

1

Qu'est-ce que l'IA générative ?

1.1. L'illusion de l'intelligence

« Mais si l'IA sait tout, elle peut me donner les chiffres du Loto ? »

Cette question, à la fois naïve et provocatrice, je l'entends souvent lors de mes formations ou interventions. Elle n'est pas anodine. Elle traduit une incompréhension fondamentale de ce qu'est réellement l'intelligence artificielle et en particulier l'intelligence artificielle générative.

Elle traduit aussi une fascination : celle d'un outil capable de parler, d'écrire, d'expliquer, de raisonner, parfois même de plaisanter, comme un humain.

Ce malentendu n'est pas nouveau. Chaque révolution technologique s'accompagne d'une forme de projection humanisée. Quand les premiers automates du XVIII^e siècle ont commencé à jouer du clavecin ou à dessiner, certains spectateurs y voyaient déjà une forme de conscience. Plus tard, lorsqu'un ordinateur a battu un champion d'échecs, le public a cru que la machine pensait. Nous confondons encore aujourd'hui la puissance de calcul avec la capacité de réflexion. L'IA générative, par son aisance linguistique, accentue cette confusion. Elle donne le sentiment qu'elle comprend, alors qu'elle ne fait qu'imiter la structure du discours.

Pour comprendre cette illusion, il faut s'attarder sur ce que fait concrètement un modèle génératif lorsqu'on lui pose une question. Prenons un exemple simple : Pourquoi le ciel est bleu ? Un être humain mobilisera son expérience, ses apprentissages scolaires, sa mémoire des phénomènes physiques. Il se souviendra peut-être d'avoir lu ou entendu parler de la diffusion de la lumière par l'atmosphère. L'IA générative, elle, ne se souvient de rien. Elle n'a pas de mémoire personnelle, pas d'expérience vécue, pas de compréhension du monde. Ce qu'elle fait, c'est prédire la suite la plus probable de la phrase en fonction de milliards d'exemples lus durant son entraînement. Elle calcule la probabilité que le mot *diffusion* suive *le ciel est bleu parce que* et que le mot *lumière* suive *diffusion*. Ce processus statistique, répété des millions de fois par seconde, produit une réponse cohérente.

Autrement dit, l'IA ne sait pas pourquoi le ciel est bleu ; elle sait seulement que, dans les textes humains, cette explication revient fréquemment lorsqu'on parle du ciel et de sa couleur. C'est une intelligence de la probabilité, pas de la compréhension. Cette distinction peut sembler subtile, mais elle change tout. Elle explique pourquoi une IA générative peut produire des réponses stupéfiantes de pertinence dans un domaine précis et, à d'autres moments, des absurdités totales dites avec une assurance déconcertante.

Pour bien comprendre cette illusion, je m'appuie souvent sur une situation vécue en formation. Lors d'un atelier, j'avais fourni à une IA générative un document PDF assez conséquent (une cinquantaine de pages) contenant des données statistiques d'un secteur d'activité. Les informations étaient structurées par mois, avec des montants et des chiffres d'affaires.

Je lui ai alors demandé d'extraire les chiffres d'affaires de ce document et d'en faire un tableau récapitulatif, avec pour chaque mois le montant correspondant. Le résultat, à première vue, était impeccable : un tableau bien structuré, lisible, cohérent, exactement ce que l'on attend d'un tel exercice. Pourtant, en relisant plus attentivement, un chiffre m'a interpellé. Un montant apparaissait pour un mois précis... alors que je savais que cette information n'existait pas dans le document source. Après vérification, le doute n'était plus permis : cette donnée n'était nulle part dans le PDF.

Je lui ai donc posé la question directement : d'où venait ce chiffre ?

La réponse a été très instructive. L'IA m'a expliqué qu'elle n'avait pas trouvé l'information dans le document et qu'elle avait donc produit une valeur *plausible*, afin de proposer un tableau complet et cohérent visuellement. Cet exemple illustre parfaitement un point fondamental : l'IA générative ne cherche pas à avoir raison. Elle ne vérifie pas la véracité des informations comme le ferait un humain. Son objectif est de produire un contenu qui a l'air correct, fluide et logique du point de vue du langage. Sa cohérence linguistique n'est en aucun cas une garantie de véracité factuelle.

Cette illusion d'intelligence provient aussi de notre propre comportement d'utilisateur. Nous avons tendance à projeter sur la machine des intentions qu'elle n'a pas. Quand une IA dit « je », nous oublions qu'il ne s'agit pas d'un sujet. Ce « je » n'est qu'une convention syntaxique, un artefact de langage qui crée l'illusion d'un interlocuteur. Lorsque ChatGPT répond avec empathie à une question sensible, il ne ressent rien ; il reproduit les schémas discursifs qu'il a appris dans des dialogues humains. Si vous lui dites « je suis triste », il répondra « je suis désolé de l'apprendre » non par compassion, mais parce que cette phrase est statistiquement la plus probable pour répondre à une déclaration de tristesse.

L'émotion n'est pas vécue, elle est simulée.

Cette simulation est d'autant plus convaincante que la machine maîtrise les codes du langage naturel. Pendant des décennies, les ordinateurs communiquaient en syntaxe rigide, sous forme de commandes. L'utilisateur devait s'adapter à la machine. Aujourd'hui, la situation s'est inversée : la machine s'adapte à notre langage. Elle parle comme nous, écrit comme nous, parfois même mieux que nous. Cette fluidité crée un brouillage entre forme et intention. Nous interprétons la qualité du texte comme le signe d'une réflexion sous-jacente, alors qu'il n'y a qu'une mécanique d'optimisation statistique. Cette confusion n'est pas sans conséquence.

Dans les entreprises où j'interviens, il n'est pas rare d'entendre des collaborateurs dire : « ChatGPT s'est trompé », ou « il a mal compris la question ». Mais en réalité, il n'a ni raison ni tort, ni compréhension ni incompréhension : il a seulement produit la suite la plus probable d'une phrase, à partir d'un contexte que nous avons, nous, défini plus ou moins clairement. L'erreur est souvent dans le **prompt**, pas dans le modèle.

Par exemple, si je demande à une IA : Écris-moi un script Python pour calculer la surface de mes polygones dans QGIS sans préciser les chemins d'accès ou le système de projection, elle produira malgré tout un script qui a l'air correct. Le code sera propre, lisible, parfois même commenté... mais potentiellement inexécutable ou faux dans mon contexte réel. L'IA ne s'est pas trompée : elle a simplement répondu à une demande incomplète.

Cette notion est essentielle pour aborder le reste de ce livre.

L'IA générative ne remplace pas l'expertise humaine ; elle la reflète et l'amplifie. Lorsque vous demandez à une IA d'écrire un script Python pour automatiser une tâche, la qualité du résultat dépendra entièrement de votre capacité à expliciter votre besoin. La machine ne comprend pas votre métier. Elle ne connaît pas vos contraintes, vos données, vos habitudes de travail. Elle ne fait que deviner, en s'appuyant sur des textes similaires. Si votre demande est vague, elle s'égarera ; si elle est précise, elle vous étonnera.

On pourrait comparer cela à un stagiaire très doué, mais sans expérience : il apprend vite, produit beaucoup, mais commet des erreurs s'il n'a pas de cadre clair. L'IA générative fonctionne de la même manière : elle excelle dans la reproduction des régularités du langage, mais échoue dès qu'on la sort de ce qu'elle peut inférer des textes qui l'ont nourrie.

Cette illusion d'intelligence est donc le miroir de notre propre projection. Nous sommes bluffés par la forme et frustrés par la limite. Mais reconnaître cette illusion, c'est déjà entrer dans une posture d'utilisateur éclairé. C'est comprendre que la valeur de l'IA générative ne réside pas dans une prétendue conscience, mais dans sa puissance à formaliser et à reformuler les savoirs humains.