

Intelligence Artificielle Générative (IAGen)

L'intelligence artificielle (IA) désigne un ensemble de méthodes informatiques qui permettent à des machines d'accomplir des tâches que l'on considère généralement comme « intelligentes » : reconnaître une image, prédire un résultat, dialoguer, etc.



Public Professionnels, Grand public, Médiateurs



Licence CC-BY



Type de
contenu

Mediation

Sommaire

Qu'est-ce que l'intelligence artificielle ?

Un mot sur le vocabulaire

Comment ça fonctionne ?

Les deux grands types d'apprentissage

Apprentissage supervisé

Apprentissage non supervisé

Et les IA génératives ?

Limites du système

Dialoguer avec une IAGen : le prompt

Comment rédiger un bon prompt ?

Les dérives et les recommandations

Les dérives les plus fréquentes

Hallucinations

Biais et discriminations

Opacité des décisions (effet "boîte noire")

Impact environnemental

Pour aller plus loin :

Commentaires

Qu'est-ce que l'intelligence artificielle ?

L'intelligence artificielle (IA) désigne un **ensemble de méthodes informatiques** qui permettent à des machines d'accomplir des tâches que l'on considère généralement comme « intelligentes » : reconnaître une image, prédire un résultat, dialoguer, etc.

Mais attention : le terme "**intelligence**" est **trompeur**. Il ne s'agit pas d'une intelligence au sens humain. Une IA ne **comprend pas** ce qu'elle fait, elle ne pense pas, ne ressent rien. Elle **calcule, statistiquement**, ce qui semble être la meilleure réponse.

Un mot sur le vocabulaire

Parler "d'une IA" n'est pas toujours juste. L'IA n'est pas une machine unique, mais une **discipline scientifique** (comme les mathématiques). On devrait dire "un outil fondé sur l'IA" ou "un système d'IA".

Médecine :
Faire une analyse d'une
IRM ou scanner



(Paros - CHU de Rennes)

Assainissement:
PurControl



Direction d'assainissement

Traitement de l'image :
Reconnaissance des plaques
d'immatriculations en ville



Comment ça fonctionne ?

L'intelligence artificielle fonctionne grâce à des **algorithmes**, c'est-à-dire des programmes informatiques capables d'apprendre à partir de données. Contrairement à un logiciel classique (comme Excel ou Word) qui suit des règles précises, une IA **s'adapte** et **s'améliore** en traitant de grandes quantités d'informations.

Les deux grands types d'apprentissage

Apprentissage supervisé

Exemple : on montre à l'IA 10 000 images de chats avec l'étiquette « chat ». L'IA apprend à reconnaître les caractéristiques d'un chat. Ensuite, on lui donne une nouvelle image, et elle tente de dire si c'est un chat ou non.

Apprentissage non supervisé

L'**apprentissage non supervisé** est une méthode où l'IA apprend **à partir de données non étiquetées**, c'est-à-dire sans que l'humain lui dise ce que contient chaque donnée. Contrairement à l'apprentissage supervisé (où l'IA apprend avec des exemples "chat", "chien", etc.), ici elle doit **trouver seule des régularités**, des similarités, des regroupements dans les données.

Exemple avec une IA générative de texte

Pour une IA comme **ChatGPT**, l'apprentissage non supervisé consiste à :

1. Lire **des milliards de phrases** issues de livres, sites, articles, sans indication du "bon mot".
2. Apprendre à **prédire le mot suivant** dans une phrase, en s'appuyant sur la logique des phrases vues.

Elle apprend ainsi à **générer du texte cohérent**, sans que personne ne lui ait dit ce qu'est un sujet, un verbe ou une idée logique. Elle a découvert les **structures du langage** toute seule.

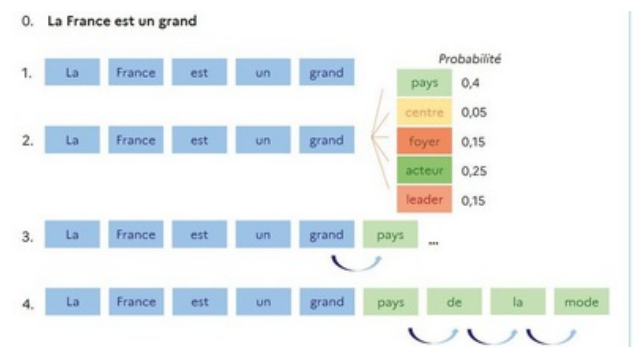
Et les IA génératives ?

Ce sont des IA qui **créent du contenu** : un texte, une image, une musique, etc. Leur fonctionnement repose aussi sur l'apprentissage, mais avec un objectif différent : **imiter** ce qu'elles ont vu, pour **générer quelque chose de nouveau**.

Exemples :

- **ChatGPT** génère du texte à partir d'une consigne.
- **DALL-E** ou **Midjourney** créent des images à partir d'une description.

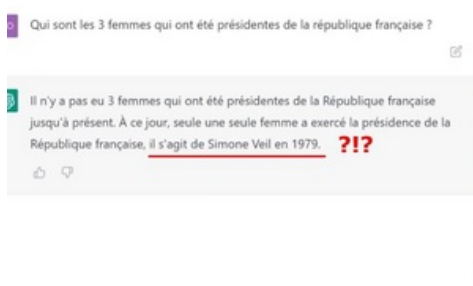
Le principe est le même : elles **prédissent** ce qui semble le plus probable. ChatGPT, par exemple, **ne réfléchit pas** : il **calcule quel mot devrait venir après l'autre**, en s'appuyant sur des milliards de phrases analysées lors de son entraînement.



Un modèle de langage ne comprend pas le sens des mots. Il est probabiliste. Ce qui n'empêche pas d'être pertinent.

Limites du système

- Les IA sont **très puissantes**, mais aussi **très fragiles**.
- Elles peuvent produire des erreurs, des "hallucinations", ou reproduire des biais humains présents dans les données.
- Elles **ne savent pas ce qu'elles ne savent pas** : elles n'ont aucun recul sur leurs propres réponses.



Dialoguer avec une IAGen : le prompt

Les intelligences artificielles génératives (comme ChatGPT) sont conçues pour interagir en **langage naturel**. On n'a plus besoin de savoir programmer : il suffit d'écrire ce qu'on veut dans une phrase claire et précise. C'est ce qu'on appelle un **prompt**.

Comment rédiger un bon prompt ?

Voici 5 astuces :

1. **Sois clair·e et précis·e** → Évite les formulations vagues ou ambiguës.
2. **Donne un rôle à l'IA** → "Tu es un expert RH", "Tu es un formateur en bureautique", etc.
3. **Indique un objectif** → "Tu dois m'aider à structurer une présentation en 3 parties."
4. **Précise le format attendu** → "Donne la réponse sous forme de liste", "Rédige un texte court de 100 mots."
5. **Ajoute le contexte si nécessaire** → Plus l'IA a d'éléments, plus sa réponse est adaptée.

Il ne faut pas hésiter à **réajuster sa demande** après une première réponse. Dialoguer avec une IAGen, c'est souvent comme un échange : on affine la consigne en plusieurs étapes.

Les dérives et les recommandations

L'intelligence artificielle offre de nombreux **potentiels professionnels**, mais son usage soulève des **risques importants**. Il ne s'agit pas de renoncer à l'outil, mais de **l'utiliser en conscience**, dans un cadre éthique et réglementaire.

Les dérives les plus fréquentes

Hallucinations

L'IA peut **inventer des faits** ou répondre à côté. Elle ne "sait" pas quand elle se trompe. Il faut donc **vérifier systématiquement** ses réponses.

Biais et discriminations

Les IA apprennent à partir de données humaines, donc souvent **biaisées**. Elles peuvent reproduire (ou amplifier) des inégalités existantes : sexisme, racisme, stéréotypes sociaux...

Opacité des décisions (effet "boîte noire")

Difficile d'expliquer comment une IA est arrivée à une décision. Cela pose des questions pour les **services publics**, qui doivent garantir **transparence et équité**.

Impact environnemental

Les IA génératives consomment énormément d'énergie (calculs, data centers). L'empreinte carbone du numérique est estimée à **3 à 4 %** des émissions mondiales, en forte augmentation.

Pour aller plus loin :

- Le service public à l'épreuve de l'intelligence artificielle : <https://www.sens-du-service-public.fr/publications>
- Le rapport "IA : notre ambition pour la France" :
<https://www.info.gouv.fr/upload/media/content/0001/09/4d3cc456dd2f5b9d79ee75feea63b47f10d75158.pdf>
- Impacts de l'intelligence artificielle : risques et opportunités pour l'environnement :
https://www.lecese.fr/sites/default/files/pdf/Avis/2024/2024_14_IA_Environnement.pdf
- Cartographie des métiers concernés par l'intelligence artificielle dans les collectivités : https://inet.cnfpt.fr/sites/default/files/2024-04/Cartographie_metiers_concernes_IA.pdf