

## L'intelligence artificielle ou IA

## L'intelligence artificielle fait partie de l'informatique.

Les trois sont des **machines artificielles programmées.**

De son côté, **le robot** test un ordinateur doté d'actionneurs et de moteurs qui lui permettent d'agir sur son environnement de manière physique.

▶▶▶ Qu'il soit classique ou qu'il utilise des fonctions d'intelligence artificielle, **tout programme informatique est un algorithme, c'est-à-dire est une séquence d'instructions qui va être exécutée pour résoudre un problème.** En tant qu'humain nous pouvons résoudre des algorithmes simples rapidement, comme par exemple trier une série de 10 chiffres du plus petit au plus grand. Lorsque nous en avons 10 000 à trier, l'ordinateur sera beaucoup plus rapide.

L'Intelligence Artificielle qualifie l'ensemble des technologies qui permettent aux machines de réaliser des tâches complexes en utilisant des moyens matériels et logiciels conçus par des humains. Ces systèmes dotés d'IA peuvent percevoir leur environnement, analyser des données, appliquer un raisonnement logique et prendre des décisions pour atteindre un objectif établi.

## FICHE SCIENTIFIQUE N°1 – SUITE

## ▶▶▶ Prenons l'exemple d'une recette de crêpes

Lorsqu'on suit une recette de crêpes traditionnelle, chaque étape est clairement définie et précise. Il faut ajouter les ingrédients dans un ordre spécifique, les mélanger, faire cuire la pâte dans une poêle et retourner les crêpes. C'est un algorithme classique, où chaque action est prévue et programmée à l'avance.

L'algorithme d'IA a besoin des mêmes ingrédients de base, mais au lieu d'avoir des instructions précises pour chaque étape, il **peut être plus flexible et s'adapter à différentes situations.**

*Par exemple, il peut prendre en compte vos préférences personnelles en matière de goût, comme la quantité de sucre ou la consistance de la pâte, et ajuster les quantités en conséquence pour vous offrir des crêpes plus personnalisées.*

*Il peut également ajuster la cuisson en fonction de facteurs externes, tels que la température de la pièce ou l'humidité ambiante.*

*Dernier exemple, il peut analyser les données en temps réel pour prendre des décisions plus intelligentes sur le moment idéal pour retourner les crêpes ou ajuster la chaleur de la poêle.*

► **L'ÉQUIPE DU QUAI DES SAVOIRS**

# FICHE SCIENTIFIQUE N°2

## Il existe plusieurs méthodes d'Intelligence Artificielle

▶▶▶ L'une de ces méthodes se nomme **l'IA symbolique**. L'idée est d'utiliser **la logique et d'apprendre à partir de connaissances données par des experts humains**. C'est la méthode la plus ancienne qui est encore utilisée aujourd'hui. Elle repose sur la manipulation de symboles et de règles logiques pour résoudre des problèmes. Elle est bien adaptée à des domaines où le raisonnement humain est essentiel, tels que le diagnostic médical, le support technique ou la planification stratégique.

*Exemple: cela peut être utile pour la planification l'ordre dans lequel un livreur doit livrer ses colis et pour savoir quels colis sont attribués à quel livreur. L'IA symbolique est celle qui nous entoure le plus, quand il y a des contraintes précises à respecter et que l'on souhaite avoir une explication et une validation de la solution proposée.*

▶▶▶ L'**IA numérique** est une autre méthode qui utilise des **statistiques et beaucoup de données** pour fonctionner. La machine apprend en analysant les informations contenues dans des données: c'est-ce qu'on appelle le Machine Learning.

**Le Machine Learning (en français, apprentissage automatique)**, est une méthode qui entraîne des algorithmes à faire ce que nous attendons d'eux à partir de données. Cette méthode s'est beaucoup développée récemment grâce à l'utilisation d'ordinateurs plus puissants et rapides, et la possibilité, avec internet, d'avoir accès à de grandes quantités de données. Ces données ce sont **des informations ou des faits** collectés ou enregistrés au format papier ou numérique. Cela peut être des nombres, des mots, des images, des sons, etc.

Les machines ont besoin d'apprendre, mais elles enregistrent différemment les informations reçues. Dans la machine, il y a de multiples algorithmes d'IA qui peuvent être combinés pour la faire fonctionner.

*Par exemple, pour qu'un algorithme reconnaisse des images de chat, on lui montre beaucoup d'images de chats. Alors qu'un bébé humain n'a besoin de regarder que 3 images de chat pour reconnaître ensuite un chat sans jamais se tromper, une machine doit en "voir" des millions. Ces images sont les **données d'apprentissage** qui permettent à la machine d'apprendre les caractéristiques communes des images de chats qui lui ont été donné.*

# FICHE SCIENTIFIQUE N°2 – SUITE

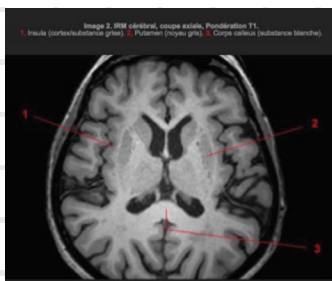
## Il existe plusieurs types d'IA numérique fonctionnant avec le Machine Learning

### ▶▶▶ L'apprentissage supervisé

Dans cette méthode, l'algorithme s'entraîne sur des images de chats qui sont étiquetées par des humains comme étant des chats. La machine analyse ces images et cherche des caractéristiques communes pour apprendre à reconnaître les images de chats.

Ce travail est souvent fait par des humains, plus ou moins qualifiés selon la difficulté des images. On les appelle les « travailleurs du clic ».

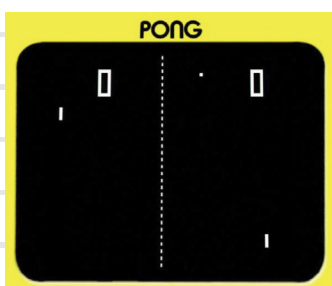
*Exemple: l'apprentissage supervisé va permettre de prédire une valeur, comme le prix d'une maison en fonction des caractéristiques ou encore d'identifier des images répréhensibles sur les réseaux sociaux.*



### ▶▶▶ L'apprentissage non supervisé

Dans cette méthode, l'algorithme doit trouver des ressemblances entre les images sans avoir d'autres informations. La machine regroupe les images reçues en fonction de formes similaires pour déterminer les caractéristiques communes des chats.

*Exemple: des centaines de données de radiographies sont collectées, elles sont analysées et pré-traitées en enlevant les radiographies avec des données trop différentes, cela va créer une radiographie modèle qui sera considéré comme normalité. Une fois le modèle entraîné, il est utilisé pour détecter les comportements anormaux dans le réseau en comparant les nouvelles observations aux schémas appris. Les activités qui ne correspondent pas aux modèles normaux sont identifiées comme des anomalies potentielles. Lorsqu'une anomalie est détectée, une alerte peut être déclenchée.*



### ▶▶▶ L'apprentissage par renforcement

Dans cette méthode, l'algorithme n'a pas non plus d'étiquettes sur les images de chats. Mais elle reçoit des récompenses positives ou négatives si elle se trompe ou non.

*Exemple: une IA va être entraînée à jouer à un jeu vidéo comme le jeu Pong, il doit donc déplacer la raquette vers le haut, vers le bas ou ne rien faire. Il va jouer et comprendre via des « récompenses » et des « sanctions » lorsqu'il joue bien ou mal. L'entraînement est terminé lorsque cette IA sait jouer au jeu avec une stratégie optimale.*

# FICHE SCIENTIFIQUE N°3

## Les Deep Learning et les Deepfakes

▶▶▶ Le **Deep Learning**, ou apprentissage profond en français, permet à la machine d'apprendre à partir de beaucoup plus de données que les autres méthodes du Machine Learning. Les **réseaux de neurones artificiels profonds** sont les algorithmes spécifiques sur lesquels repose cette méthode.

La spécificité du Deep Learning est que la machine peut apprendre à partir des **données brutes**, sans qu'il y ait besoin de les trier préalablement, comme pour les autres méthodes du Machine Learning.

Bien sûr, ce ne sont pas les mêmes neurones que ceux de notre cerveau, ils sont bien moins performants que les nôtres! Ces **réseaux de neurones** artificiels profonds sont des structures **mathématiques** qui s'inspirent du fonctionnement des neurones biologiques.

Les **neurones mathématiques** sont presque comme des personnes qui font des calculs à tour de rôle: chaque personne reçoit des informations, fait un calcul et donne le résultat à la personne suivante. La nouvelle personne fait la même chose: prendre le résultat, fait un nouveau calcul et donne au suivant un nouveau résultat. Les personnes qui calculent, ce sont les différentes couches du réseau de neurones.

▶▶▶ Il y a 2 types de réseaux de neurones: les réseaux simples qui ne contiennent qu'une seule couche entre les informations du début et les résultats définitifs; et les réseaux de neurones profonds qui comportent chacun des centaines de couches de calcul!

Revenons à nos images de chats. Un réseau de neurones profond est entraîné sur un grand nombre d'images de chats. Or, chaque image est **composée de pixels**. Le réseau de neurones analyse les motifs et les caractéristiques des pixels pour apprendre à les reconnaître. Les **premières couches** du réseau détectent **les formes** basiques, tandis que les couches suivantes identifient des **caractéristiques** plus complexes. Les dernières couches servent à déterminer avec précision si une image donnée **représente un chat ou non**.

Il faut d'abord que les algorithmes passent des **tests**, peu importe la méthode de Machine Learning utilisée. Dans notre exemple, ils doivent déterminer correctement s'il y a un chat ou non sur des images, dont les experts connaissent déjà la réponse.

Si l'algorithme fait **trop d'erreurs**, on **recommence** l'entraînement et il est modifié pour améliorer ses performances.

▶▶▶ Il existe différentes techniques de **Deep Learning**. L'une d'entre elle est particulièrement approprié à notre sujet puisqu'elle est capable de **générer des contenus multimédias**, tels que des vidéos, des images ou des audios qui peuvent sembler authentiques à première vue. Ces contenus falsifiés sont nommés **Deepfakes** (faux semblant) et peuvent être source de désinformation, de manipulation, voire d'atteintes à la vie privée.

# FICHE SCIENTIFIQUE N°3 – SUITE

## Comment fonctionnent les programmes d'intelligence artificielle qui produisent des Deepfakes et comment pouvoir différencier ces productions artificielles de la réalité ?

▶▶▶ Les Deepfakes sont produits par des programmes de type **GAN** ou **réseau antagoniste génératif**. Il s'agit de deux réseaux de neurones adversaires qui vont se perfectionner, grâce à leur mise en concurrence. L'un des deux réseaux génère une image synthétique, c'est le réseau générateur, et l'autre recherche les défauts de l'image générée en la comparant avec des données réelles, c'est le réseau discriminateur. Au fur et à mesure de l'entraînement, le générateur devient plus habile à générer des données réalistes, tandis que le discriminateur devient plus habile à les distinguer.

Ce type de programme d'intelligence artificielle s'améliore donc de jour en jour et produit donc des résultats de plus en plus réalistes. Les **dangers** associés aux IA génératives et aux Deepfakes sont la **désinformation**, par la manipulation des images et des paroles de personnalités publiques par exemple, ou encore le **harcèlement** et l'**atteinte à la vie privée**, avec la production de fausses images ou vidéos compromettantes. L'altération de preuves dans les domaines juridiques et journalistiques peut également amener à une **perte de confiance des contenus délivrés**.

▶▶▶ Cela pose également des questions d'éthique puisque si l'IA a été entraînée avec des contenus discriminants, elle les reproduira et pourra même les amplifier, sans discernement.

Tous les programmes de **Machine Learning** et de **Deep Learning** se basent sur le **Big Data**, c'est-à-dire l'ensemble des innombrables données numériques produites chaque jour dans le monde entier par nos utilisations d'internet et des réseaux sociaux. Des données qui ne sont pas exemptes de biais, ni de failles - à l'image des humains !

Ces failles sont décelées puis corrigées au fur et à mesure mais voici quelques clés qui vous aideront à **détecter un contenu d'IA** aujourd'hui :

▶▶▶ Tout d'abord, utiliser **son esprit critique** pour analyser une information et connaître **la source** est très important. Lorsque l'image que vous regardez comporte des incohérences sur l'anatomie humaine par exemple comme pour les yeux ou les mains, ou alors quand l'écrit est présentable mais peu compréhensible, cela doit éveiller vos soupçons. Il est également possible de générer des voix et de modifier des vidéos. Il existe plusieurs outils pour détecter des contenus générés par des IA qui sont sur disponibles sur internet comme les recherches d'images inversées pour détecter si c'est une image produite par une intelligence artificielle.

▶▶▶ Le problème des modèles d'IA est leur opacité ainsi que le fait qu'ils soient basés sur des calculs complexes effectués par les réseaux de neurones. C'est ce qu'on appelle le paradoxe de la **boîte noire** : Nous pouvons savoir ce qui entre dans la boîte noire mais pas ce qui s'y passe à l'intérieur. La recherche actuelle tend à rendre plus transparentes ces informations.

Après lecture de ces fiches, nous espérons que vous serez mieux armés pour comprendre et analyser les prochains contenus générés par des intelligences artificielles.

▶ L'ÉQUIPE DU QUAI DES SAVOIRS