



L'IA applicative

Découvrez la réalité derrière l'intelligence
artificielle applicative

SWIPE POUR APPRENDRE →



Qu'est-ce qu'une IA applicative ?

Une brique système

L'IA n'est pas le produit en soi, elle fait partie d'un système plus large.

Un outil intégré

Un outil intégré dans un produit, au service d'un besoin métier précis.

Contraintes fortes

Il doit respecter des contraintes de coût, de sécurité, de performance et de traçabilité des sources.



Exemples concrets d'IA applicative



Chatbot RH

Assistant virtuel pour répondre aux questions des employés.



Support client

Automatisation des réponses aux demandes courantes.



Recherche documentaire

Moteur de recherche intelligent pour documents internes.



POINT CLÉ

En entreprise, on ne fait presque jamais de recherche en IA

Pas de réentraînement

On ne réentraîne pas les modèles depuis zéro.

Poids fixes

On ne modifie pas les règles internes du modèle en permanence.

Pas d'expérimentation

On ne développe pas de nouveaux modèles ou algorithmes en continu.



Les 3 grandes phases d'un modèle IA

01

Entraînement (Training)

Le modèle est entraîné une fois sur de très grandes quantités de données.

02

Fine-tuning (Ajustement)

Ajustement léger pour adapter le modèle à un domaine précis.

03

Inférence (Utilisation)

Le modèle est utilisé en production pour générer des réponses.



PHASE 1

L'entraînement : rare et coûteux



Très coûteux

Demande énormément de calcul, temps et argent



Acteurs spécialisés

Fait par des entreprises comme OpenAI, Google, Meta



Déjà terminé

Cette phase est terminée avant que vous utilisiez le modèle

En entreprise, vous ne faites presque jamais cette étape

Le fine-tuning : pas toujours nécessaire

Le fine-tunning

- n'est pas obligatoire
- n'est pas automatique
- est coûteux
- peut générer des risques :
 - modèle trop spécialisé,
 - biais,
 - maintenance complexe



Beaucoup d'équipes préfèrent des approches plus simples comme le RAG.

PHASE 3

L'inférence : la phase principale

1

Entrée

Le modèle reçoit une demande en entrée (prompt).

2

Calcul

L'IA effectue des calculs statistiques pour déterminer la réponse la plus plausible, à partir de ce qu'elle a appris.

3

Sortie

L'IA renvoie la réponse qu'elle a générée.

- ❑ **Phrase clé :** En production, une IA est utilisée telle quelle. Elle ne s'adapte pas toute seule aux utilisateurs.

Pourquoi l'inférence seule ne suffit pas

Données internes

Le modèle ne connaît pas vos données internes spécifiques

Hallucinations

Peut inventer des informations non vérifiées

Non-conformité

Peut répondre de manière non conforme au métier

Solution : Le RAG (Génération Augmentée de Récupération) connecte les modèles IA aux bases de données internes de l'entreprise.

(Sujet que je détaillerai dans un prochain compte rendu 😊)

À retenir

Dans les logiciels, on n'entraîne pas les modèles d'IA en continu.

On les utilise en phase d'inférence.

Le modèle ne change pas

Il ne modifie pas son comportement tout seul

Adaptation via les données

On agit sur les données fournies en entrée

Le RAG comme solution

Pour adapter l'IA au contexte métier

Merci de votre attention !

Partagez ce post si vous l'avez trouvé utile !

Sarah LEON