PARTIE 1 : COMPRENDRE

# ACTEURS DE LA FORMATION : MAÎTRISEZ L'I.A

**AUTEUR : ARNAUD GUIOVANNA**
Service Transformation Numérique des Territoires
Direction Aménagement du Numérique

# SOMMAIRE

# LE PROGRAMME COMMUNOTIC

Ce livre électronique est offert **aux adhérents** du dispositif **Communotic**, piloté par la **Région Normandie, et à l'ensemble des acteurs de la formation**. Il fait suite à la conférence du **8 octobre 2024** organisée par la **Communauté urbaine Le Havre Seine Métropole** et **Communotic**. Intitulée *« Acteurs de la formation : apprenez à concevoir et former des I.A »*, cette conférence avait pour objectif de donner un large aperçu de l'Intelligence Artificielle, notamment les LLM (Large Language Models ou Grands Modèles de Langage), qui constituent la brique technologique fondamentale de l'IA générative, de leurs conceptions, de leurs usages et de leurs applications dans le secteur de la formation. Le présent document est une synthèse de cet événement et approfondit également certains concepts.

Ce travail est rendu possible grâce à **Communotic**, un programme d'accompagnement mis en œuvre depuis plus d'une décennie par la Région Normandie. Destiné à tous les acteurs de la formation en région, ce dispositif vous permet d'adhérer et de bénéficier gratuitement de nombreuses animations, ateliers ou formations. L'objectif principal de **Communotic** est d'aider les acteurs de la formation dans leur transformation digitale.

## LE PILOTE DU PROGRAMME

**Arnaud Guiovanna** travaille depuis dix ans dans le secteur de la formation, mettant à profit son expertise en numérique. Il accompagne les organisations dans l'adoption de solutions innovantes, telles que les technologies éducatives, l'Intelligence Artificielle et les outils No-Code. Actuellement, il pilote le programme Communotic au sein de la Région Normandie.

**POUR TOUTE QUESTION OU POUR ADHERER A COMMUNOTIC :**

arnaud.guiovanna@normandie.fr

ou communotic@normandie.fr

# INTRODUCTION

Dans un monde en constante évolution, l'intelligence artificielle s'impose comme un levier incontournable de transformation dans tous les secteurs, et la formation ne fait pas exception. Les modèles de langage de grande taille (LLM) redéfinissent les interactions avec les technologies, offrant aux acteurs de la formation de nouvelles perspectives pour enrichir les expériences d'apprentissage.

C'est pourquoi, nous avons conçu cette série de deux livres **« Partie 1 : Comprendre »** et **« Partie 2 : Pratiquer »** pour accompagner les acteurs de la formation de ce nouveau paradigme de l'Intelligence Artificielle. Cette série a pour objectif de faire monter en compétences les professionnels de la formation sur ce vaste sujet. Ce premier livre de la série "Acteurs de la formation : Maîtrisez l'IA" se concentre sur la compréhension des concepts fondamentaux de l'IA. Il a pour ambition de vous fournir une base solide pour mieux

appréhender vos futures applications d'intelligence artificielle dans le domaine de la formation. Nous explorerons l'histoire et les fondements de l'IA, avant de plonger dans les rouages des modèles de langage. Vous découvrirez comment ces technologies transforment le paysage éducatif et comment elles peuvent


être utilisées pour créer des environnements d'apprentissage plus personnalisés et engageants.

Que vous soyez novice ou professionnel aguerri, ce livre vous guidera pas à pas à travers des explications accessibles et des exemples concrets. Vous comprendrez non seulement le fonctionnement des LLM, mais aussi les enjeux éthiques et pratiques liés à leur utilisation dans la formation. En parcourant ces pages, vous serez en mesure de faire vos premiers pas vers l'intégration de l'IA dans vos outils pédagogiques.

Ce livre vous permettra de poser les bases pour aborder la pratique dans le second volet de la série. Pour cela, nous vous montrerons comme utiliser les outils no-code afin d'intégrer des LLM et de bâtir vos premières applications IA. Ce second volet est en cours d'écriture, il sera publié très prochainement pour le réseau Communotic.

# PARTIE 1 : COMPRENDRE L'INTELLIGENCE ARTIFICIELLE

## CHAPITRE 1 : LES FONDEMENTS DE L'IA - HISTOIRE ET ÉVOLUTION

### 1. DU MYTHE A LA REALITE SCIENTIFIQUE

L'Intelligence Artificielle, bien qu'elle semble être une technologie moderne, puise ses racines dans des mythes et des récits anciens. L'idée de créer des êtres artificiels capables de penser et d'accomplir des tâches complexes fascine l'humanité depuis des millénaires. Dès 1635 avant J.-C., **le mythe d'Atrahasis** évoquait l'idée de créer l'**Homme** pour soulager les dieux de leurs tâches. À travers l'histoire, des récits comme **Frankenstein** de **Mary Shelley** (1818) ou le film **Metropolis**

METROPOLIS, FRITZ LANG (1927)

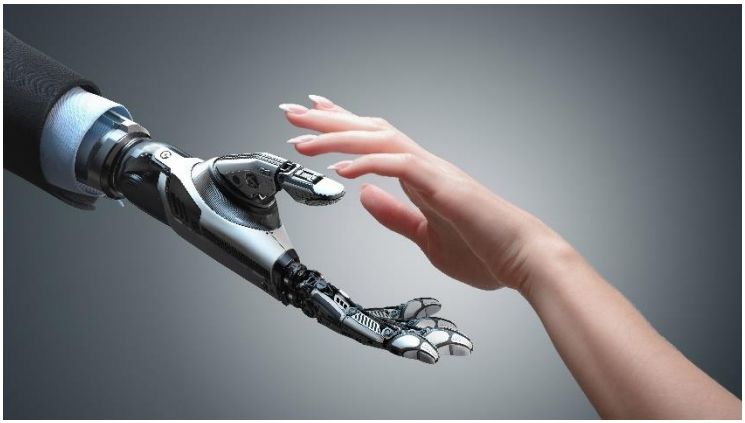
de **Fritz Lang** (1927) ont nourri l'imaginaire collectif sur la création de machines intelligentes.

Cependant, ce n'est qu'au 20e siècle que l'IA est devenue une réalité scientifique. En 1956, **la Conférence de Dartmouth**, souvent considérée comme le point de départ officiel de l'IA, a réuni des scientifiques tels qu'**Allan Newell** et **Herbert A. Simon** pour discuter de la possibilité de développer des machines capables d'imiter l'intelligence humaine. Ce fut également l'année où **Newell** et **Simon** créèrent le premier programme de raisonnement logique, le **Logic Theorist**.

CONFERENCE DE DARTMOUTH (1956)

L'histoire de l'IA se poursuit avec le développement du **Perceptron** par **Frank Rosenblatt** en 1958, le premier algorithme à apprendre à partir de données, précurseur des réseaux de neurones modernes. À partir de

là, l'IA a évolué, passant par des jalons marquants tels que **Deep Blue**, l'ordinateur d'IBM qui a battu le champion du monde d'échecs en 1997, et l'essor du **Deep Learning** en 2006, qui a propulsé les réseaux neuronaux. Aujourd'hui, les Large Language Models (LLM) comme **GPT** (2018) dominent le paysage, tandis que de futures technologies comme **JEPA** sont déjà en préparation, annonçant une nouvelle ère de capacités cognitives des machines.



La phrase « ***L'intelligence artificielle est un vieux souhait de jouer à dieu*** » citée par l'auteure de **Machines Who Think** (**Pamela McCorduck**) met en lumière l'une des motivations profondes derrière la quête humaine pour développer des technologies capables de penser, notamment l'intelligence artificielle. Pamela McCorduck, dans cet ouvrage, explore l'histoire et les perspectives de l'Intelligence Artificielle et met en lumière l'ambition ancienne de l'humanité de créer des entités pensantes. Elle explique que cette quête incarne un "vieux souhait de jouer à dieu", une idée évoquant la tentative humaine de reproduire l'intelligence, traditionnellement considérée comme un attribut divin.

***« Futur ? La panne des imaginaires »* de Nicolas Nova** est un essai qui explore la manière dont les visions du futur sont en crise dans nos sociétés contemporaines. Nicolas Nova s'intéresse à l'appauvrissement des imaginaires collectifs concernant les futurs possibles. Nous vivons dans un monde où ces promesses d'avenir semblent appartenir au passé : « Un futur antérieur ». Cela reflète une forme de stagnation dans notre capacité à imaginer des futurs véritablement novateurs. Les imaginaires collectifs du futur sont souvent basés sur des idées anciennes, qui ne correspondent plus à la réalité actuelle ou aux défis contemporains. Autrement dit, notre époque est coincée dans des récits du passé concernant le futur.

## 2. IA SYMBOLIQUE ET IA CONNEXIONNISTE

L'IA peut être classée en deux grandes approches : **Symbolique** et **Connexionniste**.

L'IA symbolique, parfois appelée IA logique ou basée sur des règles, repose sur la manipulation de symboles et l'utilisation de règles logiques pour résoudre des problèmes. Cette approche est guidée par des règles explicites définies par des humains. Par exemple, dans un système expert, les décisions sont prises selon des règles du type « Si X, alors Y ». Cette forme d'IA est souvent utilisée dans des contextes où les règles sont bien définies.

- **Exemple IA symbolique :** *Un système qui conseille de prendre un parapluie si la météo prévoit de la pluie est une application simple de l'IA symbolique.*

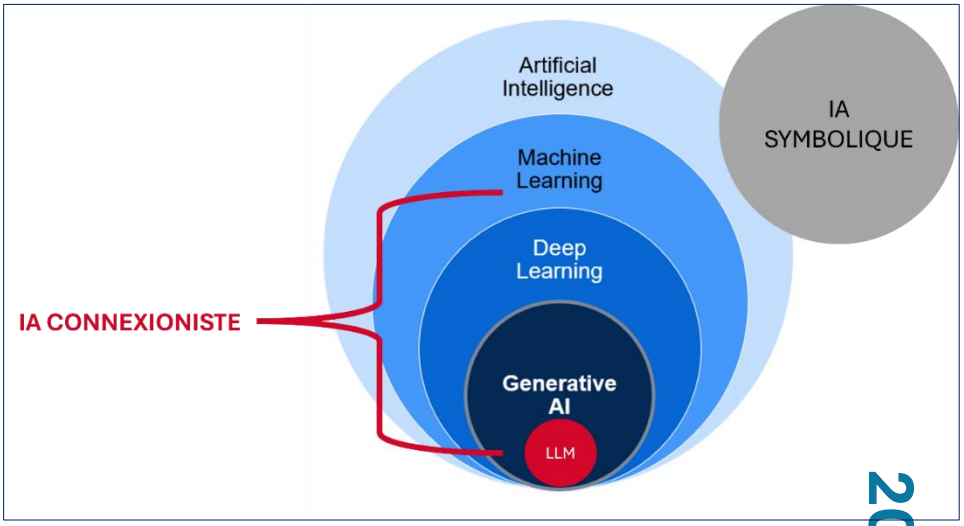
L'IA connexionniste, en revanche, s'inspire du fonctionnement du cerveau humain. Cette approche repose sur les réseaux de neurones artificiels, capables d'apprendre à partir de données. C'est la base des

technologies modernes comme le Deep Learning, qui sous-tend les modèles de langage de grande taille (LLM) tels que GPT.

- **Exemple IA connexionniste :** *Les LLM comme Chat GPT. Ils apprennent à partir d'énormes quantités de données et utilisent ces connaissances pour générer des réponses, résoudre des problèmes ou effectuer des tâches complexes comme la reconnaissance d'images ou le traitement du langage naturel.*

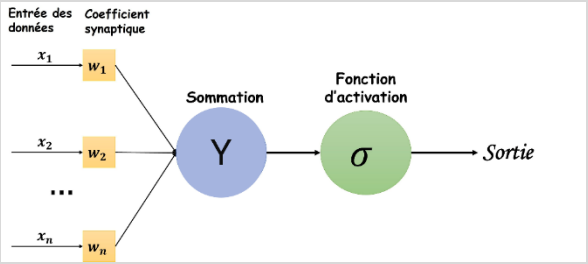
| CARACTÉRISTIQUE | IA SYMBOLIQUE | IA CONNEXIONNISTE |
|---|---|---|
| **Modèle de raisonnement** | Des règles explicites et de la logique | Des réseaux neuronaux et les données |
| **Représentation de la connaissance** | Utilise des symboles et des règles prédéfinies | Apprend des patterns à partir des données |
| **Apprentissage** | Aucun apprentissage automatique, dépend de la programmation manuelle | Apprentissage basé sur les données (supervisé ou non) |
| **Interprétabilité** | Très explicable (règles claires) | Difficile à interpréter (boîte noire) |
| **Adaptabilité** | Moins flexible dans les situations nouvelles | Très adaptable aux nouveaux contextes |
| **Domaine d'application** | Problèmes bien définis (logique, mathématiques) | Tâches complexes comme la reconnaissance d'images, le langage naturel |

## Classification des intelligences artificielles



## DEFINITION D'UN NEURONE ARTIFICIEL

Un neurone artificiel est une unité de calcul mathématique qui imite, de manière simplifiée, le comportement d'un neurone biologique. Il est conçu pour traiter des informations en appliquant une fonction mathématique à plusieurs entrées. Ces entrées sont pondérées, puis additionnées, et le résultat est passé à travers une fonction d'activation pour produire une sortie.




**Y** = (X1 * W1) + (X2*W2)

**SORTIE (z)** = F(**Y**)

Soit **SORTIE (z)** = 0 OU 1

**Voir annexe : « Comprendre la formule d'un neurone artificielle »**

## 3. L'IMPACT DE L'IA DANS LE SECTEUR DE LA FORMATION

L'IA transforme profondément le secteur de la formation, ouvrant la voie à de nouvelles méthodes d'apprentissage et à une personnalisation accrue. Grâce aux modèles de langage de grande taille (LLM) et aux technologies connexes, les organismes de formation peuvent désormais proposer des expériences plus immersives, interactives et adaptées aux besoins spécifiques des apprenants. Quelques exemples ci-dessous :

- **Personnalisation et soutien dans l'apprentissage :** L'IA permet de créer des parcours d'apprentissage sur mesure, adaptés aux compétences et aux besoins de chaque apprenant. Des systèmes basés sur l'IA peuvent analyser les retours des apprenants et ajuster les contenus pédagogiques en fonction de leur progression. L'IA peut identifier des potentiels risques

d’échec et soutenir les apprenants en difficulté.

- **Automatisation et productivité :** L'IA aide les formateurs à automatiser certaines tâches administratives, comme la correction des évaluations, la gestion des inscriptions, le suivi des apprenants. Elle peut également faciliter les tâches des autres métiers du secteur de la formation : chargé de formation, administration, ingénieur pédagogique, qualité.

- **Assistants IA au cœur du collectif :** L'intégration d'un modèle de langage (LLM) dans des plateformes collaboratives comme **Discord** ou **Teams** peut renforcer l’apprentissage collectif. Ces plateformes permettent aux apprenants de poser des questions, d'échanger des idées et de travailler en groupe, tout en bénéficiant de l’assistance continue d’un LLM. Ce dernier

peut répondre aux questions en temps réel, organiser les discussions, fournir des résumés ou des ressources, et encourager la collaboration.

Toutefois, l'utilisation de l'IA nécessite la collecte de vastes quantités de données personnelles, soulevant des préoccupations en matière de confidentialité et de sécurité. Il est essentiel de garantir que les informations des apprenants sont protégées conformément aux réglementations en vigueur. Les acteurs de la formation doivent respecter les normes RGPD et recueillir le consentement des apprenants ou des responsables légaux avant tout traitement de données personnelles.

**Le consentement** est une des bases légales prévues par le **RGPD** sur laquelle peut se fonder un traitement de données personnelles. Le **RGPD** impose que ce consentement soit libre,

spécifique, éclairé et univoque. Les conditions applicables au consentement sont définies aux articles 4 et 7 du RGPD.

Pour plus d'informations, vous pouvez vous référer aux recommandations de la CNIL : https://www.cnil.fr/fr/assurer-que-le-traitement-est-licite
https://www.cnil.fr/fr/les-fiches-pratiques-ia

L'audit par un **Délégué à la Protection des Données** (DPO) également être un moyen de prévenir toutes situations à risque.

**Enfin, l'IA Act, une nouvelle réglementation ayant un impact significatif sur les acteurs de la formation, est entrée en vigueur :**

Le règlement européen sur l'intelligence artificielle (IA) est entré en vigueur le 1er août 2024. Ce texte est nommé l'IA Act et vise à encadrer l'utilisation des technologies d'intelligence artificielle (IA) pour garantir leur utilisation éthique, sûre et transparente. Il repose sur une approche basée sur les risques, établissant quatre catégories de risques : les systèmes interdits, à haut risque, à risque limité et à risque minimal.

Les acteurs de la formation devront se conformer à cette réglementation : l'usage de l'IA concernant le domaine de la formation est considéré à haut risque (cf. **Chapitre 6 - Réglementation de l'IA ACT : IA à hauts risques**)

Enfin, l'adoption de l'IA dans le secteur de la formation s'accompagne souvent d'une dépendance significative aux technologies développées par des entreprises étrangères. Cette dépendance soulève plusieurs enjeux :

- **Souveraineté Numérique :** L'utilisation de plateformes et d'outils étrangers peut compromettre la capacité d'un pays à contrôler ses propres infrastructures éducatives et numériques. La souveraineté numérique est cruciale pour protéger les intérêts nationaux et maintenir une indépendance technologique.

- **Protection des Données :** Les technologies étrangères peuvent impliquer le transfert de données sensibles hors du pays, ce qui pose des risques en matière de confidentialité et de conformité aux réglementations sur la protection des données, comme le RGPD en Europe.

- **Sécurité et Fiabilité :** La dépendance à des fournisseurs étrangers peut exposer les institutions à des vulnérabilités, notamment en cas de tensions géopolitiques, de sanctions internationales ou de modifications unilatérales des conditions de service.

- **Adaptation Culturelle et Linguistique :** Les technologies développées à l'étranger peuvent ne pas être parfaitement adaptées aux contextes culturels, linguistiques ou pédagogiques locaux, limitant ainsi leur efficacité et leur pertinence pour les apprenants. Ces technologies peuvent amener des biais liés à d'autres contextes culturels.

# CHAPITRE 2 : LES MODELES DE LANGAGE (LLM)

## 4. QU'EST-CE QU'UN MODELE DE LANGAGE DE GRANDE TAILLE (LLM) ?

Les Large Language Models (LLM) sont des IA conçues pour comprendre et générer du texte en langage naturel. Ils sont pré-entraînés sur des vastes corpus de données textuelles afin d'apprendre les structures, les contextes et les relations entre les mots. Un modèle comme GPT (**Generative Pre-trained Transformer**) est un exemple typique d'un LLM. Ce modèle est utilisé dans l'application Chat GPT.

Un LLM est capable de réaliser des tâches telles que la génération de texte, la traduction, le résumé, la réponse à des questions, la classification de textes. **Les LLM font partie de l'IA connexionniste.**

## 5. LE FONCTIONNEMENT DES LLM

Le fonctionnement des LLM repose sur l'architecture des **Transformers**, qui est actuellement l'architecture de référence dans le traitement du langage naturel. Voici les étapes de traitement dans un LLM:

### 1. Décomposition en tokens

Lorsqu'on fournit un texte au LLM, le texte est décomposé en tokens (unités de texte), qui sont ensuite transformés en représentations numériques, appelées embeddings, permettant au modèle de traiter le langage en vecteurs de valeurs numériques.

### 2. Analyse du contexte

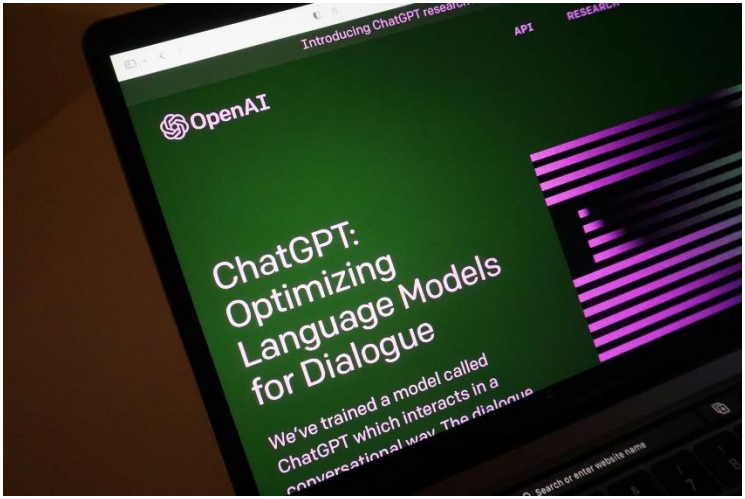
Le LLM analyse les tokens et leur ordre pour comprendre le sens de la phrase. Grâce à son entraînement sur de grandes quantités de texte, il sait comment les mots se relient entre eux.

### 3. Prédiction des tokens suivants

Pour répondre, il prédit quels tokens sont les plus probables après ceux déjà analysés. Il utilise cette prédiction pour construire une réponse cohérente.

### 4. Construction de la réponse

En assemblant les tokens prédits, le LLM génère une réponse complète et pertinente, qui respecte le sens et la structure du langage.



## 6. LE PRINCIPE DE « TOKENISATION »

**TRANSFORMATION D'UNE PHRASE EN TOKENS** :

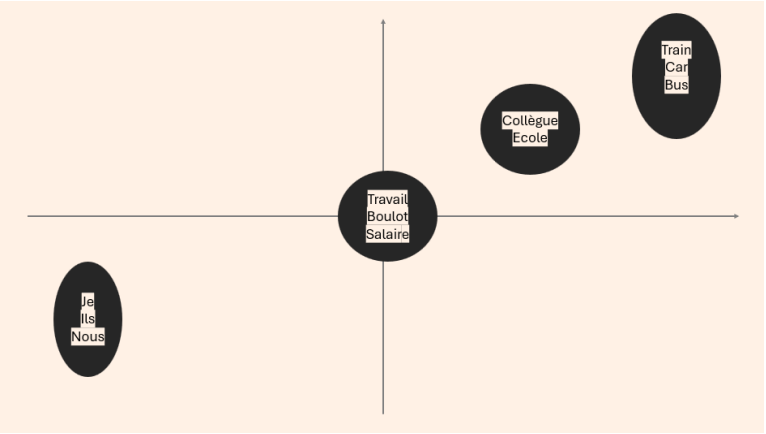
« Je vais au travail en train »

| JE | VAIS | AU | TRAVAIL | EN | TRAIN |
|---|---|---|---|---|---|

Ensuite, ces tokens sont convertis en représentation numérique, c'est l'**Embedding**. Il existe différents modèles d'Embedding, adaptés à différents corpus de textes.

| JE | VAIS | AU | TRAVAIL | EN | TRAIN |
|---|---|---|---|---|---|
| [0.2, 0.1] | [0.3, 0.2] | [0.1, 0.3] | [0.7, 0.8] | [0.1, 0.1] | [0.5, 0.8, 0.6] |

L’**Embedding** corresponds à des coordonnées dans un espace vectoriel. La distance entre les **embeddings** peut indiquer **la similarité sémantique entre les mots**.



## 7. TOUR D'HORIZON DES LLM

Il existe différents modèles de langage (LLM), ils se distinguent notamment par leur nombre de paramètres, un élément clé qui influence directement leurs performances et leurs besoins en ressources matérielles. En effet, plus un LLM possède de paramètres, plus il nécessite de puissance de calcul et de mémoire pour fonctionner. Par exemple, Llama 3.2, dans sa version à 70 milliards de paramètres, exige un ordinateur particulièrement puissant pour être exécuté localement. Pour les environnements aux ressources limitées, des versions allégées comme celle à 7 milliards de paramètres offrent une alternative moins gourmande.

## 8. LES PARAMETRES : LEUR ROLE ET LEUR IMPACT

Les paramètres d'un modèle de langage représentent les "connexions" entre les différents neurones artificiels du réseau. Ils déterminent la capacité du modèle à comprendre des relations complexes dans les données et à générer des réponses précises ou créatives. En termes simples, plus un modèle dispose de paramètres, plus il peut :

- **Capturer des subtilités dans les données :** les modèles avec un grand nombre de paramètres sont capables de comprendre des contextes complexes, des relations entre concepts, ou des nuances de langage.
- **Traiter des tâches variées :** un nombre élevé de paramètres permet au modèle d'être polyvalent et performant dans différents types de tâches (génération de texte, traduction, classification, etc.).

## 9. EFFETS DU NOMBRE DE PARAMETRES SUR LES PERFORMANCES

**Modèles avec plus de paramètres :**

**Avantages :**

- Compréhension plus fine des données complexes.
- Meilleure précision et cohérence dans les réponses.
- Aptitude à gérer des tâches plus difficiles ou à générer des textes plus longs et cohérents.

**Inconvénients :**

- Consommation accrue en ressources matérielles (CPU, GPU, RAM).
- Déploiement plus coûteux et souvent limité aux infrastructures cloud.

**Modèles avec moins de paramètres :**

**Avantages :**

- Exécution possible sur des matériels standards ou modestes (PC, VPS).
- Meilleure accessibilité pour des usages locaux ou spécifiques.

**Inconvénients :**

- Moins précis pour des tâches complexes.
- Moins performant dans la gestion de contextes longs ou subtils.

## 10. Aperçu des principaux LLM

1. **GPT-4o** (OpenAI), sorti en 2024, il est conçu pour être plus rapide et plus performant que GPT-4, tout en offrant une capacité multimodale. Cela signifie que GPT-4o peut traiter non seulement du texte, mais aussi des images et de l'audio.

2. **Gemini 1.5** (Google DeepMind) excelle dans la gestion multimodale, combinant texte, images et son.

3. **Claude 3** (Anthropic) met l'accent sur la sécurité et l'éthique.

4. **Mistral 7B** (Mistral AI) propose une efficacité optimisée avec un modèle léger et performant. Il est open-source, c'est un modèle gratuit.

5. **LLaMA 3.2 (**Meta AI) se démarque par son accessibilité open-source, favorisant l'innovation. C'est un modèle gratuit.

6. **Phi-3** (Microsoft) est un Small Language Model (SML), conçue pour être léger et performant tout en pouvant fonctionner sur des appareils

locaux de faible puissance. C’est un modèle gratuit.

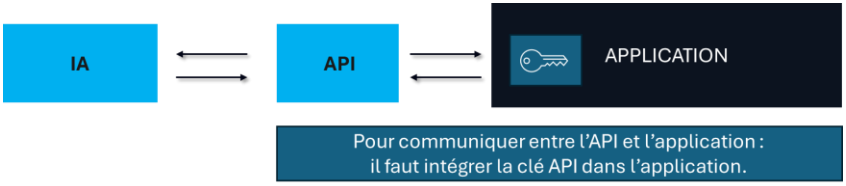
> Les SLM (Small Language Models) sont des modèles de langage de taille réduite par rapport aux LLM (Large Language Models), conçus pour offrir des performances comparables à celles des grands modèles tout en étant plus efficaces en termes de ressources. Ces modèles ont généralement un nombre réduit de paramètres, ce qui les rend plus légers et capables de fonctionner sur des dispositifs à faible puissance, comme des ordinateurs personnels ou des serveurs locaux.

Tous ces modèles supportent des fonctionnalités de **RAG** et de **Function Calling** (**Chapitre 4 et 5**), bien que certaines implémentations soient plus avancées, comme dans le cas de GPT-4o et Claude 3, qui mettent fortement l'accent sur l'intégration de ces technologies pour améliorer les réponses et les interactions.

## CHAPITRE 3 : MAITRISER L'INTEGRATION DES LLM

Pour exploiter pleinement le potentiel des LLM, il est crucial de maîtriser leur intégration dans des systèmes et des applications, que ce soit via des API payantes ou des solutions développées en interne pour réduire les coûts. Il est également important d'utiliser des modèles de prompts (prompt templates) pour orienter la génération de texte et de faire des choix stratégiques entre l'utilisation de LLM propriétaires et de modèles open-source.

**Communication entre une API et une application :**



## 11. INTEGRER DES LLM AUX APPLICATIONS GRACE AUX API PAYANTES

Les **API** (Application Programming Interfaces) permettent aux applications d'interagir avec des LLM hébergés sur des serveurs distants. L'intégration d'un LLM via une API payante est relativement simple et permet d'accéder à des modèles puissants sans avoir à gérer directement l'infrastructure complexe (serveurs distants).

**Étapes pour intégrer un LLM via une API :**

- **Choisir un fournisseur de LLM :** Des entreprises comme OpenAI, Azure AI ou Mistral fournissent des API permettant d'accéder à leurs LLM. Par exemple, OpenAI propose des API pour GPT-4.
- **Obtenir une clé API :** La clé API permet de connecter votre application aux serveurs de votre fournisseur LLM. Pour l'obtenir, il vous suffit généralement de la générer en cliquant sur un bouton dans l'interface du fournisseur. Une

fois générée, vous devez copier la clé pour l'intégrer dans votre application.

- **Intégrer une clé API :** La clé API doit être insérée dans votre application pour pouvoir envoyer des requêtes à l'API du LLM. Si vous utilisez un outil no-code, un espace dédié est prévu pour coller votre clé.

Dans le cadre de l'intelligence artificielle (IA), une API permet à une application, comme un site web ou un logiciel de formation, d'accéder à un service d'IA sans devoir tout développer en interne. Par exemple, un site éducatif peut utiliser une API pour intégrer l'IA de OpenAi (L'entreprise qui a créé ChatGPT), qui pourra répondre aux questions des utilisateurs de manière naturelle. L'API se charge de transmettre la question de l'utilisateur au système d'IA (API de OpenAI), puis de ramener la réponse en quelques secondes sur le site éducatif dans un Chatbot.

## 12. LES TARIFICATIONS DES FOURNISSEURS D'API PAYANTES : LE TOKEN

Pour utiliser une API d'intelligence artificielle, comme celles fournies par des entreprises de technologie, le modèle de tarification est souvent basé sur un concept appelé le "token". Mais qu'est-ce qu'un token ? Imaginez qu'un token soit une petite unité de texte, comme un mot ou un groupe de lettres. L'IA utilise ces tokens pour comprendre, analyser et répondre aux demandes de l'utilisateur.

Chaque fois que vous utilisez une API IA pour poser une question ou générer du texte, le système compte les tokens pour déterminer le coût. Par exemple, si vous posez une question longue ou demandez une réponse détaillée, cela nécessite plus de tokens, donc le coût est plus élevé. C'est un peu comme si vous payiez à la lettre ou au mot : plus vous en demandez, plus la "note" grimpe. Vous êtes facturés à la fois sur le nombre de tokens dans votre question, mais également sur le nombre en réponse.

**ATTENTION A LA FACTURATION AU TOKEN : DES COUTS IMPREVISIBLES**

La facturation au tokens peut vite devenir coûteuse et imprévisible. Chaque demande à l'IA consomme un nombre de tokens, et si les réponses sont longues ou détaillées, les coûts montent rapidement. Cela peut entraîner des surprises sur la facture, surtout si l'utilisation de l'IA est intensive ou que les questions sont variées. Pour éviter les coûts inattendus, il est conseillé de surveiller de près la consommation de tokens et de fixer des limites d'utilisation.

## 13. CREER SA PROPRE API LLM POUR MINIMISER LES COUTS

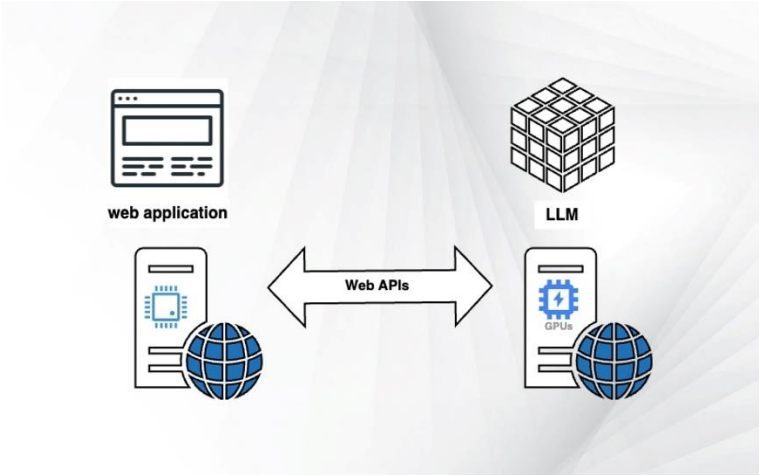
Si vous avez des besoins fréquents ou volumineux en IA, payer pour chaque token peut rapidement devenir cher. Une solution consiste alors à créer sa propre API IA en hébergeant un modèle de langage (LLM) sur vos propres serveurs. Cela signifie que, au lieu de dépendre d'un fournisseur d'API pour chaque demande (et de payer à chaque fois), vous installez et gérez votre propre IA.

Créer une API IA interne est très complexe, cela demande une certaine expertise technique, car vous aurez besoin de serveurs pour faire tourner le modèle d'IA et de personnes pour paramétrer ce serveur et optimiser le modèle pour votre usage.

L'avantage est que vous contrôlez vos coûts. Au lieu de payer à chaque utilisation, vous ne payez que pour les coûts d'hébergement (comme l'électricité et la maintenance des serveurs). En plus, vous pouvez ajuster l'IA pour qu'elle consomme moins de

ressources et réponde plus rapidement, ce qui réduit encore les dépenses.

Ainsi, en fonction de la quantité d'IA dont vous avez besoin, créer une API interne peut être bien plus rentable sur le long terme, surtout si vos utilisateurs sont nombreux ou que vous avez besoin de réponses détaillées.



## 14. LE PROMPT TEMPLATE : GUIDER LE LLM DANS SON FONCTIONNEMENT

Un prompt template est une structure préconfigurée ou un modèle qui guide l'interaction entre le LLM et l'utilisateur. Il inclut plusieurs éléments qui permettent de définir précisément comment le LLM doit répondre et quelles fonctionnalités ou comportements sont attendus. Voici un modèle de prompt template pour permettre à une IA de réagir selon des critères précis :

**1.Rôle :** « Tu joues le rôle de [nom du personnage ou fonction], avec une expertise en [domaine]. »

**2.Contexte :** « Tu te trouves dans le contexte de [description de la situation ou du cadre], interagissant avec [public cible]. Ce contexte est important pour orienter la pertinence de tes réponses. »

**3.Objectifs :** « Ton objectif principal est de [objectif principal, par ex. : fournir des réponses précises et pédagogiques, résoudre un problème, etc.]. Assure-toi que [critères de performance, par ex. : clarté, précision, interaction dynamique, etc.] sont respectés. »

**4.Fonctionnalités :** « Utilise tes compétences pour [fonctions attendues de l'IA, par ex. : répondre aux questions, générer des scénarios, simuler des interactions, etc.]. Appuie-toi sur [types de ressources ou informations à utiliser]. Tu dois aussi [instructions supplémentaires pour maximiser les capacités]. »

**5.Exemples :** « Voici quelques exemples de ce à quoi doivent ressembler tes réponses : [exemples précis]. »

## 15. LLM PROPRIETAIRES ET LLM OPEN-SOURCE

La décision d'utiliser un modèle de langage propriétaire ou open-source est une considération essentielle lors de l'intégration des LLM dans une application.

### LLM PROPRIETAIRES

Les modèles propriétaires comme GPT-4, Claude ou Gemini sont proposés par des entreprises avec des API bien développées et un support technique important. Ces modèles sont souvent formés sur de vastes ensembles de données, ce qui leur permet d'offrir des réponses de haute qualité sur une large gamme de sujets. Voici quelques points à considérer :

- **Performance élevée :** Les modèles propriétaires sont généralement bien optimisés et bénéficient de mises à jour régulières pour améliorer leur performance.
- **Facilité d'intégration :** Grâce aux API robustes, ils s'intègrent facilement dans des

systèmes ou des applications déjà en place sans nécessiter une infrastructure dédiée.

- **Coût :** L'un des principaux inconvénients des modèles propriétaires est leur coût. L'accès aux API peut rapidement devenir coûteux à grande échelle, surtout si vous traitez des milliers de requêtes par jour.
- **Confidentialité :** L'utilisation de modèles propriétaires signifie que les données que vous envoyez aux serveurs de ces entreprises peuvent être traitées et potentiellement stockées. Pour des entreprises avec des besoins stricts en matière de confidentialité, cela peut poser un problème.
- **Personnalisation limitée :** Vous êtes souvent limité à l'utilisation du modèle tel quel, avec peu de contrôle sur la façon dont il a été entraîné ou les données qu'il a utilisées.

Pour utiliser ces modèles, vous devez utiliser les API fournis par les constructeurs comme OpenAI ou Google.[1]

## LLM OPEN-SOURCE

Les modèles open-source comme Llama 3, Mistral ou Gemma offrent une flexibilité accrue, bien que l'effort d'intégration soit plus conséquent. En effet, c'est l'utilisateur qui doit intégrer lui-même ces modèles dans son application directement ou via la construction d'une API :

- **Gratuité :** Les modèles open-source sont généralement gratuits à utiliser et peuvent être déployés sur votre propre infrastructure, ce qui permet de réduire les coûts à long terme.
- **Contrôle total :** Avec un modèle open-source, vous avez un contrôle total sur la manière dont



[1] Pour un exemple détaillé, reportez-vous au **Chapitre 3**, section 8 *« OBTENIR UNE CLE API PAR OPENAI : FOURNISSEUR DES LLM DE LA FAMILLE GPT. »*.

le modèle est déployé et utilisé. Vous pouvez l'entraîner sur vos propres données, l'adapter à vos besoins spécifiques, et même l'optimiser pour certaines tâches.

- **Confidentialité :** Étant donné que vous pouvez héberger le modèle en interne, vous avez un contrôle total sur les données qui sont utilisées et traitées, ce qui est crucial pour les secteurs ayant des exigences strictes en matière de conformité et de confidentialité.

- **Infrastructure nécessaire :** L'inconvénient majeur est la complexité technique associée à l'hébergement et à l'optimisation de ces modèles. Ils nécessitent souvent des ressources informatiques puissantes et une expertise technique pour leur gestion.

- **Personnalisation :** Vous pouvez adapter les modèles open-source en fonction de vos besoins spécifiques, ce qui est un atout considérable. Vous pouvez par exemple ajuster les paramètres d'entraînement ou intégrer des ensembles de données personnalisés pour obtenir des résultats optimisés pour votre cas d'usage.

Pour utiliser des LLM open-source en local (directement sur votre ordinateur), vous pouvez utiliser Ollama (cf. Un guide pour utiliser Ollama est disponible en annexe de ce livre).

# CHAPITRE 4 : LE RAG - ENRICHIR LES LLM DE VOS DONNEES

## 16. DECOUVRIR LE RAG ET SES BENEFICES

Le **Retrieval Augmented Generation** (RAG) est une approche innovante qui combine les modèles de langage de grande taille (LLM) avec des techniques avancées de récupération d'informations. En intégrant des données externes pertinentes lors de la génération de réponses, le RAG permet de surmonter les limitations des LLM traditionnels, tels que la connaissance statique ou les informations obsolètes.

### Qu'est-ce que le RAG ?

Le RAG est une architecture hybride qui, face à une requête, utilise un module de recherche pour extraire des informations pertinentes à partir d'une base de données ou d'un corpus spécifique. Ces informations sont ensuite combinées avec les capacités génératives du LLM pour produire des réponses plus précises, contextualisées et riches en contenu.

**Les bénéfices du RAG**

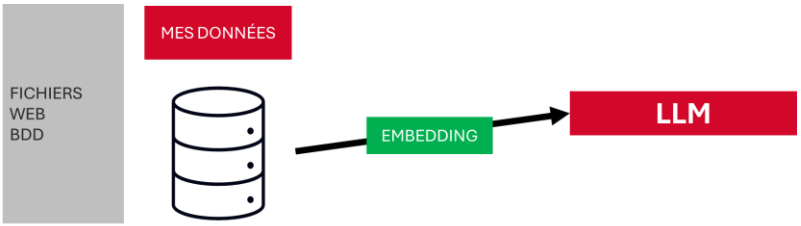
- **Actualisation des connaissances :** Accès à des informations récentes, dépassant les limitations de la date de coupure des LLM.
- **Personnalisation accrue :** Adaptation des réponses en fonction de sources de données spécifiques au domaine ou à l'utilisateur.
- **Réduction des erreurs factuelles :** Diminution des "hallucinations" des LLM en s'appuyant sur des données vérifiées.
- **Explicabilité et traçabilité :** Possibilité de fournir des sources ou des références, renforçant la confiance des utilisateurs.
- **Efficacité opérationnelle :** Pas besoin de réentraîner entièrement le modèle pour intégrer de nouvelles informations.

## 17. AMPLIFIER LES CAPACITES DES LLM GRACE AU RAG

En intégrant le RAG, les LLM peuvent offrir des réponses non seulement basées sur leur entraînement initial, mais également enrichies par des données spécifiques, récentes et pertinentes pour l'utilisateur issus de base de données, de corpus de textes, de divers fichiers.

**Comment le RAG enrichit les LLM**

- **Contextualisation avancée :** Les réponses sont adaptées au contexte actuel ou au domaine spécifique de l'utilisateur grâce aux bases de données ou aux fichiers injectées.
- **Flexibilité :** Capacité à intégrer rapidement de nouvelles informations sans nécessiter de longues phases d'entraînement.
- **Personnalisation :** Possibilité d'adapter les réponses en fonction des besoins individuels des apprenants ou des objectifs pédagogiques des formateurs.




**Mise en œuvre du RAG**

- **Construction d'une base de connaissances :** Compilation de documents pertinents pour le domaine éducatif concerné.
- **Intégration technique :** Mise en place d'un système efficace de récupération d'informations compatible avec le LLM utilisé.
- **Optimisation continue :** Ajustement des paramètres pour améliorer la pertinence et la qualité des réponses fournies.

## 18. CAS D'USAGE POUR LES ACTEURS DE LA FORMATION

L'application du RAG dans le secteur de la formation offre des opportunités significatives pour améliorer l'expérience d'apprentissage des apprenants et faciliter le travail des formateurs.

### 1. Personnalisation de l'apprentissage avec les données de l'apprenant

**Pour les apprenants :**

- **Parcours d'apprentissage sur mesure :** L'IA utilise le RAG pour accéder aux données spécifiques de l'apprenant (résultats antérieurs, styles d'apprentissage, objectifs personnels) afin de personnaliser les contenus pédagogiques.
- **Feedback ciblé :** Fourniture de commentaires précis basés sur les travaux et les performances passées de l'apprenant, avec des

suggestions pour améliorer les domaines spécifiques.

**Exemple :** Un étudiant en langues étrangères reçoit des exercices de grammaire personnalisés sur les points où il a montré des difficultés, grâce à l'analyse de ses précédents tests et devoirs.

**Pour les formateurs :**

- **Suivi individualisé :** Le RAG permet aux formateurs d'accéder rapidement aux progrès et aux défis de chaque apprenant, en utilisant les données collectées pour adapter leur enseignement.
- **Adaptation des stratégies pédagogiques :** En comprenant les besoins individuels, les formateurs peuvent ajuster leurs méthodes pour soutenir efficacement chaque étudiant.

**2. Intégration des contenus propriétaires de l'organisme de formation**

**Pour les apprenants :**

- **Accès à des ressources exclusives :** L'IA récupère des documents internes, des études de cas et des recherches propres à l'organisme, enrichissant l'apprentissage avec du contenu que les apprenants ne trouveraient pas ailleurs.

**Exemple :** Dans une école de commerce, les étudiants utilisent une IA qui intègre les dernières analyses réalisées par l'organisme de formation pour étudier les tendances économiques actuelles.

**Pour les formateurs :**

- **Mise à jour dynamique des supports :** Les formateurs peuvent incorporer facilement les nouvelles informations internes dans leurs cours, en s'assurant que le contenu est toujours à jour.
- **Conformité aux normes internes :** Le RAG garantit que les informations diffusées

respectent les politiques et les standards de l'organisme.

**3. Utilisation des connaissances spécifiques du formateur**

**Pour les apprenants :**

- **Accès aux expertises du formateur :** L'IA utilise le RAG pour intégrer les publications, les recherches et les notes du formateur, offrant un apprentissage enrichi par l'expertise spécifique de l'enseignant.

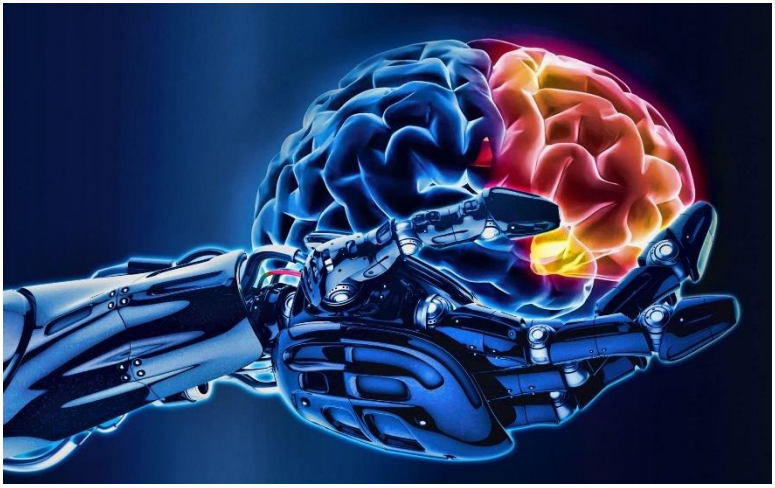
**Exemple :** Un professeur d'histoire de l'art partage ses recherches inédites via l'IA, permettant aux étudiants d'accéder à des analyses approfondies non disponibles dans les manuels standards.

**Pour les formateurs :**

- Diffusion efficace des connaissances : Les formateurs peuvent partager leurs travaux et insights personnels à grande échelle, sans

multiplier les efforts.

- Interaction enrichie avec les apprenants : L'IA peut servir de relais pour répondre aux questions des étudiants en s'appuyant sur les connaissances spécifiques du formateur.



**4. Support administratif et orientation personnalisée**

**Pour les apprenants :**

- **Assistance personnalisée :** L'IA peut répondre aux questions administratives ou d'orientation en se basant sur le dossier spécifique de l'étudiant et les procédures internes de l'organisme.

**Exemple :** Un étudiant souhaite changer de filière. L'IA, en accédant à son dossier académique et aux politiques de l'établissement, lui fournit un guide personnalisé sur les étapes à suivre.

**Pour les formateurs et l'organisme :**

- **Optimisation du temps :** En déléguant les questions récurrentes à l'IA, le personnel peut se concentrer sur des tâches à plus forte valeur ajoutée.
- **Cohérence des informations :** L'IA fournit des réponses standardisées et conformes aux

politiques internes, réduisant les erreurs et les malentendus.

**5. Formation professionnelle continue avec des données d'entreprise**

**Pour les apprenants (professionnels) :**

- **Mise à jour des compétences internes :** Les employés accèdent à des formations intégrant les données et les procédures spécifiques de leur entreprise.

**Exemple :** Un technicien reçoit une formation sur une nouvelle machine installée dans son entreprise, avec des instructions et des protocoles spécifiques récupérés via le RAG.

**Pour les formateurs et l'entreprise :**

- **Déploiement rapide des formations :** Intégration des nouveaux processus ou outils dans les modules de formation sans délai.

- **Adaptation aux besoins de l'entreprise :** Les formations sont alignées sur les objectifs stratégiques et les particularités de l'entreprise.

**Avantages pour les organismes de formation**

- **Valorisation des ressources internes :** Le RAG permet d'exploiter pleinement les documents, les connaissances et les données internes de l'organisme.
- **Expérience apprenant améliorée :** En offrant des contenus personnalisés et pertinents, les apprenants sont plus engagés et satisfaits.
- **Efficacité opérationnelle :** L'automatisation de la récupération et de la diffusion des informations spécifiques réduit les charges administratives et optimise les processus.

L'utilisation du RAG pour injecter des données spécifiques transforme la façon dont les LLM peuvent être utilisés dans le domaine de la formation. En intégrant les données propres aux apprenants, aux formateurs et aux organismes, il est possible de créer des expériences d'apprentissage hautement personnalisées, pertinentes et efficaces. Cette approche pragmatique du RAG offre des avantages tangibles pour tous les acteurs de la formation, en alignant les technologies d'IA sur les besoins réels du terrain.

## EXEMPLE UTILISATION D'UNE IA RAG

**Contexte :**

Dans une institution de formation en santé, les protocoles médicaux changent fréquemment. Un formateur dispense un cours sur les procédures d'hygiène hospitalière.

**Mise à jour dynamique des supports :**

Grâce au RAG, le formateur intègre les nouvelles directives sanitaires dans l'IA. L'IA récupère les informations pertinentes des documents officiels et internes. Elle peut ainsi aider le formateur à générer du contenu pédagogique adapté et actualisé.

**Résultat :**

**Pour le formateur :** Supports actualisés sans effort manuel, conformes aux normes.

**Pour les apprenants :** Accès à des informations récentes, formation alignée sur les pratiques actuelles du secteur.

# CHAPITRE 5 : FUNCTION CALLING - CONNECTER LES LLM A VOS SERVICES TIERS

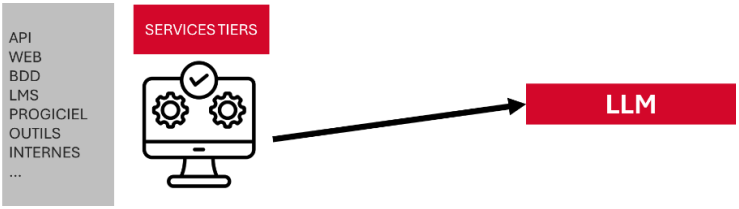
## 19. INTRODUCTION AU FUNCTION CALLING

Les modèles de langage comme ChatGPT peuvent déjà répondre à des questions, mais imaginez s'ils pouvaient agir pour vous : chercher des infos dans vos systèmes, planifier des tâches ou interagir avec vos outils. C’est ce que permet le Function Calling. Il connecte les LLM à des services externes pour exécuter des actions, comme accéder à vos plateformes de formation (LMS), automatiser des tâches ou générer des rapports, rendant l'IA encore plus utile et proactive dans votre quotidien.

## 20. COMMENT LES LLM PEUVENT APPELER DES SERVICES EXTERNES

Pour que cela fonctionne, voici les grandes étapes :

- **Définir les actions possibles :** On indique au LLM quelles tâches il peut effectuer, par exemple "récupérer les notes" ou "programmer une réunion".
- **Faire le lien avec vos outils :** Le LLM va communiquer avec vos systèmes, comme votre plateforme d'apprentissage, en utilisant les API liées à vos services ou systèmes tiers.
- **Utiliser les résultats dans les réponses :** Une fois qu'il a obtenu les informations ou exécuté l'action, le LLM les utilise pour répondre de manière plus pertinente et personnalisée.




**Exemple concret de Function Calling**

Imaginons que vous demandiez à une IA :

- Vous : "Où en est l'apprenant Pierre dans le module de gestion de projet ?"
- IA : [Appelle la plateforme LMS pour récupérer les informations]
- IA : "Pierre a terminé 80 % du module avec une moyenne de 88 % aux évaluations."

Le Function Calling permet à l'IA de chercher directement dans vos outils ces informations sans que vous ayez besoin de le faire vous-même.

## 21. CAS D'USAGE POUR LES ACTEURS DE LA FORMATION

### 1. Suivi personnalisé des apprenants

Le LLM peut accéder directement aux données de votre LMS pour suivre le parcours de chaque apprenant. Cela signifie que vous pouvez demander à l'IA de :

- Récupérer les résultats d'évaluations.
- Suivre la progression des compétences.
- Fournir des recommandations personnalisées aux apprenants en fonction de leurs performances.

**Exemple :** « Lance un rapport sur les progrès de Julie dans le cours de gestion de projet. » Le LLM va interroger le système, puis fournir les résultats et suggestions.

### 2. Automatisation des tâches administratives

Le Function Calling peut automatiser certaines tâches chronophages comme :

- Planifier des cours ou des réunions avec les apprenants.
- Envoyer des notifications automatiques avant les échéances.
- Générer des rapports sur les performances des apprenants ou l'avancement des formations.

**Exemple :** « Programmer une session de révision pour le groupe A vendredi à 15h." »
L'IA peut s'occuper de tout cela, y compris envoyer les invitations.

**3. Gestion des compétences et certifications**

Le LLM peut aussi aider à gérer les compétences des apprenants en se connectant à vos systèmes de suivi. Il peut :

- Identifier les compétences acquises et celles à renforcer.
- Proposer des formations complémentaires pour obtenir une certification.

- Suivre l’évolution des compétences en fonction des évaluations.

**Exemple :** « Quelles compétences me manquent pour obtenir ma certification en cybersécurité ? » Le LLM va consulter vos outils et vous fournir une réponse claire.

Le Function Calling, c’est comme donner des super pouvoirs à l'IA que vous utilisez déjà dans la formation. Non seulement elle peut répondre à des questions, mais elle peut aussi agir directement dans vos outils pour automatiser des tâches, récupérer des informations précises et vous faire gagner du temps. Que ce soit pour suivre les progrès des apprenants, planifier des formations ou gérer les compétences, c'est un gain d'efficacité énorme pour les formateurs et les organismes de formation.

# CHAPITRE 6 : LES GRANDS ENJEUX

## 22. UN NOUVEAU STANDARD DE PRIX : LE TOKEN

L'introduction des modèles de langage de grande taille (LLM) dans les applications a transformé la manière dont nous abordons les coûts d'utilisation. Contrairement aux approches traditionnelles basées sur des licences ou des abonnements fixes, les LLM s'appuient de plus en plus sur une unité économique spécifique : le token. Un token représente une portion de texte (mot, caractère ou partie de phrase) et l'utilisation d'un LLM est généralement facturée en fonction du nombre de tokens utilisés pour traiter une requête.

Ce modèle de tarification par token s'applique lorsque l'on utilise les API de fournisseurs de LLM tels qu'OpenAI, Google, Azure AI Services, ou encore Anthropic. Ces services facturent chaque token utilisé, que ce soit en entrée (la question posée par un utilisateur) ou en sortie (la réponse générée par le modèle). Cela signifie que le coût peut varier en fonction de la longueur et de la complexité des requêtes soumises.

Pour les acteurs de la formation, cette approche granulaire des coûts peut s'avérer intéressante puisqu'elle permet une facturation plus fine basée sur l'utilisation réelle. Cependant, cela nécessite également une gestion attentive pour éviter les dérives budgétaires. Dans le cadre d'une formation où l'IA serait utilisée régulièrement, maîtriser cette tarification devient essentiel pour garantir que les coûts ne deviennent pas prohibitifs.

## 23. RETARD EUROPEEN ET DEPENDANCE TECHNOLOGIQUE

L'Europe accuse un certain retard dans le domaine de l'intelligence artificielle, en particulier face aux géants technologiques américains et chinois qui dominent le marché des LLM. Cet écart se traduit par une dépendance croissante aux technologies étrangères, qu'il s'agisse des modèles d'OpenAI, Google ou des entreprises chinoises comme Baidu. Bien que certains pays européens aient pris des initiatives pour encourager le développement de leurs propres IA, comme la France avec ses laboratoires de recherche en IA ou l'Allemagne avec ses innovations industrielles, ces efforts restent insuffisants face à la rapidité des avancées des autres régions du monde.

Cette situation pose des enjeux stratégiques pour les acteurs européens de la formation. Utiliser des technologies étrangères signifie souvent se plier à des conditions d'utilisation, des coûts et des normes qui ne correspondent pas toujours aux besoins ou aux

réalités locales. De plus, cela entraîne des préoccupations en matière de souveraineté numérique et de protection des données.

### 24. UN COUT ENVIRONNEMENTAL

Le déploiement de modèles d'IA de plus en plus puissants, comme les LLM, n'est pas sans conséquence pour l'environnement. En effet, l'entraînement et l'utilisation de ces modèles nécessitent des quantités massives de puissance de calcul, ce qui entraîne une consommation d'énergie importante. L'entraînement des modèles comme GPT-4 ou Mistral peut prendre des semaines, voire des mois, sur des milliers de serveurs fonctionnant en continu, consommant ainsi une énorme quantité d'électricité.

Les acteurs de la formation, soucieux d'adopter des pratiques durables, doivent être conscients de ces coûts environnementaux lorsqu'ils intègrent des LLM dans leurs infrastructures. Il est donc essentiel d'évaluer les impacts environnementaux de l'utilisation de

l'IA et d'envisager des solutions pour minimiser ces effets, telles que l'utilisation de modèles plus efficaces, le recours à des approches plus responsables en termes d'utilisation de ces LLM.

## 25. REGLEMENTATION DE L'IA ACT : IA A HAUTS RISQUES

L'IA Act, ou *Artificial Intelligence Act*, est un règlement de l'Union européenne visant à encadrer l'utilisation des technologies d'intelligence artificielle (IA) afin de garantir une utilisation éthique, sûre et transparente. Adopté en mai 2024 et entré en vigueur le 1er août 2024, ce règlement introduit une classification des systèmes d'IA basée sur les risques, allant des applications interdites aux applications à risque minimal. Les obligations spécifiques pour les systèmes d'IA à haut risque entreront en vigueur le 2 août 2026.

## 26. STRUCTURE DU REGLEMENT ET IMPLICATIONS

1. **Systèmes interdits** : Certaines applications de l'IA sont interdites, notamment celles qui exploitent la vulnérabilité des personnes pour manipulation, ou qui affectent les droits fondamentaux. Cela comprend, par exemple, des systèmes de surveillance massive ou de scoring social.

2. **Systèmes à haut risque** : Les systèmes d'IA déployés dans des domaines critiques, tels que la santé, la sécurité et l'éducation, sont classés comme à haut risque et doivent répondre à des exigences rigoureuses de transparence, de robustesse et de sécurité. Dans le secteur de la formation, cela pourrait inclure des outils d'évaluation ou de sélection pour des formations. Ces systèmes nécessitent des contrôles de qualité, la traçabilité des données et des garanties de transparence

algorithmique. Les fournisseurs doivent s'assurer que ces systèmes respectent ces normes avant leur déploiement.

3. **Systèmes à risque limité et minimal** : Les systèmes d'IA présentant un risque faible, comme certaines applications éducatives standards, sont soumis à des exigences de transparence réduites. Les utilisateurs doivent être informés de l'utilisation de l'IA dans ces applications.

### 27. Rôles et responsabilités

L'IA Act définit des rôles spécifiques pour chaque acteur de la chaîne de valeur de l'IA, avec des obligations distinctes, ces rôles peuvent se cumuler pour une même entité :

**Fournisseur :** une personne physique ou morale, une autorité publique, une agence ou un autre organisme qui développe ou fait développer un système d'IA ou un modèle d'IA à usage général et le met sur le marché ou met le système d'IA en service sous son propre nom

ou sa propre marque, que ce soit à titre onéreux ou gratuit ;

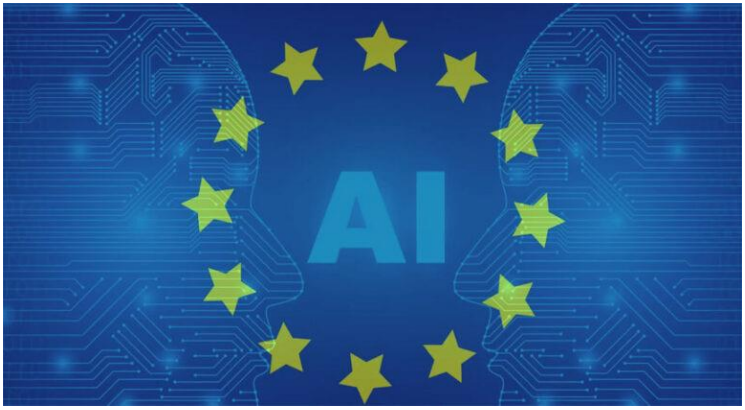
**Déployeur :** une personne physique ou morale, une autorité publique, une agence ou un autre organisme utilisant un système d'IA sous son autorité, sauf si le système d'IA est utilisé dans le cadre d'une activité personnelle non professionnelle :

**Représentant autorisé :** une personne physique ou morale située ou établie dans l'Union qui a reçu et accepté un mandat écrit d'un fournisseur de système d'IA ou de modèle d'IA à usage général pour, respectivement, exécuter et mener à bien en son nom les obligations et les procédures établies par le présent règlement ;

**Importateur :** une personne physique ou morale située ou établie dans l'Union qui met sur le marché un

système d'IA portant le nom ou la marque d'une personne physique ou morale établie dans un pays tiers ;

**Distributeur :** une personne physique ou morale de la chaîne d'approvisionnement, autre que le fournisseur ou l'importateur, qui met un système d'IA à disposition sur le marché de l'Union ;



## 28. IMPLICATIONS POUR LES ACTEURS DE LA FORMATION

Dans le secteur de la formation, les implications de l'IA Act sont profondes :

1. **Transparence** : Les institutions doivent informer les apprenants et le personnel sur la présence d'IA dans les systèmes de formation et expliquer leur finalité et leurs limitations.

2. **Sécurité des données** : Étant donné que les données personnelles et éducatives sont sensibles, des mesures strictes de protection et de gestion des données doivent être mises en place, tant par les fournisseurs que par les utilisateurs.

3. **Suivi de conformité** : Les systèmes à haut risque, tels que ceux utilisés pour évaluer les performances des apprenants, nécessitent des contrôles réguliers pour garantir une évaluation équitable et sans biais, avec des audits

de conformité et des processus de documentation rigoureux.

4. **Formation des utilisateurs** : Il est essentiel que les formateurs et les responsables pédagogiques soient formés à l'utilisation de ces technologies pour s'assurer qu'elles sont employées de manière éthique et en accord avec les objectifs pédagogiques.

L'IA Act impose donc aux acteurs de la formation un devoir de vigilance et de transparence pour s’assurer que les outils IA respectent les valeurs fondamentales du secteur.

**OBLIGATIONS POUR LES ACTEURS DE LA FORMATION :**

- **Identification des systèmes à haut risque :** Les organismes doivent recenser les outils d'IA utilisés, en particulier ceux qui influencent l'accès aux formations, l'évaluation des compétences ou la surveillance des apprenants.

- **Mise en conformité :** Il est crucial que ces systèmes respectent les normes de transparence, de robustesse et de supervision humaine fixées par le règlement.

- **Formation du personnel :** Les formateurs et responsables doivent être formés à l'usage de l'IA pour garantir un usage approprié et aligné sur les directives.

- **Protection des droits des apprenants et transparence** : L'utilisation de l'IA doit respecter les droits fondamentaux, en évitant notamment les biais et discriminations qui pourraient affecter les décisions. Les acteurs de la formation doivent indiquer comment et en quoi l'IA agit dans le processus de formation.

- **Conformité réglementaire** : Le non-respect du règlement peut entraîner des sanctions financières élevées, ce qui crée un enjeu de taille pour les acteurs de la formation.

## 29. OBLIGATIONS DE TRANSPARENCE POUR LA FORMATION

Le règlement européen sur l'intelligence artificielle, connu sous le nom d'IA Act, impose des obligations spécifiques aux acteurs de la formation en matière de transparence lors de l'utilisation de systèmes d'intelligence artificielle (IA). Ces exigences visent à garantir une utilisation éthique et responsable de l'IA dans le domaine éducatif.

**Obligations de transparence pour les acteurs de la formation :**

- **Information des apprenants** : Les organismes de formation doivent informer clairement les apprenants lorsqu'un système d'IA est utilisé dans le cadre de leur parcours éducatif. Cette information doit préciser la nature et le rôle de l'IA dans les processus d'apprentissage ou d'évaluation.

- **Explicabilité des décisions** : Les décisions prises par des systèmes d'IA, notamment en matière d'évaluation ou de recommandation de contenus, doivent être compréhensibles pour les apprenants. Les organismes doivent être en mesure d'expliquer les critères et les données utilisés par l'IA pour parvenir à ces décisions.
- **Documentation accessible** : Une documentation détaillée sur le fonctionnement des systèmes d'IA doit être mise à disposition des parties prenantes, y compris les apprenants, afin de favoriser une compréhension approfondie des mécanismes en jeu.

**Avant tout projet en IA, il est nécessaire de se faire accompagner en droit numérique par des spécialistes.**

# ANNEXES

## CHAPITRE 7 : COMPRENDRE LA FORMULE D'UN NEURONNE ARTIFICIEL

Un neurone artificiel suit une formule mathématique de type : **Sortie (z) = F(Y)**, où :

- **X1 et X2** sont les entrées du neurone, représentant les données brutes reçues par le neurone. Dans un réseau de neurones pour la classification d'images, par exemple, x1x1x1 pourrait correspondre à l'intensité d'un pixel, et x2x2x2 à celle d'un autre pixel.
- **W1 et W2** sont les poids associés aux entrées. Chaque poids détermine l'importance de l'entrée sur la sortie finale. Un poids élevé signifie que l'entrée correspondante influence davantage le résultat. Ces poids sont ajustés au cours de la phase d'apprentissage.
- **Y** est la somme pondérée des entrées avant application de la fonction d'activation.

## 30. COMMENT FONCTIONNE CETTE FORMULE ?

1. **Multiplication des entrées par leurs poids :** Le neurone prend chaque entrée (x1) et la multiplie par son poids respectif (w1). Exemple :
    - Si **x1**= 0.5 et **w1** = 0.8, alors **x1·w1** = 0.4.
    - Si **x2** = 0.3 et **w2** = 0.6, alors **x2·w2** = 0.18.
2. **Somme des produits :** Ensuite, le neurone additionne les résultats des multiplications pour obtenir la somme pondérée Y.
    - Dans notre exemple : **y** = 0.4+0.18 = 0.58.
3. **Application de la fonction d'activation :** Une fois la somme pondérée YYY calculée, le neurone passe cette valeur à travers une fonction d'activation. Cette étape est cruciale pour permettre au neurone de prendre une décision non linéaire (comme activer ou désactiver le neurone). Selon le type de fonction d'activation choisie, le neurone peut produire une sortie continue ou discrète :

- Par exemple, si on utilise une fonction seuil, le neurone renvoie z=1 (activé) si y dépasse un certain seuil, ou z=0 (désactivé) sinon.

4. **Sortie et décision finale :** Le 0 ou 1 en sortie permet de classer les données de manière linéaire. Par exemple, nous fixons un seuil à 0.5 pour notre exemple, cela signifie que la sortie sera 1. Donc le résultat est vrai. Exemple, indiquer si la photo est un chat ou non :
5. **Application concrète :** Si l'on utilise ce neurone pour déterminer si une photo représente un chat :
    - Si **y** > 0.5 alors la sortie, z = 1 -> la photo est un chat
    - Si **y** <0.5 alors la sortie, z = 0 -> la photo n'est pas un chat

**N.B :** Cet exemple avec une sortie 0 ou 1 est une version simplifiée du neurone artificiel. Dans les réseaux modernes, on utilise des fonctions d'activation plus complexes comme ReLU ou Sigmoïde, qui permettent au réseau d'apprendre des relations plus riches et de résoudre des problèmes plus avancés, comme la reconnaissance d'images. Bien que cet exemple aide à comprendre les bases, les réseaux réels utilisent ces fonctions pour traiter des données plus complexes.

## CHAPITRE 8 : UTILISATION DE OLLAMA

Ollama est un outil informatique qui permet d'utiliser des modèles de langage (LLM) open-source (gratuit), directement sur son propre ordinateur. Contrairement à d'autres solutions nécessitant une connexion internet pour accéder à des serveurs distants, Ollama fonctionne localement, offrant ainsi plus de contrôle sur les données et une meilleure confidentialité. Il est compatible avec différents systèmes d'exploitation, notamment macOS, Linux et Windows.

Vous pouvez télécharger un guide en cliquant sur ce lien.

*Le second volet « Acteurs de la formation : maîtrisez l'IA – Partie 2 : Pratiquer » sera prochainement disponible.*

PARTIE 1 : COMPRENDRE

# ACTEURS DE LA FORMATION : MAÎTRISEZ L'I.A

COMMUNOTIC
L'ère numérique de la formation

**EDITION 2024**
Service Transformation Numérique des Territoires
Direction Aménagement du Numérique

RÉGION NORMANDIE