

Réflexions sur les impacts d'ordre sociétal, environnemental et managérial des IA génératives

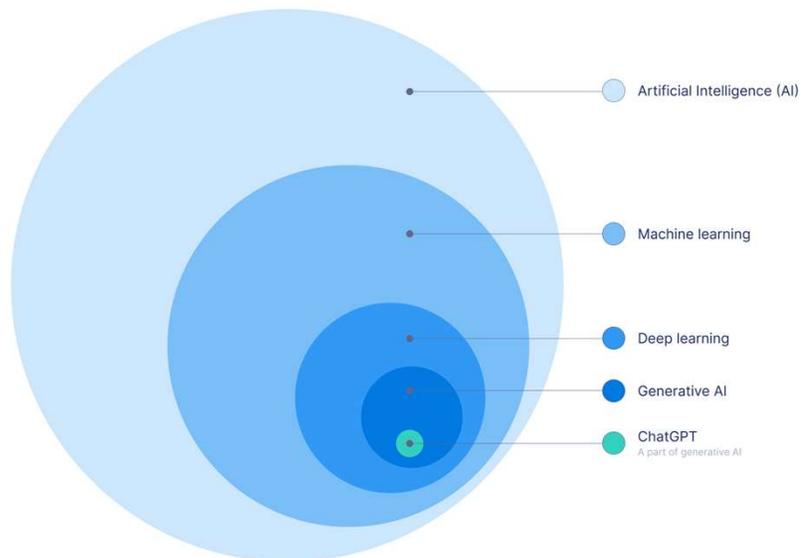
Nicolas Neysen

Liège, le 14 juin 2024



1

Large Language Model

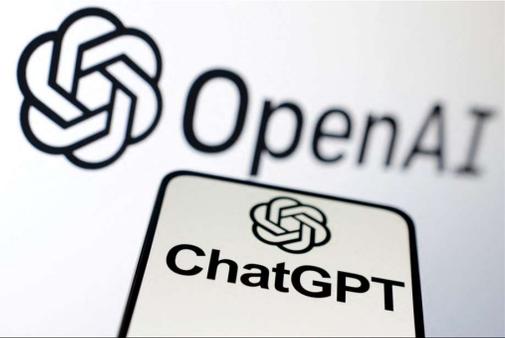


2

Large Language Model

Un LLM est un type de modèle en IA, basé sur une très grande quantité de texte existant. Le LLM utilise des techniques d'apprentissage profond (deep learning) pour prédire la séquence de mots à suivre. C'est donc un calcul de probabilité à la fois très complexe et très rapide, permettant de générer du texte qui semble naturel et cohérent.

L'algorithme derrière Chat-GPT3, contient 175 milliards de paramètres, ce qui en fait l'un des plus grands modèles de type LLM à ce jour.



3

ChatGPT went from 0 to 100M users in 60 days

Time it took for selected online services to reach 1 million users¹



¹ Source: Statista

4

Besoin d'aide !

N Vous

Agis comme un orateur talentueux avec le style des keynotes TedX. Donne moi une structure de présentation d'une durée de 30 minutes sur le thème "Réflexions sur les impacts d'ordre sociétal, environnemental et managérial des IA génératives". La présentation est destinée à un public formé d'agents de l'Institut Scientifique de Service Public.

5

1. Introduction (2 min)

- Établir le contexte : L'avènement des IA génératives et leur rôle croissant dans notre quotidien.
- Définir les IA génératives : Qu'est-ce que c'est et comment fonctionnent-elles ?
- Annonce du plan : Explorer les impacts sociétaux, environnementaux et managériaux de ces technologies.

2. Impacts Sociétaux des IA Génératives (7 min)

- Révolution de la créativité : Comment les IA génératives transforment l'art, la musique et la littérature.
- Questions éthiques : Les implications de l'utilisation d'IA génératives pour la création de contenu et les droits d'auteur.
- Nouveaux paradigmes de communication : L'influence des IA génératives sur la façon dont nous interagissons en ligne.

3. Impacts Environnementaux des IA Génératives (7 min)

- Consommation énergétique : Analyse de l'empreinte carbone des infrastructures nécessaires au fonctionnement des IA génératives.
- Gestion des données : L'impact de la collecte et du traitement massif de données sur l'environnement.
- Stratégies de durabilité : Explorer des solutions pour atténuer l'impact environnemental tout en continuant à innover dans le domaine des IA génératives.

4. Impacts Managériaux des IA Génératives (7 min)

Revenir sur les points clés de la présentation et encourager le public à prendre des mesures concrètes.

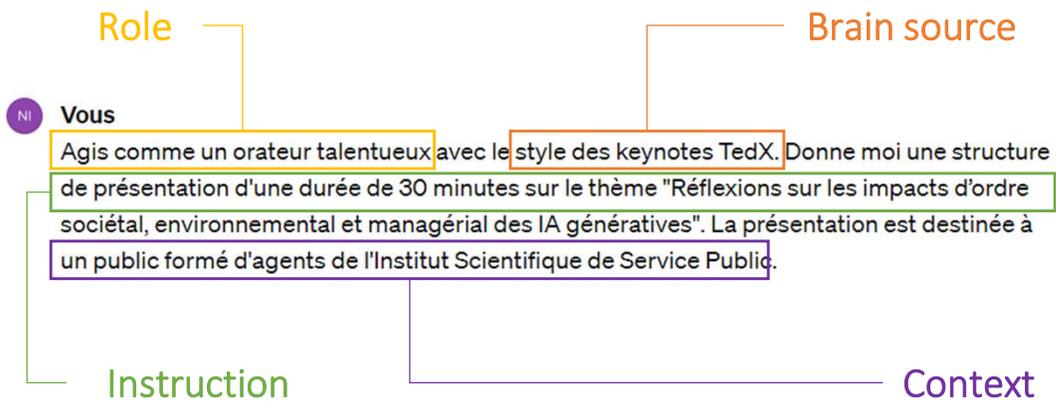
- Automatisation des processus : Comment les IA génératives transforment les pratiques commerciales et les processus de production.
- Gestion des talents : Les défis et opportunités liés à l'intégration des IA génératives dans les équipes de travail.
- Adaptabilité organisationnelle : Les entreprises doivent-elles se réinventer pour s'adapter à l'ère des IA génératives ?

5. Conclusion (7 min)

- Récapitulation des principaux points abordés.
- Appel à l'action : L'importance de rester conscients des implications des IA génératives et de travailler ensemble pour façonner un avenir éthique et durable.
- Invitation au dialogue : Ouvrir la discussion et encourager la collaboration pour explorer davantage ces questions cruciales.

6

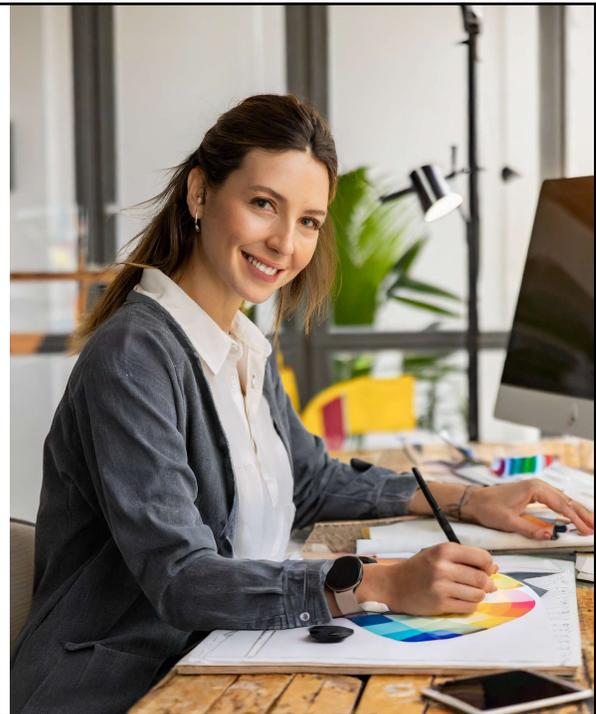
Prompt Manager... un métier qui n'existait pas voici 1 an



7

Des métiers voués à disparaître ?

- Traducteur/rice
- Graphiste
- Designer
- Programmeur/euse
- Recruteur/euse
- Rédacteur/rice
- Dessinateur/rice
- Acteur/rice
- ...
- Et le vôtre ?



8

L'impact des LLM sur les métiers et les tâches

Une étude américaine récente¹ menée sur base des données de l'O*NET, le système d'information US sur les professions (+1000 titres et +/- 25.000 tâches) indique que (seulement) **19% des travailleurs ont plus de la moitié de leurs tâches qui sont exposées (+ & -) à l'IA générative**

La thèse du remplacement massif (disparition) de métiers ne tiendrait donc pas. L'usage de l'IA générative est davantage vu comme un **moyen d'augmenter la productivité** du travailleur humain (e.g. GitHub Copilot).

Mais pour la première fois dans l'histoire industrielle, ce sont les cols blancs qui sont en première ligne...

¹ Source: McAfee A., Rock D., Brynjolfsson E., (Nov. 2023), Harvard Business Review.

9

LLM dans un service de support client

Une grande entreprise américaine dans le secteur des logiciels a implémenté une solution AI hybride (LLM+ML) pour ses 1.500 agents du département service client¹.

- Nombre de cas résolus / heure : +15%
- Nombre de cas traités en simultané : +15%
- Durée moyenne de conversation : -10%
- Diminution des signaux de FRUSTRATION !!!!
- Taux de résolution des employés:
 - 20% bas : +35%
 - 20% haut : +0%

¹ Source: McAfee A., Rock D., Brynjolfsson E., (Nov. 2023), Harvard Business Review.



10

LLM dans le conseil en stratégie

Question : Comment l'introduction de LLM dans les tâches de travailleurs du savoir hautement qualifiés influence-t-elle leur travail traditionnel ?

758 consultants BCG en stratégie répartis en 3 groupes : sans support IA, avec accès à Chat GPT-4, avec accès à Chat GPT-4 + guide pratique



¹ Source: Dell'Acqua *et al.*, Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality (2023). Harvard Business School Technology & Operations Mgt. Unit Working Paper No. 24-013.

11

La frontière en dents de scie de l'IA

Innovation & développement de produit. La tâche d'évaluation initiale demandait aux participants de réfléchir à des concepts de boissons innovantes (marque, slogan, communiqué de presse, analyse de marché, etc.)

Diagnostic et recommandations stratégiques. La tâche demandait aux participants de résoudre un cas pratique (« consulting case interview ») sur base d'une documentation variée (résultats financiers, retranscription d'interviews, communiqués d'entreprise, etc.)

Source: Dell'Acqua *et al.*, 2023.

— AI Abilities
- - - Equal Difficulty Tasks

Task on the Frontier

12

« Inside the frontier »

Qualité des réponses :

- Groupe « sans IA » : 4.10
- Groupe « avec IA » : 5.66 (+38%)
- Groupe « avec IA et guide » : 5.85 (+43%)

Quantité de tâches accomplies :

- Groupe « sans IA » : 82%
- Groupe « avec IA » : 91%
- Groupe « avec IA et guide » : 93%

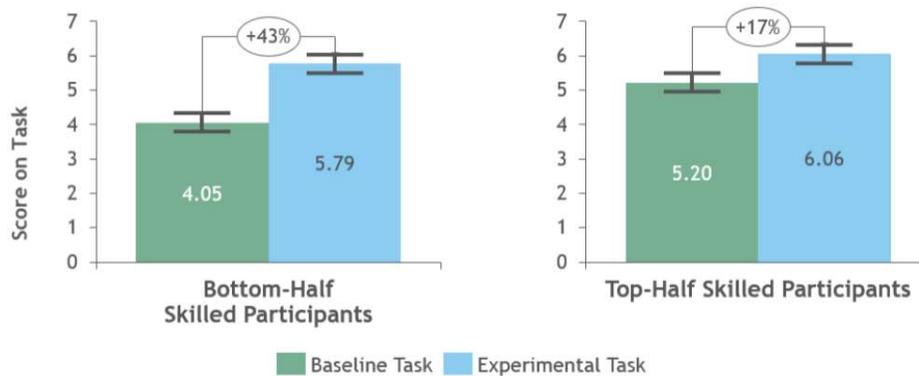
« L'utilisation de l'IA [Chat-GPT4] a entraîné une augmentation du nombre de sous-tâches complétées de 12,5 % en moyenne, mais elle a également amélioré la qualité des réponses de plus de 40 % en moyenne. »

Source: Dell'Acqua *et al.*, 2023.

13

« Inside the frontier »

Les sujets les moins compétents ont enregistré l'augmentation la plus importante de leurs performances (43 %), contre 17 % pour les sujets les plus compétents.



Source: Dell'Acqua *et al.*, 2023.

14

« Outside the frontier »

Exactitude des réponses :

- Groupe « sans IA » : 84.5%
- Groupe « avec IA » : 71.5%
- Groupe « avec IA et guide » : 60.5%

Qualité des réponses fournies :

- Groupe « sans IA » : 5.86
- Groupe « avec IA » : 6.91 (+18%)
- Groupe « avec IA et guide » : 7.33 (+25%)

« L'usage de l'IA réduit significativement l'exactitude des recommandations concernant la résolution d'un problème d'entreprise, mais la qualité a augmenté, que la recommandation sous-jacente soit correcte ou non. »

Source: Dell'Acqua *et al.*, 2023.

15

LLM dans le conseil en stratégie : Conclusions

- Pour les tâches situées à l'intérieur de la frontière (pertinent), l'IA a considérablement amélioré les performances humaines, surtout auprès des sujets moins performants
- En dehors de cette frontière (non-pertinent), les humains se sont trop fiés à l'IA et ont été plus enclins à commettre des erreurs
- Cette étude souligne l'importance de valider et de questionner l'IA en exerçant un effort cognitif continu et en soumettant les résultats au jugement d'experts

Source: Dell'Acqua *et al.*, 2023.

16

Risks and Problems

- AI Hallucinations
- Mis- & Disinformation
- Privacy Concerns
- Biased Information
- Copyright and IP rights
- Energy Consumption
- Information security
- Mental health



17

AI “hallucinations”

“LLMs are designed to produce convincing responses without any guarantees regarding their accuracy. [...] People using LLMs often anthropomorphize the technology, where they trust it as a human-like information source. [...] users can easily be convinced that responses are accurate, even when they have a biased or partial version of the truth.”¹

Prof. Brent Mittelstadt (Oxford Internet Institute)

¹ Source: Mittelstadt B., Wachter S., Russel C., *To protect science, we must use LLMs as zero-shot translators*, Nature Human Behavior, 7., 2023.



18



Vous

C'est quoi l'ISSeP ?



ChatGPT

L'ISSeP, ou Institut Supérieur de l'État de la Province de Namur, est une institution d'enseignement supérieur belge située à Namur, en Belgique. Il propose des formations dans les domaines de l'ingénierie, de l'environnement, de l'architecture, de la construction, de l'informatique et des sciences de la santé. L'ISSeP est reconnu pour son expertise dans les domaines de la gestion de l'eau, de l'énergie et de l'environnement.



¹ Requête réalisée sur ChatGPT3.5, le 14/05/2024.

19

Privacy and Disinformation

“Approximately 11% of the information shared with ChatGPT by employees consists of sensitive data, including client information. [...] Companies such as Samsung, Amazon and JPMorgan have reportedly barred employees from uploading confidential information, including code, to LLMs.”¹

Researchers Poke Holes in Safety Controls of ChatGPT and Other Chatbots

A new report indicates that the guardrails for widely used chatbots can be thwarted, leading to an increasingly unpredictable environment for the technology.

FORBES > INNOVATION > VENTURE CAPITAL

The Incoming Tidal Wave Of Data Pollution In AI

Health IT - 2 Min Read

Researchers find AI can be used to generate volumes of misleading content on critical health topics, call for vigilance

¹ Source: Forbes, *10 Ways Cybercriminals Can Abuse Large Language Models*, M. Drollet (30 June 2023).

20

AI bias... 'garbage in, garbage out'

"The training data used to develop LLMs are often biased because they're based on humanity's past written records. [...] The situation is compounded by LLMs being trained on data from the Internet, where more men than women spend time."¹

Alex Hanna (Director of research)

"Do we have to reproduce society as it is? Or do we have to represent society as we would like it to be? But then, according to whose imagination?" ”

- Giada Pistilli -
Principal Ethicist at Hugging Face

NOVEMBER 22, 2023 | 3 MIN READ

ChatGPT Replicates Gender Bias in Recommendation Letters

A new study has found that the use of AI tools such as ChatGPT in the workplace entrenches biased language based on gender

¹ Source: Scientific American, *ChatGPT Replicates Gender Bias in Recommendation Letters* (22 Nov. 2023).

21

Copyright infringement

"AI art generators are trained on enormous datasets, containing millions upon millions of copyrighted images, harvested without their creator's knowledge, let alone compensation or consent. This is effectively the **greatest art heist in history**. It's daylight robbery."¹

Molly Crabapple (artist)

AI art tools Stable Diffusion and Midjourney targeted with copyright lawsuit



/ The suit claim tools violate cc scraping artist: web without tj

NEWS

OpenAI and Microsoft face new copyright lawsuit over AI training



Published 5 days ago on November 23, 2023
By Chuks Nnabuanyi Jr

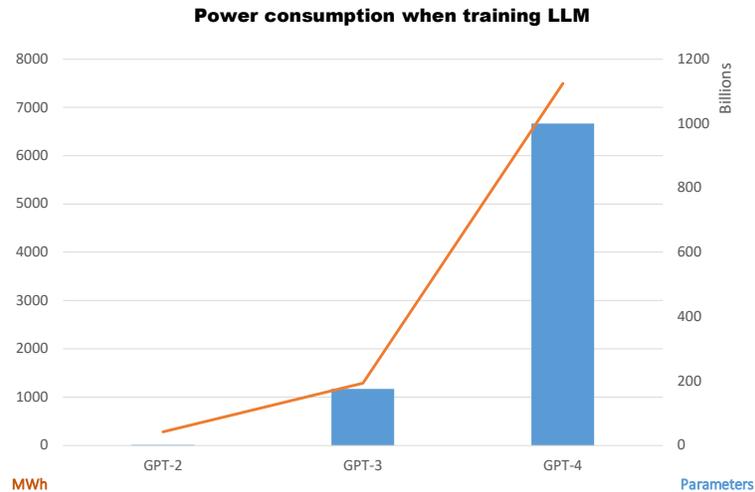
¹ Source: The Guardian, *AI machines aren't 'hallucinating'. But their makers are*, N. Klein (8 May 2023).

22

Energy Consumption

Training GPT-4 is comparable to the annual energy consumption of around 700 households in the United States.

The estimated energy consumption of a single query is 0.01kWh.



¹ Source: Nature Communications / Statista, 2023.

23

Que faire en tant que manager ou leader d'organisation ?

STEP 1 >

Inventaire rapide des fonctions reposant davantage sur le savoir que le savoir-faire (« knowledge jobs ») au sein de l'organisation

STEP 2 >

Pour chacune des fonctions retenues, se poser la question suivante : « L'ajout d'un **assistant compétent mais naïf** pour ce job serait-il un plus ? »

STEP 3 >

Prioriser l'implémentation en commençant par les fonctions pour lesquelles les gains de productivité sont évidents (tâches assistées par l'IA générative)

¹ Source: McAfee A., Rock D., Brynjolfsson E., (Nov. 2023), Harvard Business Review.

24

