

# Glossaire de l'IA

## 40 termes essentiels pour comprendre l'Intelligence Artificielle

De LLM à AGI — tous les concepts clés expliqués simplement.

### Fondamentaux

<b>Intelligence Artificielle (IA)</b> Techniques permettant aux machines de simuler des capacités cognitives : apprentissage, raisonnement, résolution de problèmes et génération de langage.	<b>Machine Learning (ML)</b> Les algorithmes apprennent automatiquement à partir de données, sans être explicitement programmés. Plus les données sont nombreuses, plus le modèle est performant.
<b>Deep Learning</b> Réseaux de neurones profonds (plusieurs couches). Révolutionnaire pour la vision, le langage et la génération de contenu. Base du boom IA depuis 2012.	<b>Réseau de neurones</b> Architecture inspirée du cerveau humain. Neurones artificiels connectés en couches successives capables d'apprendre des représentations très complexes.
<b>Algorithme</b> Suite d'instructions logiques pour résoudre un problème. En IA, il définit comment le modèle apprend à partir des données et fait des prédictions.	

### Grands Modèles de Langage (LLM)

<b>LLM — Large Language Model</b> Modèle entraîné sur des milliards de textes pour comprendre et générer du langage naturel. Base de Claude, ChatGPT, Gemini et Mistral.	<b>Token</b> Unité de texte traitée par un LLM (env. 3/4 de mot). Les modèles ont une fenêtre de contexte limitant le nombre de tokens traités simultanément.
<b>Transformer</b> Architecture Google 2017 ("Attention Is All You Need"). Base de tous les LLM modernes grâce au mécanisme d'attention multi-tête.	<b>Contexte (Context window)</b> Quantité max de texte traitée à la fois. Claude : 200k tokens. Gemini 2.0 : 1 million de tokens — record du marché en 2026.
<b>Hallucination</b> L'IA génère des informations fausses de manière convaincante, sans le savoir. Vérifier toujours les faits importants avant utilisation.	<b>Paramètres</b> Les "connexions" d'un modèle ajustées pendant l'entraînement. GPT-4 : ~1 700 milliards. Plus il y en a, plus le modèle est puissant et coûteux.
<b>Inference</b> Phase d'utilisation du modèle (votre question = inférence). Oppose à l'entraînement, beaucoup plus coûteux en calcul et en énergie.	

## Techniques & Methodes

<p><b>Prompt</b></p> <p>Instruction donnée à une IA. La qualité du prompt détermine la qualité de la réponse. Maîtriser le prompting, c'est maîtriser l'IA.</p>	<p><b>Prompt Engineering</b></p> <p>Art de formuler des instructions optimales. Techniques : roleplay, few-shot, chain of thought, décomposition de tâches, format imposé.</p>
<p><b>Chain of Thought (CoT)</b></p> <p>Demander à l'IA de raisonner étape par étape avant de conclure. Améliore significativement les réponses complexes.</p>	<p><b>Few-shot prompting</b></p> <p>Donner 2-3 exemples du format attendu dans le prompt avant de poser la vraie question. Guide l'IA vers le résultat souhaité.</p>
<p><b>RAG</b></p> <p>Retrieval-Augmented Generation. Connecter un LLM à vos propres documents. L'IA cherche l'info pertinente puis génère une réponse ancrée dans vos sources.</p>	<p><b>Fine-tuning</b></p> <p>Reentraîner un modèle sur des données spécifiques pour le spécialiser. Adapter un LLM aux documents RH ou à une terminologie métier.</p>
<p><b>Embedding</b></p> <p>Représentation mathématique d'un texte en vecteurs. Mesure la proximité sémantique entre concepts. Base technique du RAG.</p>	<p><b>RLHF</b></p> <p>Reinforcement Learning from Human Feedback. Des humains notent les réponses de l'IA pour l'améliorer progressivement par renforcement.</p>

## IA Generative

<p><b>IA Generative</b></p> <p>IA qui crée du contenu original : texte, images, audio, vidéo, code. Oppose à l'IA discriminative qui classe l'existant.</p>	<p><b>Multimodal</b></p> <p>Modèle traitant texte + images + audio + vidéo simultanément. GPT-4o, Gemini 2.0 et Claude 3 sont multimodaux.</p>
<p><b>Text-to-Image</b></p> <p>Générer une image depuis une description textuelle. Exemples : Midjourney, DALL-E 3, Stable Diffusion, Adobe Firefly.</p>	<p><b>Deepfake</b></p> <p>Contenu synthétique imitant une vraie personne (image, vidéo, voix). Risques majeurs de désinformation et de manipulation.</p>

## Agents & Automatisation

<p><b>Agent IA</b></p> <p>Système capable d'actions autonomes : naviguer sur le web, exécuter du code, gérer des fichiers, interagir avec des services.</p>	<p><b>Workflow d'automatisation</b></p> <p>Enchaînement automatisé de tâches IA. Exemple : recevoir un email, l'analyser, rédiger une réponse, l'envoyer. Sans intervention humaine.</p>
<p><b>MCP — Model Context Protocol</b></p> <p>Standard d'Anthropic connectant les LLM à des outils externes (Google Drive, Gmail, bases de données). Fondement des agents IA.</p>	<p><b>Orchestration</b></p> <p>Coordination de plusieurs agents IA pour accomplir une tâche complexe. Outils : LangChain, LangGraph, n8n, Make.com.</p>

## Ethique & Gouvernance

<b>Biais algorithmique</b> Erreur systematique dans les predictions d'une IA causee par les biais dans les donnees. Risque de discrimination en RH, credit, justice.	<b>Alignement IA</b> Defi : s'assurer que les IA agissent conformement aux valeurs humaines. Enjeu central de la securite des IA avancees.
<b>AI Act</b> Reglementation europeenne 2024. Premier cadre legal mondial sur l'IA. Classe les usages par niveau de risque et impose des obligations.	<b>RGPD et IA</b> Le RGPD s'applique aux IA traitant des donnees personnelles. Droits : acces, rectification, effacement, opposition, portabilite.
<b>AGI</b> Artificial General Intelligence. IA theorique capable de toutes les taches cognitives humaines. Non atteinte en 2026. Sujet de recherche fondamentale.	<b>Empreinte carbone</b> Entrainer de grands modeles consomme enormement d'electricite et d'eau. Enjeu environnemental majeur et croissant pour l'industrie IA.

## Modeles Open Source

<b>Open Source (modele)</b> Poids et code publics et librement utilisables. Exemples : Llama (Meta), Mistral 7B, Falcon. Utilisable en local sans envoyer de donnees.	<b>Ollama</b> Outil pour telecharger et executer des LLM sur votre ordinateur. Simple, gratuit, supporte Llama, Mistral, Phi et de nombreux autres.
<b>Quantification</b> Reduire la taille d'un modele (16 bits vers 4 bits) pour le rendre utilisable sur du materiel grand public avec peu de perte de qualite.	