des sous-ensembles de données légèrement différents.

### **GABARIT FACIAL**

En matière de reconnaissance faciale, un « gabarit » désigne les mesures qui sont mémorisées lors de l'enregistrement des caractéristiques d'un visage. Les données extraites pour constituer ce gabarit sont des données biométriques au sens du RGPD (article 4-14).

### **GAN** (Generative Adversarial Network, Réseaux antagonistes génératifs)

Les GAN (Generative Adversarial Networks, réseaux antagonistes génératifs) sont une classe d'algorithmes d'apprentissage automatique utilisés pour générer de nouvelles données qui ressemblent à un ensemble de données d'entraînement. Les GAN sont composés de deux réseaux de neurones, un générateur et un discriminateur, qui sont entraînés ensemble dans un processus compétitif

### **GÉNÉRATION DE TEXTE** (Text Generation)

La génération de texte est un sous-domaine du traitement du langage naturel qui se concentre sur la création de textes en langage humain par des machines. Les applications incluent la rédaction automatique d'articles, la création de résumés, et la génération de contenu pour les chatbots ou les assistants vocaux.

### **GENERATIVE AI**

La Generative AI, ou intelligence artificielle générative, désigne un ensemble de techniques d'apprentissage automatique qui permettent aux systèmes d'IA de créer du contenu nouveau et original à partir de données d'entraînement. Les modèles de Generative AI, tels que GPT, sont capables de générer du texte, des images, de la musique et d'autres types de contenu en imitant les caractéristiques et les styles présents dans les données d'entraînement. La Generative AI est utilisée dans diverses applications, allant de la génération de texte à la création artistique et à la conception assistée par ordinateur.

### **GPT**

GPT (Generative Pre-trained Transformer) est une architecture de modèle de langage développée par OpenAI. Elle utilise un mécanisme d'auto-attention pour apprendre et générer du texte de manière cohérente et pertinente. GPT est largement utilisé pour diverses tâches

de traitement du langage naturel, telles que la génération de texte, la traduction automatique et la réponse aux questions.

## **GRADIENT**

Le gradient est, en mathématiques, un vecteur représentant la variation d'une fonction au voisinage d'un point donné (en pratique, lorsqu'on dessine une courbe, plus le gradient est élevé, plus la « pente » de la courbe est forte). Appliqué au cas de l'apprentissage d'un modèle d'IA, le gradient est utilisé pour mettre en œuvre l'algorithme de descente de gradient (ou algorithme de la plus forte pente). Ce dernier permet d'obtenir un résultat optimal selon certains critères (par exemple : minimisation d'une fonction de perte), de manière itérative, c'est-à-dire par une succession d'étapes.

## **Grand modèle de langage LLM (Large Language Model)**

Un grand modèle de langage (LLM) est un modèle de langage de grande envergure, généralement basé sur des réseaux de neurones profonds, qui a été entraîné sur un vaste ensemble de données textuelles. Les LLM sont capables de comprendre, de générer et de répondre à des questions en langage naturel avec une précision et une cohérence élevées. Ces modèles ont révolutionné le domaine du traitement du langage naturel et sont utilisés dans diverses applications, telles que les assistants vocaux, les chatbots et la génération de texte.

## **HEURISTIQUE (Heuristic)**

Une heuristique est une méthode de résolution de problèmes basée sur des règles approximatives ou des principes pratiques plutôt que sur des algorithmes rigoureux. Les heuristiques sont souvent utilisées en **intelligence artificielle** pour résoudre des problèmes pour lesquels une solution optimale est difficile à trouver, en fournissant des solutions acceptables en un temps raisonnable.

## **HYPERPARAMÈTRE**

Élément indépendant de l'apprentissage tels que le nombre de nœuds et la taille des couches cachées du réseau de neurones, l'initialisation des poids, le coefficient d'apprentissage, la fonction d'activation, etc.

## **IA FAIBLE:**

correspond au périmètre de l'intelligence artificielle tel qu'on le connaît, selon l'UNESCO. En résumé, une IA faible est capable d'exécuter une (ou plusieurs) tâches de façon autonome, mais dans un cadre défini par l'homme. Elle n'a pas de conscience ou de sensibilité à l'inverse de l'**IA forte** : un concept qui relève de la fiction à ce jour, rappelle l'organisme.

## **IA FORTE**

L'IA forte, ou intelligence artificielle forte, est un concept d'IA qui décrit un système informatique capable de posséder une compréhension et une conscience similaires à celles des humains. Contrairement à l'IA faible, l'IA forte serait capable de raisonner, d'apprendre et de s'adapter de manière générale, et pourrait potentiellement surpasser l'intelligence humaine dans tous les domaines.

## **INTELLIGENCE ARTIFICIELLE**

L'intelligence artificielle (IA) est un processus d'imitation de l'intelligence humaine qui repose sur la création et l'application d'algorithmes exécutés dans un environnement informatique dynamique. Son but est de permettre à des ordinateurs de penser et d'agir comme des êtres humains.

## **MACHINE LEARNING** (apprentissage automatisé)

Le Machine Learning (apprentissage automatique) est une méthode d'analyse de données qui permet aux machines d'apprendre sans être explicitement programmées. Il s'agit d'un sous-domaine de l'intelligence artificielle..

## **METADONNEES**

Données qui permettent de définir, de contextualiser ou de caractériser d'autres données.

## **MIND UPLOADING**

Selon les transhumanistes\*, nos sensations, idées et émotions se résument à des connections neuronales. Le téléchargement de l'esprit (mind uploading) est l'idée transhumaniste selon laquelle le « contenu » du cerveau humain est réductible à un ensemble d'informations que l'on pourrait traduire dans le code binaire informatique, et donc télécharger (upload) dans un ordinateur.

## **MODÈLE (IA)**

Le modèle d'IA est la construction mathématique générant une déduction ou une prédiction à partir de données d'entrée. Le modèle est estimé à partir de données annotées lors de la phase d'apprentissage (ou d'entraînement) du système d'IA.

## **MODÈLE DE LANGAGE**

Modèle statistique de la distribution d'unités linguistiques (par exemple : lettres, phonèmes, mots) dans une langue naturelle. Un modèle de langage peut par exemple prédire le mot suivant dans une séquence de mots. On parle de modèles de langage de grande taille ou « Large Language Models » (LLM) en anglais pour les modèles possédant un grand nombre de paramètres (généralement de l'ordre du milliard de poids ou plus) comme GPT-3, BLOOM, Megatron NLG, Llama ou encore PaLM.

## **MODÈLE DISCRIMINATIF**

Modèle capable de réaliser une prédiction quant à l'appartenance à une classe pour des données nouvelles sur la base d'un apprentissage réalisé auparavant sur un jeu de données d'entraînement.

## **MODÈLE GÉNÉRATIF**

Modèle défini par opposition à un modèle discriminatif. Il permet à la fois de générer de nouveaux exemples à partir des données d'entraînement et d'évaluer la probabilité qu'un nouvel exemple provienne ou ait été généré à partir des données d'entraînement.

## **MULTI MODAL**

Le terme multi modal fait référence aux systèmes d'intelligence artificielle capables de traiter et de comprendre différents types de données, tels que le texte, les images, les vidéos et les sons. Les modèles multi modaux sont conçus pour apprendre des représentations communes à partir de différentes sources de données, ce qui leur permet d'être plus polyvalents et performants dans diverses tâches et applications.

## **NÉO-CONNEXIONNISME**

Théorie née dans le domaine des sciences cognitives et des neurosciences, le néo-connexionnisme propose d'élaborer des modèles informatiques qui visent à simuler des phénomènes d'apprentissage par



des réseaux de neurones formels\*, dont l'organisation et le fonctionnement ont été pensés par analogie avec les systèmes neuronaux physiologiques.

## **NEURONE ARTIFICIEL**

Un neurone artificiel fonctionne d'une manière inspirée de celle d'un neurone biologique : un nœud d'un réseau de plusieurs neurones reçoit généralement plusieurs valeurs d'entrée et génère une valeur de sortie. Le neurone calcule la valeur de sortie en appliquant une fonction d'activation à une à une somme pondérée des valeurs d'entrée.

## **ONTOLOGIE**

En informatique, une ontologie est la modélisation d'un ensemble de données par des concepts et relations issues de connaissances dans un domaine donné (par exemple, géographie, médecine, agriculture, etc.).

## **OPEN DATA**

Le terme désigne la mise à disposition publique, par téléchargement, de bases de données structurées. Ces données sont ré-employables de manière non-onéreuse dans les conditions d'une licence spécifique, pouvant notamment préciser ou prohiber certaines finalités de réemploi. L'open data n'est pas à confondre avec de l'information publique unitaire disponible sur des sites internet, dont l'intégralité de la base n'est pas téléchargeable (par exemple les bases de données de jurisprudence). Il ne se substitue pas aux modes de publicité obligatoire de certaines mesures ou décisions administratives ou judiciaires déjà édictés par certaines lois ou règlements.

## **PaLM-2 :**

acronyme de Pathways Language Model-2, modèle de langage développé par Google et entraîné sur plusieurs milliards de paramètres. Il est destiné à alimenter Google Bard, le générateur de texte lancé par la firme en février 2023.

## **PARAMÈTRE (IA)**

Dans le domaine de l'intelligence artificielle, le paramètre est la propriété apprise des données utilisées pour l'entraînement (par exemple le poids de chaque neurone d'un réseau).

## **PARTITIONNEMENT DE DONNÉES**

Le partitionnement de données (clustering en anglais) est une méthode ayant pour but de diviser un ensemble de données en différents sous-ensembles homogènes, c'est-à-dire partageant des caractéristiques communes. Ces caractéristiques reposent sur des critères de proximité définis en introduisant la notion de distance entre les objets.

## **PROFILAGE**

Selon l'article 4(4) du RGPD, il s'agit d'un traitement de données à caractère personnel réalisé afin d'évaluer certains aspects de la vie d'une personne physique (situation économique, santé, préférences personnelles, etc.)

## **PSEUDONYMISATION**

Selon l'article 4 du RGPD, il s'agit d'un traitement de données à caractère personnel de telle façon que celles-ci ne puissent plus être attribuées à une personne concernée précise sans avoir recours à des informations supplémentaires, pour autant que ces informations supplémentaires soient conservées séparément et soumises à des mesures techniques et organisationnelles afin de garantir que les données à caractère personnel ne sont pas attribuées à une personne physique identifiée ou identifiable.

## **RECHERCHE ARBORESCENTE (Tree search)**

La recherche arborescente est une méthode de résolution de problèmes en intelligence artificielle qui explore les différentes possibilités de manière systématique en utilisant une structure d'arbre. La recherche arborescente est couramment utilisée pour résoudre des problèmes de jeu, de planification et de prise de décision. Les algorithmes de recherche arborescente, tels que la recherche en profondeur, la recherche en largeur et l'algorithme  $A^*$ , parcourent l'arbre en examinant les nœuds et en choisissant la meilleure option en fonction d'une heuristique ou d'une fonction d'évaluation. Cette méthode permet d'identifier les solutions

optimales ou approximatives aux problèmes complexes en limitant l'espace de recherche et en réduisant le temps de calcul.

## **RECONNAISSANCE D'ENTITÉS NOMMÉES**

En anglais « Named-entity recognition » (NER), sous-tâche d'extraction d'informations qui cherche à localiser et classier les mentions d'entités nommées dans du texte non structuré en catégories prédéfinies, emplacements, codes médicaux, expressions de temps, quantités, valeurs monétaires,...

## **RÉCURSIVITÉ (Recursivity)**

La récursivité est un concept en informatique et en intelligence artificielle où une fonction ou un algorithme fait appel à lui-même pour résoudre un problème. La récursivité est souvent utilisée pour résoudre des problèmes complexes de manière élégante et efficace, tels que la recherche d'arbres, la factorielle et la suite de Fibonacci.

## **RÉDUCTION DE DIMENSION OU DIMENSIONNALITÉ** Méthode

permettant de diminuer la quantité d'information en ne conservant que le strict nécessaire, permettant ainsi d'obtenir plus d'efficacité en termes de résultats et de temps d'analyse. Cette réduction de l'information utile peut se faire par sélection des caractéristiques les plus appropriées que celles de départ.

## **RÉGRESSION**

La régression est un ensemble de méthodes d'analyse statistique permettant d'approcher une variable à partir d'autres qui lui sont corrélées. En apprentissage automatique, on distingue les problèmes de régression des problèmes de classification. Ainsi, on considère que les problèmes de prédiction d'une variable quantitative sont des problèmes de régression tandis que les problèmes de prédiction d'une variable qualitative sont des problèmes de classification....

## **Représentation sémantique des connaissances**

Algorithmes\* permettant de formaliser une phrase écrite dans une langue quelconque (exemple : « Paul prend le bus pour Berlin ») sous forme logique et ainsi de la rendre interprétable par un ordinateur. La machine peut alors réaliser des inférences logiques (comme une

déduction), qui lui permettent de classer les mots dans différentes catégories et d'analyser les phrases qu'on lui soumet.

## **RÉSEAU DE NEURONES ARTIFICIELS (ARTIFICIAL NEURAL NETWORK)**

Dans le domaine de l'intelligence artificielle, un réseau de neurones artificiels est un ensemble organisé de neurones interconnectés permettant la résolution de problèmes complexes tels que la vision par ordinateur ou le traitement du langage naturel. Il s'agit d'un type particulier d'algorithmes d'apprentissage automatique (comme les machines à vecteur de support, arbres de décision, K plus proches voisins, etc.) caractérisés par un grand nombre de couches de neurones, dont les coefficients de pondération sont ajustés au cours d'une phase d'entraînement (apprentissage profond)

## **RÉSEAU NEURONAL CONVOLUTIF(Convolutional Neural Network, CNN)**

Un réseau neuronal convolutif (Convolutional Neural Network, CNN) est un type de réseau de neurones artificiels spécialisé dans le traitement et l'analyse des images. Les CNN sont largement utilisés dans la reconnaissance d'images, la classification et la détection d'objets.

## **RLHF: Reinforcement Learning from Human Feedback (Apprentissage par renforcement à partir des retours humains)**

Le RLHF, ou Reinforcement Learning from Human Feedback, est une méthode d'apprentissage automatique qui combine l'apprentissage par renforcement et les retours humains pour améliorer les performances d'un agent d'intelligence artificielle. Dans cette approche, un agent apprend à travers un processus itératif en recevant des récompenses ou des punitions basées sur les retours humains, plutôt que de simplement apprendre à partir de données préexistantes. Les retours humains fournissent un signal de récompense précieux qui aide l'agent à mieux comprendre et à se conformer aux intentions de l'utilisateur et aux objectifs à long terme. Le RLHF est particulièrement utile dans les situations où il est difficile de définir des récompenses précises ou d'obtenir des données d'entraînement de haute qualité.

## **ROBOTIQUE**

La robotique est un domaine de recherche et d'ingénierie qui vise à concevoir, construire et utiliser des robots pour accomplir diverses tâches, souvent dans le but de remplacer ou d'assister l'humain. La robotique fait appel à l'intelligence artificielle pour permettre aux robots de prendre des décisions et de s'adapter à leur environnement.

## **ROBUSTESSE (IA)**

Dans le domaine de l'intelligence artificielle, la résilience est la capacité du système à maintenir sa conformité à des exigences de performance et/ou de sécurité en présence de données d'entrée extérieures à son domaine d'emploi (par exemple en raison d'un défaut sur un capteur).

## **SCIENCES DES DONNEES (data science)**

Regroupement large des mathématiques, des statistiques, des probabilités, de l'informatique, de la visualisation des données afin d'extraire des connaissances d'un ensemble hétérogène de données (images, son, texte, données génomiques, liens dans les réseaux sociaux, mesures physiques, etc.).

## **SEGMENTATION DES DONNÉES**

La segmentation des données est une méthode permettant la division d'un corpus de données en plusieurs ensembles (par exemple d'entraînement, de validation et de test), soit à partir de critères objectifs (date, âge, etc.) soit de manière aléatoire.

## **SEMI-SUPERVISED LEARNING (Apprentissage semi-supervisé)**

L'apprentissage semi-supervisé est une méthode d'apprentissage automatique qui combine les approches supervisées et non supervisées. Un algorithme apprend à partir d'un ensemble de données partiellement étiquetées, utilisant à la fois les exemples étiquetés et non étiquetés pour améliorer la performance du modèle.

## **Stochastic Gradient Descent (SGD, Descente de gradient stochastique)**

La descente de gradient stochastique (Stochastic Gradient Descent, SGD) est une méthode d'optimisation utilisée pour minimiser une

fonction objectif en mettant à jour les paramètres du modèle par petites étapes proportionnelles au gradient négatif de la fonction objectif. La SGD est couramment utilisée pour entraîner des réseaux de neurones et des modèles d'apprentissage automatique. Contrairement à la descente de gradient classique, la SGD met à jour les paramètres du modèle en utilisant un sous-ensemble aléatoire des données d'entraînement à chaque étape, ce qui permet de réduire le temps de calcul et d'éviter les minimums locaux.

### **Support Vector Machine (SVM, Machine à vecteurs de support)**

Une Support Vector Machine (SVM, machine à vecteurs de support) est un algorithme d'apprentissage supervisé utilisé pour la classification et la régression. Les SVM fonctionnent en trouvant le meilleur hyperplan qui sépare les classes ou prédit les valeurs cibles avec la plus grande marge possible.

### **SURAPPRENTISSAGE (OVERFITTING)**

Dans le domaine de l'intelligence artificielle, le surapprentissage entraîne un modèle qui correspond trop précisément à une collection particulière de données utilisées pour l'entraînement. Cette analyse risque de ne pas permettre une utilisation fiable du système d'IA.

### **SUPERVISED LEARNING :**

relatif à un procédé de machine learning baptisé « apprentissage supervisé » en français. Le principe : entraîner un programme à prédire un résultat en fonction d'occurrences constatées sur une base d'exemples. L'apprentissage supervisé est principalement utilisée pour classifier et détecter des anomalies ou établir des probabilités dans un contexte donné. Exemple : la détection de spams.

### **SYSTÈME D'IA À USAGE GÉNÉRAL**

Un système d'IA à usage général peut être utilisé et adapté à un large éventail d'applications pour lesquelles il n'a pas été conçu intentionnellement et spécifiquement. Un tel système est destiné à exécuter des fonctions génériques telles que la reconnaissance d'images et de paroles, la génération d'images audio et vidéo, la détection de formes, la réponse à des questions, la traduction, etc.

## **TAUX D'APPRENTISSAGE (LEARNING RATE)**

Dans le domaine de l'intelligence artificielle, le taux d'apprentissage est le facteur multiplicatif appliqué au gradient. À chaque itération, l'algorithme de descente de gradient multiplie le taux d'apprentissage par le gradient. Le taux d'apprentissage est un hyperparamètre qui joue sur la rapidité de la descente de gradient : un nombre d'itérations plus ou moins important est nécessaire avant que l'algorithme ne converge, c'est-à-dire qu'un apprentissage optimal du réseau soit réalisé.

## **TEST (IA)**

Dans le domaine de l'intelligence artificielle, le test est un processus consistant à évaluer les performances d'un système et à rechercher des erreurs liées à l'exécution d'un algorithme ou d'un programme en s'appuyant sur des jeux de données d'entrée n'ayant pas été utilisés lors de la phase d'entraînement.

## **TOKEN (Jeton)**

Un token (jeton en français) est une unité de texte utilisée pour représenter et traiter des séquences de mots ou de caractères dans les modèles de langage et les chatbots. Les tokens permettent aux modèles de langage, tels que GPT, d'apprendre et de générer du texte de manière plus efficace et précise en traitant de plus petites unités de texte.

## **TRAITEMENT AUTOMATIQUE DE LA PAROLE**

Ensemble de disciplines dont l'objectif est la captation, la transmission, l'identification et la synthèse de la parole. Ces disciplines rassemblent notamment la reconnaissance de la parole, la synthèse de la parole, l'identification du locuteur ou encore la vérification du locuteur.

## **TRAITEMENT AUTOMATIQUE DU LANGAGE NATUREL (NATURAL LANGUAGE PROCESSING OU NLP)**

Le traitement automatique du langage naturel est un domaine multidisciplinaire impliquant la linguistique, l'informatique et l'intelligence artificielle. Il vise à créer des outils de capable d'interpréter et de synthétiser du texte pour diverses applications.

## **TRANSFORMER**

correspond à un réseau de neurones artificiels développé initialement pour le traitement automatique du langage, notamment pour la traduction ou la génération de textes. Les Transformers peuvent « réaliser l'essentiel, voire tout leur apprentissage de manière non supervisée, c'est-à-dire sur des textes « bruts » sans plus d'information ou d'annotation sur leur contenu » selon Michel Poujol,

## **UNSUPERVISED LEARNING :**

aussi appelée « apprentissage non-supervisé » en français, cette méthode d'apprentissage automatique doit, comme son nom l'indique, permettre à un programme de produire un résultat sans supervision à partir de données brutes. Selon la CNIL, le principe s'articule autour de « la détection de similarités » entre les données étudiées. L'apprentissage non-supervisé peut être, par exemple, utilisé par les algorithmes de recommandations, à même de prédire un comportement ou besoin en fonction des habitudes de navigation.

## **VALIDATION (IA)**

Dans le domaine de l'intelligence artificielle, la validation est un processus consistant à expérimenter, observer et optimiser (en modifiant les hyperparamètres notamment) le comportement du système lors de son exécution. Elle permet de s'assurer, dans le domaine d'emploi, de l'adéquation des données de sortie avec les résultats attendus.

## **VECTEUR (Vector)**

Dans le domaine de l'intelligence artificielle (IA), un vecteur est une structure de données fondamentale qui se présente comme une liste ordonnée de nombres. Ces nombres représentent des caractéristiques ou des traits dans un espace multidimensionnel. Dans l'apprentissage automatique, par exemple, les vecteurs peuvent représenter les traits d'un individu ou d'un objet, comme l'âge, le poids et la taille d'une personne. En traitement du langage naturel, les mots ou les phrases peuvent être convertis en vecteurs pour un traitement plus efficace. L'importance du vecteur en IA découle de sa capacité à représenter des données de manière structurée et à permettre des manipulations mathématiques complexes.



## **VISION PAR ORDINATEUR (COMPUTER VISION)**

La vision par ordinateur est une branche de l'intelligence artificielle dont le principal but est de permettre à une machine d'analyser et traiter une *ou* plusieurs images ou vidéos prises par un système d'acquisition.

## **VIE ARTIFICIELLE**

Champ de recherche interdisciplinaire visant à créer des systèmes artificiels s'inspirant des systèmes vivants, sous la forme de programmes informatiques ou de robots.

## Annexe : Bibliographie

- [Glossaire de l'intelligence artificielle \(unesco.org\)](https://unesco.org)
- <https://www.coe.int/fr/web/artificial-intelligence/glossary>
- [Glossaire de l'intelligence artificielle \(IA\) | CNIL](https://www.cnil.fr)
- [https://www.cnil](https://www.cnil.fr)
- <https://www.cnil.fr/fr/intelligence-artificielle/glossaire-ia>