

<https://larcenciel.be/spip.php?article1462>



# CHACUN CHERCHE SON "CHAT"

- MATIÈRE À PENSER - ENJEUX DU NUMÉRIQUE - ENJEUX DE L'INTELLIGENCE ARTIFICIELLE -

Date de mise en ligne : lundi 27 mars 2023

---

Copyright © LARCENCIEL - site de Michel Simonis - Tous droits réservés

---

## La ruée vers l'intelligence artificielle

**Après ChatGPT, dans laquelle Microsoft a largement investi, c'est au tour du géant Google d'annoncer la sortie imminente de son propre agent conversationnel à base d'intelligence artificielle : Bard, attendu au mois de mars.**

**Sur France Culture, le 13 février 2022, le point de vue de Laurence Devillers, professeur d'intelligence artificielle à la Sorbonne, chercheur au CNRS et auteure de l'essai Les robots Émotionnels (Editions de l'Observatoire)**

## Sommaire

- [A - Pourquoi les moteurs de recherche ont été dévoyés de leurs objectifs initiaux ?](#)
  - [Rappel en bref : Précisions sur les moteurs de recherche](#)
  - [L'échec de Google \(comment google à foiré\)](#)
  - [les moteurs de recherche à la recherche de nouvelles modalités](#)
- [B - Le deep learning, qu'est-ce que c'est ?](#)
- [C - "ChatGPT est un affabulateur qui répond toujours même quand il ne sait pas"](#)

Bard, Ernie, ChatGPT... A travers le monde, un nombre important de robots conversationnels voient le jour au même moment.

Les chercheurs travaillent ensemble, depuis 2017, sur l'utilisation de grands modèles génératifs capables d'explorer des milliards de données et d'arriver à créer des modèles avec des réseaux de neurones.

### ChatGPT : un système performant mais dangereux ?

"Ce système ne raisonne pas", rappelle Laurence Devillers. Localement, la syntaxe est bonne. Donc on peut être très leurré par cela. Et on a l'impression que ce qu'il dit est vrai, alors qu'il raconte n'importe quoi. Il n'a pas de compréhension temporelle, il peut mélanger des dates, il n'a pas de compréhension de logique, pas de raisonnement d'induction/déduction, c'est vraiment des probabilités de succession de mots." La professeure admet qu'il faut quand même "reconnaître la performance. Mais il faut absolument éduquer tout le monde. On oublie que cela génère des choses fausses (...)

"Pour l'instant, ChatGPT qui parle notre langue, parle en fait avec 50% de connaissances en anglais." "On risque d'être submergés par du contenu sémantique qui vient soit de l'Asie, soit des États Unis. (...) Il est urgent que l'on comprenne qu'il faut absolument réglementer, comprendre ce que sont les notions d'éthique, mais en même temps créer et innover très fortement en Europe" conclut-elle.

Tout cela pose un certain nombre de questions qui permettent d'apporter quelques informations complémentaires.

Et pour commencer :

# A - Pourquoi les moteurs de recherche ont été dévoyés de leurs objectifs initiaux ?

## Rappel en bref : Précisions sur les moteurs de recherche

### Qu'est-ce qu'un moteur de recherche ?

Un moteur de recherche est un outil qui permet de parcourir le web à la recherche de pages et de documents correspondant à un ou plusieurs mots-clés.

L'utilisateur, qui cherche des informations ou à effectuer un achat, est censé trouver rapidement la réponse à sa question grâce au moteur de recherche.

Celui-ci possède un index d'une gigantesque quantité de documents web (pages, images, vidéos, etc.), et propose des résultats qui sont classés en fonction de critères qu'il a mis en place dans un algorithme.

Ces résultats de recherche s'affichent en général sur une ou plusieurs pages (les SERPs), sous forme de liste de liens cliquables.

Dans ce domaine, c'est Google qui a réussi à imposer un quasi monopole ces 20 dernières années.

### Il subsiste cependant des alternatives...

Parmi les moteurs de recherches alternatifs au géant américain, on trouve des entreprises implantées régionalement pour des raisons linguistiques ou même politiques.

En dehors des géants russe ou chinois, la plupart des alternatives à Google proposent deux choses majeures que Google n'est pas en mesure d'apporter :

- La confidentialité et la protection des informations personnelles.
- Le reversement de bénéfices publicitaires issus des résultats de recherche à des associations écologiques et / ou solidaires.

Qu'est-ce qu'un moteur de recherche écologique ? Allons voir de plus près qui sont ces moteurs de recherche alternatifs. 1461'>VOIR

## L'échec de Google (comment google à foiré)

Vous entrez un mot ou une phrase dans un moteur de recherche et vous obtenez une pile d'informations souvent non pertinentes.

Google est excellent pour répondre à des questions dont la réponse est objective, comme "Nombre de milliardaires dans le monde" ou "Quelle est la population de l'Islande ?". Il est plutôt mauvais pour répondre à des questions qui requièrent un jugement et un contexte.

La mission déclarée d'une entreprise valant près de deux mille milliards de dollars est d'"organiser l'information du monde" et pourtant l'Internet reste mal organisé. Ou, autrement dit, dans un monde où l'information est infinie, il ne suffit plus d'organiser l'information mondiale. Il devient important d'organiser **l'information digne de confiance** du monde. [1]]

Ce qui était au départ (en 1998) un moyen bien intentionné d'organiser les informations du monde entier s'est transformé en une entreprise qui consacre l'essentiel de ses ressources à la monétisation des clics pour soutenir les annonceurs, au lieu de privilégier la fourniture de résultats de recherche fiables aux internautes.

Avec l'avènement de Google AdWords, il est devenu rentable de diffuser du contenu de faible qualité qui passait pour informatif et remplissait les résultats du moteur de recherche de Google.

Le problème est aujourd'hui très différent de celui d'il y a vingt-cinq ans. Il n'est plus de savoir quoi lire/acheter/manger/regarder/etc. mais de trouver la meilleure chose à lire/acheter/manger/regarder/etc. avec un temps et une attention limités.

Nous sommes encore loin de réaliser la grande vision de l'Internet. Le projet de connaissance humaine, tel qu'il se présente aujourd'hui, est un vaste océan d'informations et d'idées éphémères et fragmentées, dont les meilleures sources sont difficiles à trouver.

Des initiatives, comme DuckDuckGo et Neeva, tentent de concurrencer Google de front en construisant de robustes moteurs de recherche horizontaux (VOIR...).

Plutôt que d'explorer et d'indexer les données à leur manière, elles se placent au-dessus des sources de données existantes et se positionnent comme une alternative à Google axée sur la protection de la vie privée.

Pour la grande majorité des internautes, le fait de pouvoir contrôler leurs propres données n'est pas un argument de vente, surtout si cela implique de payer pour quelque chose qu'ils ont l'habitude d'obtenir gratuitement.

## les moteurs de recherche à la recherche de nouvelles modalités

Dès lors, maintenant, on cherche à construire des moteurs de recherche spécialisés qui indexent, conservent et organisent les choses d'une nouvelle manière. La solution est une meilleure recherche et une meilleure sélection, où la valeur réside dans ce qui est exclu autant que dans ce qui est inclus : une combinaison que les anglo-saxons nomment [boutique search engines](#).

L'opportunité dans le domaine de la recherche n'est donc pas d'attaquer Google de front avec un puissant agrégateur horizontal, mais plutôt de construire des moteurs de recherche spécialisés qui indexent, conservent et organisent les choses d'une nouvelle manière.

## B - Le deep learning, qu'est-ce que c'est ?

Utilisé par Siri, Cortana et Google Now pour comprendre la voix ou être capable d'apprendre à reconnaître des visages, le deep learning (apprentissage profond en français) est souvent confondu avec le concept d'intelligence artificielle (IA). À tel point que l'on croit que ces deux mots sont des synonymes.

Pourtant, ce n'est pas du tout le cas. Le deep learning est une branche de l'IA.

Explications. L'intelligence artificielle est née avec le développement de l'ordinateur et de l'informatique à la fin des années 1980. Différents courants ont vu le jour. L'un d'eux consistait à s'inspirer du fonctionnement du cerveau humain afin de tenter de créer des neurones artificiels. La première machine a été créée en 1951 par deux chercheurs de l'université de Harvard.

Mais ce n'est que récemment, grâce à l'avancée des performances de calcul des ordinateurs, que s'est développé le concept de **deep learning**. Il s'agit de réseaux de neurones disposant de nombreuses couches cachées.

Ceci mérite quelques explications...

1 - Concrètement, le deep learning est une technologie qui apprend à une machine à se représenter le monde. C'est une technique d'apprentissage permettant à un programme de reconnaître le contenu d'une image ou de comprendre le langage parlé. Dans le passé, pour faire cela, les ingénieurs expliquaient à la machine comment représenter une image. Avec le deep learning, elle le fait désormais toute seule.

### 2 - Des couches de neurones pour un apprentissage supervisé décuplé

Pour comprendre comment une machine parvient à cette prouesse, il faut évoquer l'**apprentissage supervisé**. Il s'agit d'une technique courante en IA, qui consiste à nourrir la machine de nombreuses informations. Par exemple, pour qu'un programme apprenne à reconnaître une voiture, on le "nourrit" de dizaines de milliers d'images de voitures, étiquetées comme telles. Une fois l'entraînement achevé (celui-ci peut durer des heures, voire des jours), le programme peut reconnaître des voitures sur de nouvelles images.

*Le deep learning utilise lui aussi l'apprentissage supervisé mais c'est l'architecture interne de la machine qui est différente : chacune des milliers d'unités composant le réseau de neurones effectue des petits calculs simples.*

Yann Ollivier, chercheur au CNRS, explique le procédé avec un exemple : *"Comment reconnaître une image de chat ? Les points saillants sont les yeux et les oreilles. Comment reconnaître une oreille de chat ? L'angle est à peu près de 45 degrés. Pour reconnaître la présence d'une ligne, la première couche de neurones va comparer la différence des pixels au-dessus et en dessous : cela donnera une caractéristique de niveau 1. La deuxième couche va travailler sur ces caractéristiques et les combiner entre elles. S'il y a deux lignes qui se rencontrent à 45°, elle va commencer à reconnaître le triangle de l'oreille de chat. Et ainsi de suite".* A chaque étape, le réseau de neurones approfondit sa compréhension de l'image.

---

Et voici donc les questions qui s'enchaînent et dont les réponses apparaissent dans les articles qui suivent. voir notamment les articles "[Déclic. Le tournant GPT](#)" et "[Le sens des mots](#)"

### 3 - et l'apprentissage supervisé ?

– C'est quoi exactement ces "couches de neurones pour un apprentissage supervisé décuplé" ?

### 4 - Des algorithmes toujours plus gourmands en énergie...

A-t-on une idée de l'empreinte carbone qui se cache derrière un algorithme de "deep learning" ?

### 5 - que penser de l'alternative "Carbontracker, pour une puissance de calcul plus verte" ?

### 6 - A propos, quels sites internet polluent le plus ? Les mauvais et les bons élèves

### 7 - ChatGPT et l'E-commerce

L'entreprise de commerce en ligne Shopify, utilise aussi l'API de ChatGPT pour son nouveau chatbot sur son application mobile "Shop".

Et les adeptes d'Instacart, qui permet de faire ses courses et de se les faire livrer, pourront bientôt demander des recommandations à l'appli pour des recettes de cuisine qui se transformeront directement en listes de courses, toujours grâce à l'API d'OpenAI.

### 8 - Microsoft ?

Microsoft a créé la surprise début février en intégrant ChatGPT à son moteur de recherche Bing, sous la forme d'un chatbot capable de répondre aux questions des utilisateurs mais aussi de générer différents textes sur commande.

Le groupe informatique américain avait déjà investi un milliard de dollars dans OpenAI, la start-up à l'origine des modèles de langage qui alimentent ChatGPT, et a passé cette année un nouvel accord de plusieurs milliards avec cette entreprise californienne à la pointe de l'IA générative.

Les technologies du nouveau Bing, en phase de test, doivent être intégrées à Windows 11 (la dernière version de son système d'exploitation sur les PC), à sa suite bureautique (dans des logiciels comme Word ou Teams) et sur son navigateur Edge.

### 9 - quel jeu joue Elon Musk ?

Le multi-millionnaire, qui a racheté Twitter dans l'idée d'y libérer la parole, estime que ChatGPT restreint trop la liberté d'expression en faveur des opinions de la gauche américaine "woke", d'après le site spécialisé.

Il aurait donc sollicité des chercheurs ces dernières semaines pour constituer un laboratoire concurrent.

### 10. Et encore :

Le chatbot d'OpenAI répond depuis le mois de novembre à toutes les questions des utilisateurs. L'entreprise américaine OpenAI – dont l'infrastructure de calcul est financée par Microsoft qui vient de réinvestir plusieurs milliards – a en effet bien "marketé" son lancement en offrant un essai gratuit de la version beta. L'outil aurait dépassé les 100 millions d'utilisateurs en janvier !

Une version payante ChatGPT Plus, avec notamment des temps de réponse rapides, a quant à elle été lancée aux Etats-Unis et sera "prochainement" proposée à d'autres pays pour une vingtaine de dollars par mois. Microsoft a également annoncé intégrer des fonctionnalités du chatbot dans sa nouvelle version premium de l'application Teams.

## C - "ChatGPT est un affabulateur qui répond toujours même quand il ne sait pas"

Lancé en novembre, il a été élaboré par la société OpenAI co-crée par Elon Musk en 2015, et dans laquelle Microsoft a réinvesti plusieurs milliards en janvier.

*"Tout d'abord, ChatGPT fait inéluctablement peser une menace sur l'emploi avec la prise en charge et l'automatisation accrue de tâches, notamment intellectuelles et créatives",* entame Denys Malengreau.

*"Par extension à ChatGPT, il existe également une menace au niveau des droits d'auteur",* explique M. Malengreau. *"Par exemple, le chatbot est capable de reproduire du code informatique sous licence... sans inclure la licence ou les auteurs originaux !",* illustre Gilles Louppe, professeur en intelligence artificielle à l'ULiège.

D'ailleurs, d'autres intelligences artificielles en cours de développement représentent une menace similaire.

il faut savoir que quand ChatGPT ne connaît pas la réponse à une question, il répond quand même. *"ChatGPT est un affabulateur qui va toujours répondre même quand il ne sait pas. Car il n'a pas de représentation interne du monde, de raisonnement ou de conscience comme un humain : il ne sait pas qu'il ne sait pas. Ses réponses, qu'elles soient justes ou fausses sont purement mécaniques",* éclaire Gilles Louppe.

Autrement dit, parfois, ChatGPT invente. *"L'outil facilite ainsi la production intentionnelle de faux contenu, ce qui présente un vrai risque d'accroître la diffusion de fausses informations en ligne",* lance Denys Malengreau. D'autant que lorsqu'on lui demande de citer ses sources, le chatbot mentionne parfois des sources fausses ou qui n'existent pas comme de faux liens URL. *"Cela aggrave la crise actuelle de l'autorité institutionnelle, politique, médiatique, éducative et même parentale, c'est-à-dire une forme de défiance à l'égard des communications que l'on reçoit. Ce manque de confiance – 'plus personne ne peut plus rien croire' – conjugué à la production de fake news est même un risque d'atteinte au tissu démocratique",* accentue encore l'expert en économie numérique.

Ainsi, le fact-checking est fondamental. *"OpenAI travaille à ce que ChatGPT invente de moins en moins de réponses",* partage Gilles Louppe. *L'entreprise a même lancé un 'AI classifier' pour que les utilisateurs puissent repérer si un texte a été écrit par une intelligence artificielle" (ChatGPT mais pas seulement, NdlR) ou par un humain. Ce sont des outils qui vont commencer à émerger en masse".*

A noter l'expérience de Tay, l'IA de Microsoft qui rapidement après son introduction sur Twitter en 2016 a tenu des propos insultants, méchants, vulgaires ou encore racistes. Cette opération colossale a pris énormément de temps et s'est déroulée au Kenya, où la main-d'œuvre est payée environ 2 euros par jour.

*"La dépendance manuelle à des décisions humaines et à leur jugement fait que l'intelligence artificielle peut nécessairement être biaisée dans ses réponses. Les équipes d'OpenAI font des choix de développement, ont un certain rapport au monde, et c'est cette idéologie qui peut ressortir dans les réponses de ChatGPT. Il n'y a pas de technologie neutre, l'outil reste imprégné de l'idéologie dont il a pris sa source",* conclut M. Malengreau.