

*Un livre qui rend simple
ce que les autres rendent compliqué*

ChatGPT

POUR
LES PROS

***Tout comprendre
En 45 min.***

***ChatGPT
expliqué aux
aux dirigeants***



**CLOUD BUSINESS
School**

Jean-François MESSIER



BIENVENUE

ChatGPT est l'application dont tout le monde parle et qui a atteint plus d'un milliard d'utilisateurs en 11 mois.

Problème : ChatGPT ou ChatGPT Turbo, LLM, prompts, prompt engineering, renforcement, custom instructions, BARD, RLHF, tokens, corpus.... Il y a vraiment de quoi ne rien comprendre, se sentir dépassé, ou avoir le sentiment de ne pas connaître autant qu'il le faudrait.

Si tel est votre cas, si vous ne maîtrisez pas tous les termes qui précèdent ou que vous avez des doutes, investissez 20 min à lire ce qui suit : vous allez tout comprendre !

Spécifiquement si vous êtes dirigeant : vous ne pouvez passer à côté de cette révolution :

- Ces applications dites d'IA générative, malgré leurs défauts actuels, vont plus changer le monde dans les 3 ans à venir que ne l'a fait Internet + les PC + les mobiles cumulés.... Depuis 30 ans.
- Les champs d'application sont colossaux, ils concernent toutes les entreprises, ils permettent dès maintenant des sauts de performance colossaux, et/ou des économies non négligeables.

Bienvenue dans le 21° siècle !

Ex-Directeur des Nouvelles Technologies chez Mercuri International, Enseignant à l'ESCP Europe, auteur de nombreux programmes de formations sur les nouvelles technos et la vente....

En 20 ans Jean-François a accumulé une expérience riche, auprès de 900 entreprises de 22 pays.

Les plus gros sont des géants mondiaux et nombre d'entreprises du CAC 40, les plus petits font moins d'un million d'euros de CA, le plus petit a 5 salariés.

10 000

Participants en conférences

20 000

Commerciaux, managers et dirigeants

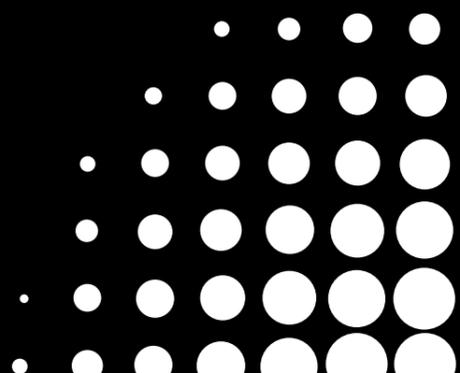
1 000

Entreprises accompagnées

CA 191 MILLIARDS € SAMSUNG Fabricant de produits électroniques 2011, 12	CA 81,8 MILLIARDS € Nestlé Produits parpharmacie de Nestlé Health Science 1999	CA 81,6 MILLIARDS € Johnson-Johnson Matériel médical, produits d'hygiène, cosmétiques 1998, 1999, 2000	CA 76,0 MILLIARDS € Carrefour Grande distribution. Division achats Belgique 2007, 08, 09	CA 68,7 MILLIARDS € ArcelorMittal N°1 mondial de la sidérurgie 1998, 99	CA 62,7 MILLIARDS € BASF The Chemical Company Plus grand groupe de chimie au monde 1996, 97, 98
CA 57,9 MILLIARDS € AEGON Compagnie d'assurances 2007, 08, 09	CA 52,1 MILLIARDS € Allergan Produits d'ophtalmologie, neurologie et dermatologie 1998, 99	CA 51,9 MILLIARDS € NOVARTIS Groupe pharmaceutique suisse 1998, 99	CA 51,6 MILLIARDS € Pfizer Groupe pharmaceutique américain 1998, 99	CA 50,0 MILLIARDS € hp Fabricant d'ordinateurs, serveurs, imprimantes 2005, 06, 07, 08	CA 41,0 MILLIARDS € SAINT-GOBAIN Transformation et distribution de matériaux 1996, 97, 98, 2018
CA 40,2 MILLIARDS € orange Opérateur téléphonique 1999, 2000, 13, 14, 18	CA 39,6 MILLIARDS € BAYER Bayer Diagnostic, fabricant de matériel médical 1997, 98, 99	CA 34,5 MILLIARDS € SANOFI Groupe pharmaceutique, leader français de la santé 1998, 99, 2000, 18	CA 31 MILLIARDS € essilor Verres et équipements d'optique ophtalmique 2004	CA 30,6 MILLIARDS € TOSHIBA Fabricant de matériel électronique et informatique 2016, 17, 18, 19	CA 25,7 MILLIARDS € Schneider Electric Produits de gestion de l'électricité et automatisés 2006, 07, 08, 09
CA 24,7 MILLIARDS € LA POSTE Services postaux, banque, assurance, marketing 2008, 09	CA 24,5 MILLIARDS € AXA Compagnie d'assurances 2006, 07, 08, 09	CA 22,6 MILLIARDS € Bristol-Myers Squibb Divulsième groupe pharmaceutique mondial 1999, 2000	CA 21,0 MILLIARDS € Air Liquide Gaz, technologies et services pour l'industrie et la santé 2019	CA 20,7 MILLIARDS € KUEHNE + NAGEL Transport et logistique 2012, 13, 18	CA 18,7 MILLIARDS € Henkel Détergents, produits entretiens, colles et adhésifs 2008
CA 18,2 MILLIARDS € PHILIPS Éclairage, électroménager, équipement médical 2012	CA 17,7 MILLIARDS € ING Banque Assurances 2012	CA 15,9 MILLIARDS € BD Fabricant de matériel médical 2002, 05	CA 8,7 MILLIARDS € Pernod Ricard Industrie des vins et spiritueux 2012, 13, 14	CA 6,3 MILLIARDS € Alcon Leader mondial en produits de soins oculaires 1999, 2000	CA 6,1 MILLIARDS € DACHSER Intelligent Logistics Logistique — Transport 2017, 18, 19
CA 5,8 MILLIARDS € ZEISS Fabricants verres et matériel d'optique de précision 2017, 18, 19	CA 5,7 MILLIARDS € QUEBECOR Médias, journaux, TV, radio, Internet, téléphonie 2007, 08, 09	CA 5,5 MILLIARDS € Groupe CASTEL Producteur et distributeur de vins français 2016, 17, 18, 19	CA 5,5 MILLIARDS € MICHELIN Fabricant de pneumatiques 2006, 07, 08, 09, 10, 11	CA 5,5 MILLIARDS € PANALPINA Logistique — Transport 2012	CA 5,3 MILLIARDS € PIRELLI Fabricant de pneumatiques 2010, 13, 14
CA 5,1 MILLIARDS € RHENUS LOGISTICS Logistique — Transport 2015	CA 5,0 MILLIARDS € KNAUF Matériaux d'isolation thermique et phonique 1996, 97	CA 5,0 MILLIARDS € NYCOMED Groupe pharmaceutique passé sous Talca en 2011 1998, 99	CA 5,0 MILLIARDS € CIC Banque, assurances 2013	CA 4,9 MILLIARDS € Bausch+Lomb Fabricant ophtalmologie et matériel médical optique 1997, 98	CA 3,5 MILLIARDS € GRDF Filiale d'Engie, distributeur de gaz naturel 2018
CA 3,4 MILLIARDS € LEXMARK Fabricant d'imprimantes 2004, 2010	CA 3,2 MILLIARDS € FESTO Leader mondial en technique d'automatisation 2018, 19	CA 3,1 MILLIARDS € Wienerberger Fabrication de briques, tuiles et bardage 2015, 16	CA 3,1 MILLIARDS € GRUNDOS Spécialiste dans les pompes hydrauliques, chauffage, etc. 1999, 2000, 01	CA 2,4 MILLIARDS € REXEL Distribution de matériel électrique, chauffage, etc. 1996-97, 2006, 07, 08	CA 2,4 MILLIARDS € COOEDIM Construction, extension et promotion immobilière 2006
CA 2,2 MILLIARDS € sonepar Distribution de matériel électrique, chauffage, etc. 1997, 98, 2002	CA 2,2 MILLIARDS € IPSEN Groupe biopharmaceutique en oncologie, neurosciences... 2000	CA 2,0 MILLIARDS € hager Fabricant de matériel électrique 2014	CA 1,7 MILLIARD € SCA Produits d'hygiène, couches, mouchoirs, protège-slips 1999, 2000, 01, 03	CA 1,6 MILLIARD € J&S Fret d'entreprise, logistique et support à la supply chain 2018	CA 1,1 MILLIARD € MAISONS DU MONDE Enseigne de 250 magasins aménagement et décoration 2014, 15
CA 1,0 MILLIARD € +GF+ AgieCharmilles Machines pour la fabrication de moules et d'outils... 2013	CA 938 MILLIONS € GEODIS Transport logistique 2013, 14, 15	CA 862 MILLIONS € MERIAL Leader mondial en santé animale 1999, 2000	CA 726 MILLIONS € Butagaz Fournisseur d'énergie pour particuliers et pros 2000	CA 610 MILLIONS € BWT Fabricant de systèmes de traitement de l'eau 2017, 18	CA 576 MILLIONS € Ceva Laboratoire vétérinaire 2007, 08, 09
CA 567 MILLIONS € CATALINA Service de marketing internet 2013, 14, 15	CA 520 MILLIONS € EUROMASTER Montage pneus et entretien courant automobile 1999, 2000, 01	CA 507 MILLIONS € LDL n°1 du high-tech et du matériel informatique 2018	CA 488 MILLIONS € MANE Un des leaders mondiaux en flagrance et parfums 2016, 17	CA 465 MILLIONS € PUM PLASTIQUES Négoce produits plastiques à destination du BTP 1998, 99	CA 450 MILLIONS € YVES ROCHER Fabrication et distribution de produits cosmétiques 2012, 13, 14
CA 311 MILLIONS € Mutualia Mutuelle santé 2018, 19	CA 300 MILLIONS € Aprolis Spécialiste de la location de matériels de manutention 2018, 2019	CA 255 MILLIONS € Insight Distribution de matériel informatique, logiciels... 2018	CA 249 MILLIONS € Valeo Fabricant d'accessoires automobile 2e monte 2009, 12, 13	CA 119 MILLIONS € Reed Exhibitions Organisation de foires et salons professionnels 1992, 93, 94, 95, 2018, 19	CA 20 MILLIONS € Solocal Publicité et marketing numérique 2015

SOMMAIRE

1. <u>Qu'est-ce que ChatGPT ?</u>	4
▪ Qu'est-ce que l'intelligence artificielle ?	4
▪ Si ce n'est de l'intelligence artificielle, c'est quoi ?	5
2. <u>Qu'est-ce que les LLM — modèles de langage ?</u>	6
▪ Les modèles autorégressifs : celui de ChatGPT.	6
▪ Les modèles de langage bidirectionnels.	7
▪ Les modèles de langage mixtes	7
3. <u>Comment fonctionne ChatGPT ?</u>	8
▪ 0 % compréhension, 100 % « calculation »	8
▪ Qu'est-ce que la température ?	11
4. <u>Comprendre les tokens</u>	13
▪ Pourquoi vous devez comprendre les tokens	13
▪ Correspondance entre Token — mots — caractères	14
▪ Comprendre et visualiser les tokens	14
▪ Pourquoi mesurer les tokens plutôt que les mots ou caractères ?	15
5. <u>La base de données d'information de ChatGPT</u>	16
▪ Le corpus d'informations	16
▪ Limitations du corpus d'informations de ChatGPT	16
▪ Le renforcement	18
6. <u>Prompts et prompt engineering.</u>	21
▪ Les 7 couches obligatoires d'un prompt	22
▪ Custom Instructions pour les prompts récurrents	24
▪ GPTs pour les prompts récurrents	26
▪ Éviter les 4 pièges les plus fréquents	27
<u>Conclusion</u>	29



1 QU'EST-CE QUE CHATGPT ?

Une application dite d'« intelligence artificielle » basée sur un « modèle de langage ».

Voici dès la première phrase 2 concepts : IA pour « intelligence artificielle » et « modèle de langage »... accrochez la ceinture, c'est parti !

1. Qu'est-ce que l'intelligence artificielle ?

L'intelligence artificielle (IA) consiste à reproduire à l'aide de machines, des activités mentales humaines, qu'elles soient de l'ordre, de la compréhension, de la perception, ou de la décision. Par là même, l'IA est distincte de l'informatique, qui traite, trie et stocke les données et leurs algorithmes.

C'est pour cette raison que la formulation « Intelligence artificielle » est largement usurpée pour des logiciels comme ChatGPT et autres IA dites « générative ». (les IA génératives sont une catégorie d'IA qui se concentre sur la génération de code informatique, de texte, d'images, etc.).

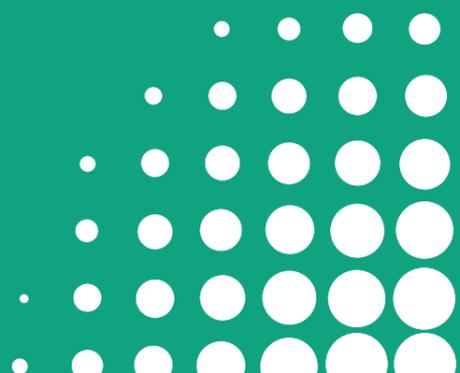
En effet ces applications ne sont rien de plus que des applications informatiques sans aucune capacité à réfléchir, comprendre, ni même utiliser leurs connaissances spontanément. Voir nos articles sur ces sujets :

- [Décryptage et exemples amusants des limites cachées de ChatGPT !](#)
- [On nous aurait menti ? L'intelligence artificielle n'existe pas.](#)

2. Alors si ce n'est de l'intelligence artificielle, c'est quoi ?

ChatGPT est une application informatique qui traite des données avec des structures de contrôle conditionnelles (If... then... Else.... Loop...), comme le faisaient les SGBD (système de gestions de base de données) dès les années 60.

Le changement n'est pas une révolution technologique, mais un changement d'échelle : aujourd'hui la puissance est vertigineuse : les données sont du BigData (l'intégralité du web et plus), le nombre de paramètres est colossal (175 Milliards en janvier 2024). Ce qui confère un potentiel colossal à ce type d'application basée sur un modèle de langage...



2 QU'EST-CE QUE LES MODÈLES DE LANGAGE LLM ?

ChatGPT est basé sur un modèle de langage capable de générer du contenu (texte, code, images, etc.). On parle de LLM pour « large language model » en anglais. Ce sont des modèles de langage de grande taille, possédant un grand nombre de paramètres (généralement de l'ordre du milliard ou plus).

En réalité il existe une multitude de modèles de langage :

1. Les modèles autorégressifs : ChatGPT.

Les modèles de langage autorégressifs génèrent du texte en prédisant le mot suivant en calculant la probabilité maximum d'occurrence basée sur ceux qui viennent avant.

C'est à cette catégorie qu'appartient ChatGPT. Il gère à l'heure où j'écris cet article 175 milliards de paramètres.

Mais il en existe d'autres : Transformer-XL, CTRL de Salesforce (Conditional Transformer Language Model), etc.

2. Les modèles de langage bidirectionnels.

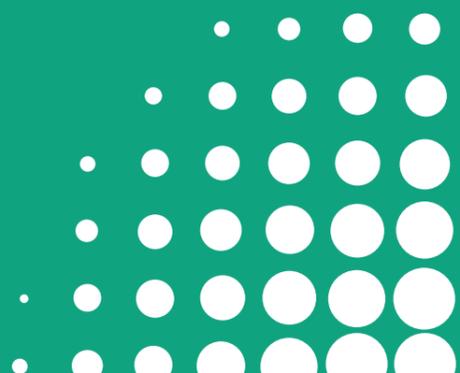
Ces modèles se distinguent par leur capacité à comprendre le contexte des mots dans les deux sens, ce qui améliore significativement la précision de la compréhension du langage naturel par rapport aux modèles unidirectionnels. Exemples : BERT de Google (Bidirectional Encoder Representations from Transformers), XLNet, ERNIE de Baidu (Enhanced Representation through Knowledge Integration), RoBERTa de Facebook (Robustly Optimized BERT Approach).

3. Les modèles de langage mixtes

À la fois Autorégressifs et Bidirectionnel : T5 (Text-to-Text Transfer Transformer) de Google, BART (Bidirectional and Auto-Regressive Transformers) de Facebook, Electra de Google

Et plein d'autres...

- des modèles spécialisés pour les dialogues – DialoGPT, Meena -,
- des modèles pour des tâches spécifiques,
- des modèles spécialisés en classification
- etc.



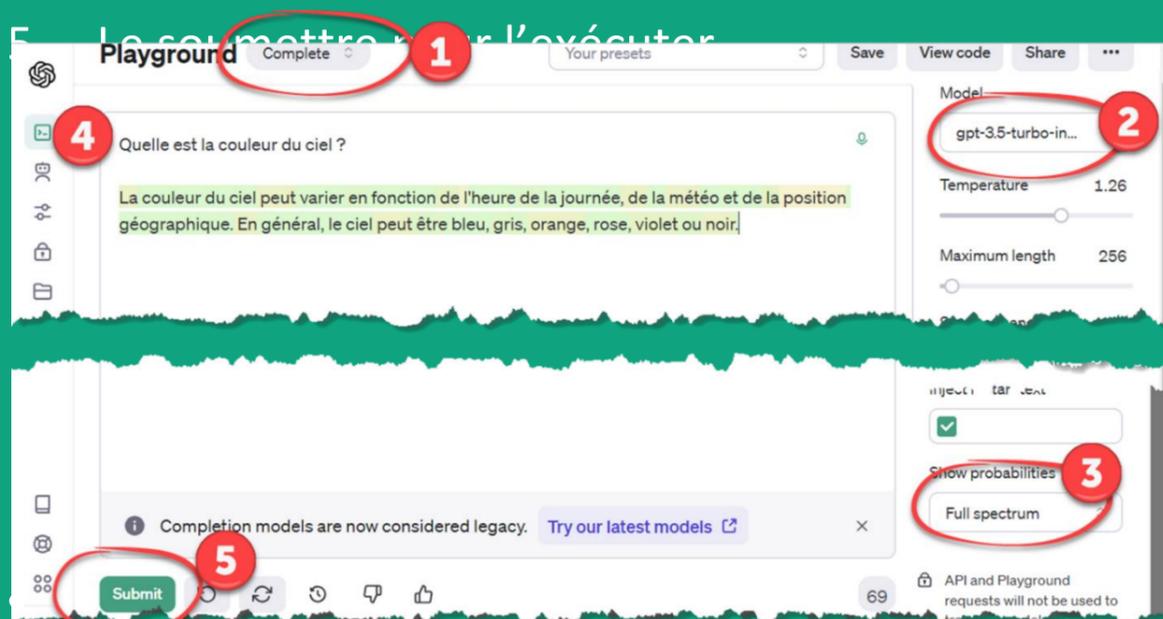
3 COMMENT FONCTIONNE CHATGPT ?

Vous l'avez compris, ChatGPT est basé sur un LLM (Large Language Model) autorégressif : cela signifie juste qu'il génère du texte en prédisant la probabilité maximum du mot suivant en fonction des mots précédents.

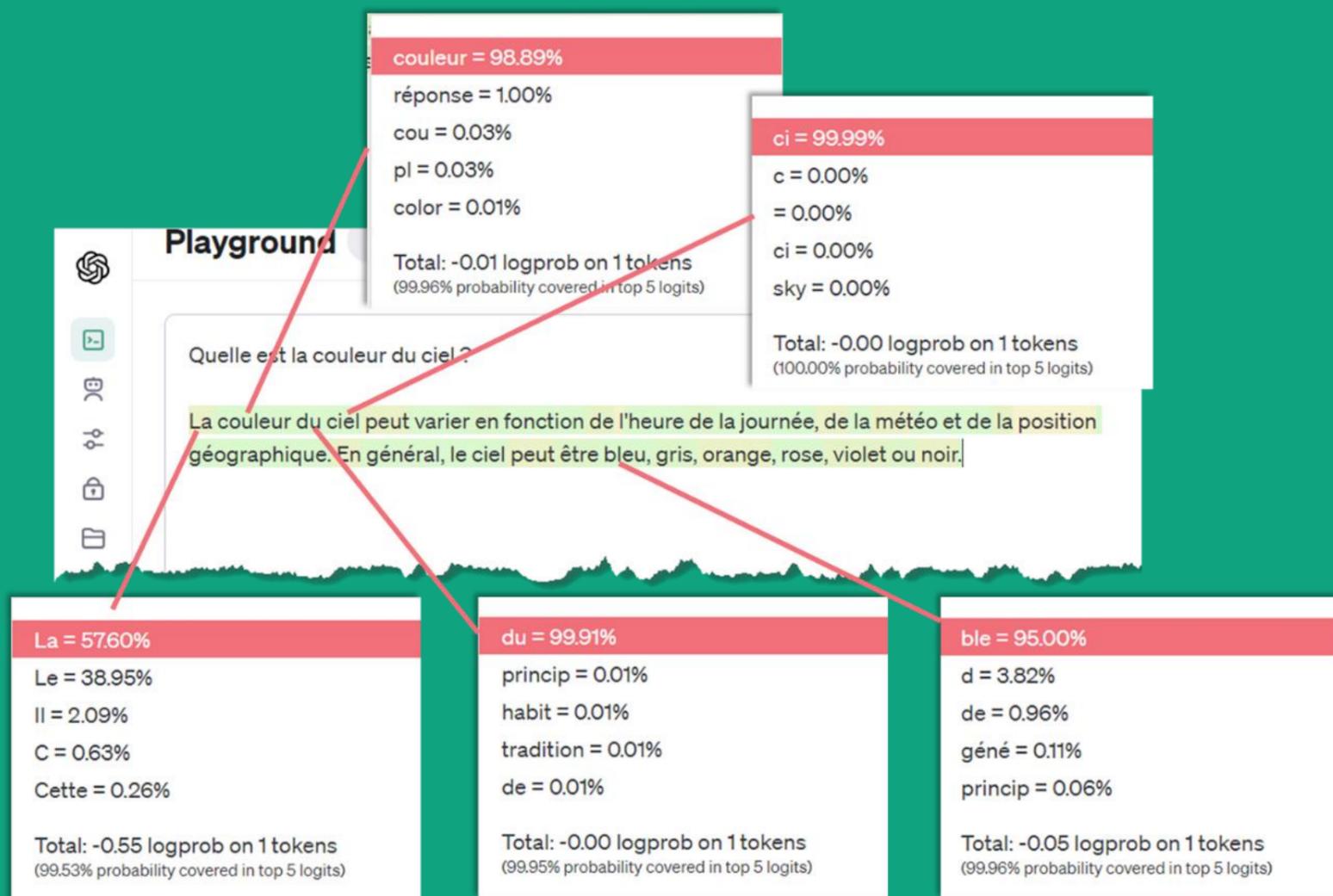
1. 0 % compréhension, 100 % « calculation »

Le playground d'OpenAI facilite la compréhension : vous devez créer un compte pour y avoir accès. Voici le setup pour comprendre le fonctionnement de ChatGPT :

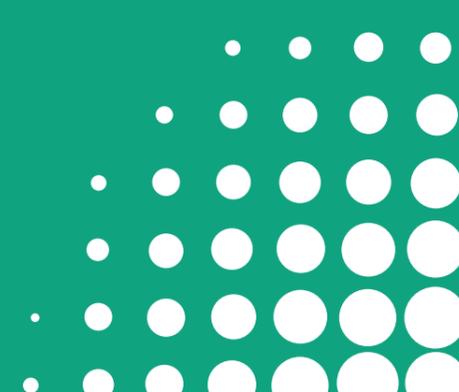
1. Sélectionner le mode « complete »
2. Sélectionner le dernier modèle de langage disponible
3. Opter pour l'affichage des probabilités « Full Spectrum »
4. Saisir un prompt



Cette interface montre bien le mode de calcul de la probabilité du mot suivant en fonction des précédents, prompt compris ; au survol de chaque mot apparaît une fenêtre avec la probabilité d'occurrence qui se présente pour le mot suivant ; et on voit que ChatGPT opte à chaque fois pour la probabilité la plus élevée.



La couleur derrière chaque mot indique la probabilité du mot : le vert pour les hautes probabilités, le rouge pour les faibles probabilités, et l'orange pour les probabilités moyennes.



Ce qui met en évidence trois réalités importantes à retenir :

1. **Quand ChatGPT commence à répondre, il n'a aucune idée de ce que sera la réponse !**
La réponse sera construite au fur et à mesure, mot après mot. Ce qui signifie que ChatGPT n'a aucune conscience de la réponse qu'il construit, et donc aucune compréhension réelle de la question qui lui a été posée.
2. C'est une autre manière de comprendre ce que nous expliquons dans plusieurs numéros de cette news : **ChatGPT ne raisonne pas, ChatGPT ne comprend pas, Chat GPT n'a rien à voir avec l'intelligence.** C'est une application qui applique bêtement un algorithme calculant une probabilité de résultat.
3. ChatGPT est donc programmé pour deviner les mots les plus probables, **pas nécessairement les plus exacts.** Étant donné la vaste étendue d'informations disponibles en ligne : contenus fictifs, humoristiques ou tout simplement erronés ; on comprend aisément pourquoi ChatGPT se trompe régulièrement et fait des erreurs.

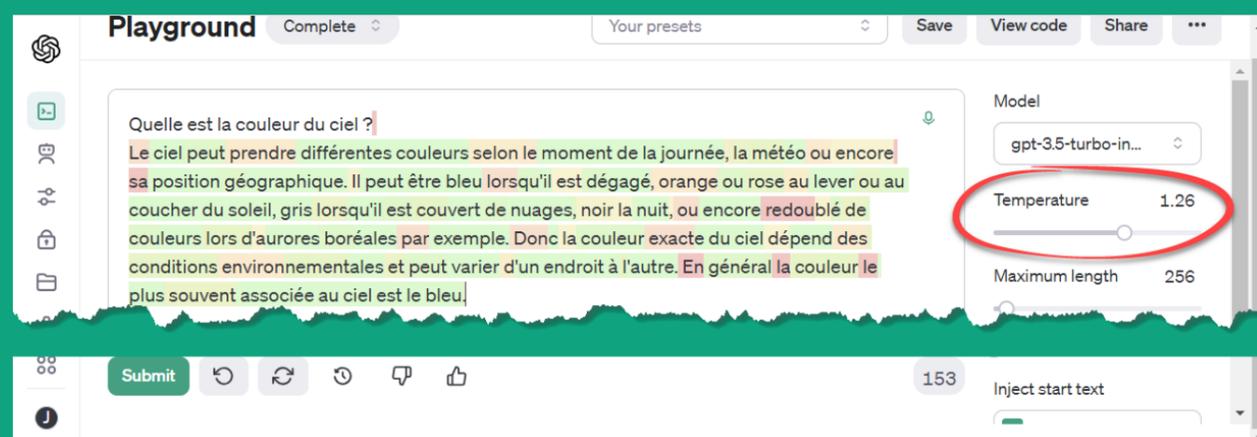
Toutefois il est possible sur demande de faire opter l'application pour des choix de mots qui n'ont pas la plus forte probabilité d'occurrence : cela dépend du réglage de la « **température** ».

2. Qu'est-ce que la température ?

La température est un paramètre ajustable de 0 à 2 : on peut le préciser dans un prompt, ou l'ajuster par un curseur dans le playground de ChatGPT. À une température proche de 0, le modèle tend à choisir les mots à plus forte probabilité d'occurrence. Cela donne des résultats plus déterministes et souvent plus cohérents. À l'inverse, une température proche de 2 pousse le modèle à prendre des libertés, explorant des choix des mots plus improbables, introduisant plus de variété dans les résultats.

Le test précédent page 9 avait été réalisé avec une température réglée sur 0 : dans cette option, ChatGPT optait systématiquement pour la probabilité la plus élevée, et on obtenait le résultat le plus conforme au corpus d'information de ChatGPT. La couleur à l'arrière-plan de chaque mot était principalement verte.

Réglons maintenant la température sur 1,26 : au premier coup d'œil, il apparaît que la couleur change : il y a plus de rouge et orange, mettant en évidence le fait que ChatGPT opte pour les mots aux probabilités d'occurrence plus faibles.



Les mots utilisés sont moins prévisibles, moins conventionnels.

On peut observer qu'avec une température supérieure à 1, la probabilité retenue n'est plus systématiquement la plus élevée. La réponse va s'éloigner de la conformité au corpus d'information initial.

The screenshot displays a language model interface with several components:

- Token Probability Lists:**
 - Top-left: être = 53.00%, avoir = 25.47%, prendre = 11.24% (highlighted), app = 6.20%, var = 1.31%. Total: -2.19 logprob on 1 tokens (97.21% probability covered in top 5 logits).
 - Top-right: en = 67.65%, selon = 25.35% (highlighted), , = 6.01%, := 0.22%, au = 0.16%. Total: -1.37 logprob on 1 tokens (99.39% probability covered in top 5 logits).
 - Bottom-left: la = 86.43%, l = 9.46%, les = 1.61%, le = 1.49%, selon = 0.47%, sa = 0.24% (highlighted). Total: -6.05 logprob on 1 tokens (99.70% probability covered in top 6 logits).
- Model Output:** "Quelle est la couleur du ciel ?" followed by a paragraph of text: "Le ciel peut prendre différentes couleurs selon le moment de la journée, la météo ou encore sa position géographique. Il peut être bleu lorsqu'il est dégagé, orange ou rose au lever ou au coucher du soleil, gris lorsqu'il est couvert de nuages, noir la nuit, ou encore redoublé de couleurs lors d'aurores boréales par exemple. Donc la couleur exacte du ciel dépend des conditions environnementales et peut varier d'un endroit à l'autre. En général la couleur le plus souvent associée au..."
- Model Settings:** Model: gpt-3.5-turbo-in..., Temperature: 1.26 (circled in red), Maximum length: 256.
- Interface Elements:** "Submit" button, "Inject start text" field, and a "153" token count.

4 COMPRENDRE LES TOKENS

Le token est l'unité de mesure qui mesure la quantité de texte que ChatGPT peut traiter en une seule fois : la mesure prend en compte les caractères, les espaces, la ponctuation. Un token est plus qu'un caractère, moins qu'un mot le plus souvent, genre une syllabe, mais les découpages ne correspondent pas au découpage de syllabes. Par exemple :

- **Environnemental** = 3 tokens :
Environ – nemen – tal
- **Géographique** = 3 tokens :
Gé - ograph – ique

Attention : parfois un mot de plusieurs syllabes peut n'être associé qu'à un token :

- **Différent** = 1 token :
Différent

1. Pourquoi c'est important ?

Il est important de comprendre les tokens, car ChatGPT est limité en capacité : il ne peut traiter qu'un nombre de tokens limité au cumul des instructions et de la réponse. Par exemple, un prompt de 50 tokens suivis d'une réponse de 150 tokens consomme un total de 200 tokens.

Voici les limites actuelles en janvier 2024 :

Modèle	Limite de Tokens	Équivalent # mots approximativement.	
		français	anglais
ChatGPT-3.5	4 096 tokens	2 800 à 3 000	3 000 à 3 300
ChatGPT-4	8 192 tokens	5 700 à 6 100	6 100 à 6 500
ChatGPT-4 Turbo spé	128 000 tokens	89 000 à 96 000	96 000 à 120 000
API	4 096 tokens	2 800 à 3 000	3 000 à 3 300

2. Correspondance tokens — mots — caractères

La correspondance entre # tokens et # mots dépend de la langue : en anglais, un token correspond généralement à un peu moins d'un mot, car l'anglais a tendance à avoir des mots plus courts. En français, les mots sont souvent plus longs, donc le nombre de mots par 100 tokens peut être légèrement inférieur.

Quant à la correspondance entre token et caractères, les mots français étant généralement plus longs que les mots anglais, 100 tokens en français peuvent représenter plus de caractères que la même quantité de tokens en anglais.

Langue	Tokens	Mots	Caractères	Mots	Tokens	Caractères
Anglais	100	75-80	500-600	100	125-135	600-700
Français		70-75	600-700		130-140	750-850

3. Comprendre et visualiser les tokens

Le **tokenizer d'OpenAI** permet de comprendre facilement le concept de tokens. Rentrez un texte dans la zone de dialogue, et vous obtenez :

- le nombre de tokens,
- le nombre de caractères.
- Une visualisation des tokens par code couleur.

GPT-3.5 & GPT-4 GPT-3 (Legacy)

Quelle est la couleur du ciel ?

Le ciel peut prendre différentes couleurs selon le moment de la journée, la météo ou encore sa position géographique. Il peut être bleu lorsqu'il est dégagé, orange ou rose au lever ou au coucher du soleil, gris lorsqu'il est couvert de nuages, noir la nuit, ou encore redoublé de couleurs lors d'aurores boréales par exemple. Donc la couleur exacte du ciel dépend des conditions environnementales et peut varier d'un endroit à l'autre. En général la couleur le plus souvent associée au ciel est le bleu.

Clear Show example

Tokens	Characters
144	537

Quelle est la couleur du ciel ?
 Le ciel peut prendre différentes couleurs selon le moment de la journée,
 la météo ou encore sa position géographique. Il peut être bleu lorsqu'il
 est dégagé, orange ou rose au lever ou au coucher du soleil, gris lorsqu'
 'il est couvert de nuages, noir la nuit, ou encore redoublé de couleurs
 lors d'aurores boréales par exemple. Donc la couleur exacte du ciel dé
 pend des conditions environnementales et peut varier d'un endroit à l'aut
 re. En général la couleur le plus souvent associée au ciel est le bleu.

TEXT TOKEN IDS

Les codes couleur permettent de bien visualiser les « tokens ».

3. Pourquoi ChatGPT mesure les tokens plutôt que les mots ou caractères ?

Les tokens permettent au LLM de mieux gérer les nuances linguistiques que s'il gérait des mots. Plutôt que d'avoir un token unique par mot, le mot est décomposé en tokens plus petits permettant au modèle d'apprendre séparément la signification et l'utilisation de ces sous-éléments. Cela permet de gérer de nombreux mots et leurs variations sans avoir besoin d'un vocabulaire immense.

Et un modèle basé sur les tokens aura plus de repères face à un mot inconnu que s'il était basé sur des caractères : il peut décomposer ce mot en tokens qu'il connaît, pour en déduire une signification ou, au moins, le traiter de manière cohérente, beaucoup plus rapidement que s'il devait le faire caractère par caractère.

5 LA BASE DE DONNÉES D'INFORMATIONS DE CHATGPT

La base de données d'information de ChatGPT provient de son entraînement sur un large éventail de textes disponibles : des livres, des extraits de livres, des articles, des sites web, et d'autres types de sources pour couvrir une vaste gamme de sujets.

1. Qu'appelle-t-on corpus d'informations

Cette accumulation d'informations nommée « **corpus d'information** » est stockée dans le modèle, lui permettant d'avoir accès à de nombreux sujets.

2. Limitations du corpus d'informations de ChatGPT

Il est important d'être conscient que le corpus d'informations dont dispose ChatGPT n'est pas à jour, que de nombreuses informations font défaut, et qu'il contient beaucoup d'erreurs !

❑ Les informations du corpus ne sont pas à jour :

- Celles de GPT-3-5 s'arrêtent en décembre 2021. Celles de ChatGPT-4 s'arrêtent en avril 2023.
- Les 2 modèles de ChatGPT ne possèdent pas d'informations sur des événements ou des développements postérieurs à ces dates.
- ChatGPT-3 est gratuit, mais n'a aucune possibilité d'accès à internet : cela signifie qu'il ne peut pas récupérer ou fournir des informations actualisées après décembre 2021.

- Pour GPT-4, il est possible d'incorporer dans les prompts une instruction demandant à l'application d'aller chercher sur le web des informations en temps réel avant de répondre, soit par des plugins gratuits, soit par la fonction Browse With Bing. (les 2 sont inclus gratuitement dans GPT-4)

- ❑ **Les informations dont dispose ChatGPT sont régulièrement fausses :** ChatGPT intègre dans son corpus le web et donc toutes les erreurs et/ou données obsolètes qui circulent sur le web. En conséquence de quoi il va régulièrement fournir des informations incorrectes ou obsolètes et ne devrait pas être utilisé sans vérification supplémentaire.
- ❑ **Les informations dont dispose ChatGPT contiennent de nombreuses lacunes :** Il est important de noter que ChatGPT ne « lit » pas les textes de la manière dont un humain le ferait. Plutôt, il analyse les textes pour apprendre des modèles de langage, des structures de phrases, des informations factuelles, et des styles d'écriture. Cette analyse est effectuée sur des extraits de textes plutôt que sur des œuvres complètes, en raison de contraintes légales et éthiques, ainsi que de la nature de l'entraînement des modèles de langage.
- ❑ **Les informations ne s'accompagnent pas d'une véritable compréhension de ces connaissances...** voir l'exemple du prompt 2 utilisé en exemple dans cet article : « [Un escargot met en échec ChatGPT et révèle ses limites.](#) »

Dans cet exemple du paradoxe de l'escargot, on voit la réponse de ChatGPT rephraser le bon mode de calcul, mais l'incapacité à mettre en œuvre cette connaissance dans les lignes suivantes pour fournir la réponse.

❑ **ChatGPT n'utilise pas spontanément sa base de connaissance.** Par exemple :

- Si l'on demande à ChatGPT de chercher dans son corpus l'âge maximum d'un cheval, et l'âge maximum pour pratiquer le CSO, il va répondre 40 ans pour l'âge maximum, et 20 ans pour participer à des CSO.
- Mais si on lui demande un programme d'alimentation pour un cheval de CSO de 50 ans, il répond sans aller vérifier si cela fait sens.

Ces limitations sont autant de raisons expliquant les nombreuses erreurs que les novices génèrent avec ChatGPT : éviter ces réponses absurdes demande d'être conscient de cette faiblesse du modèle et de l'envoyer chercher l'information nécessaire avant de répondre à une question. C'est le rôle du prompt qui sera expliqué plus loin dans cet article.

Et le **renforcement** dont bénéficie ChatGPT est insuffisant pour résoudre ces limites.

3. Qu'est-ce que le renforcement (RLHF)

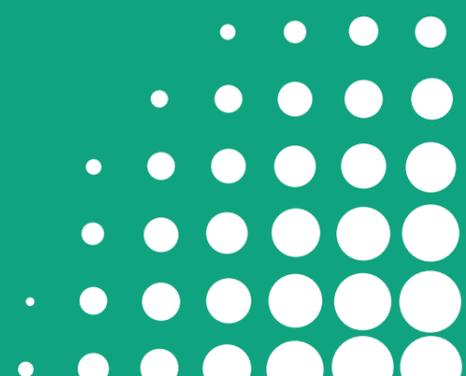
ChatGPT, utilise une approche basée sur l'apprentissage par renforcement à partir de retours humains, aussi connu sous le sigle RLHF (Reinforcement Learning from Human Feedback). Ce dernier s'appuie sur les retours des annotateurs humains pour ajuster les réponses du modèle.

Concrètement des opérateurs humains d'OpenAI sont chargés de classer des sorties proposées par ChatGPT par ordre de pertinence. En collectant un grand nombre de feed-back et en utilisant des systèmes de ranking comme ELO, le procédé améliore considérablement la pertinence de ChatGPT :

- Il permet d'intégrer des connaissances humaines et des préférences subjectives dans le processus d'apprentissage. Cela peut conduire à des modèles plus alignés avec les valeurs humaines et à une meilleure performance dans des tâches complexes.
- Il permet d'éliminer des réponses objectivement fausses ou enfreignant les règles d'éthiques de ChatGPT liées à la sexualité, au racisme, au sexisme, à la protection de la vie privée, à la violence, etc.
- Il est particulièrement utile dans les situations où un arbitrage objectif est difficile ou impossible, comme des tâches impliquant des valeurs humaines.

Pour autant il ne résout pas la multitude des erreurs produites par ChatGPT, pour 2 raisons :

- La limitation des coûts en ressources humaines ne permet de couvrir qu'un pourcentage insuffisant de l'ensemble des sujets qui le nécessiteraient.
- Les annotateurs humains chargés de donner des retours au modèle ont pu valoriser des réponses bien rédigées, n'enfreignant pas les règles éthiques de l'application, sans pour autant que celles-ci soient justes.



D'où toutes ces erreurs et leur fréquence comme illustré dans nos numéros précédents :

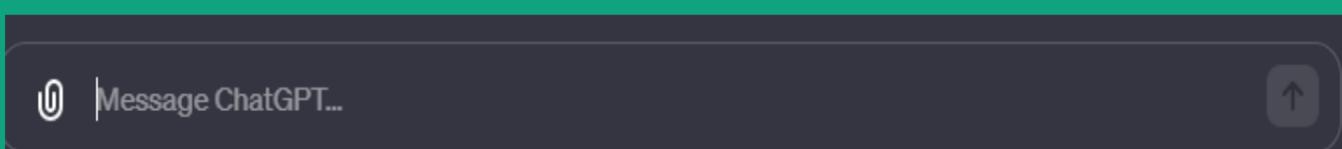
- [Décryptage et exemples amusants des limites cachées de ChatGPT !](#)
- [On nous aurait menti ? L'intelligence artificielle n'existe pas.](#)

Le moyen ultime de se préserver des erreurs de ChatGPT et des réponses inappropriées et la technique utilisée pour lui poser des questions via ce que l'on appelle un **prompt** : Rédiger le prompt demande de maîtriser au minimum les techniques pour guider ChatGPT dans la résolution de la question ou du problème, ainsi que pour contourner ses faiblesses. C'est notre dernier chapitre.

6 COMPRENDRE LE PROMPT ENGINEERING

Jusque là, nous avons vu ce qu'il est important de connaître pour ne pas être « perdu » dans une conversation ou une lecture au sujet de ChatGPT.

Voyons maintenant le plus important pour ce qui est de l'utilisation de ChatGPT dans votre entreprise : En tant qu'utilisateurs « grand public », vous et vos salariés allez passer par l'interface de ChatGPT qui vous demande de saisir un message afin de pouvoir répondre :



Ce message que vous allez saisir en tant qu'utilisateur est le point de départ de la réponse de ChatGPT et il ne doit pas se limiter à poser une question : cela aurait pour conséquence de risquer toutes les erreurs précédemment évoquées. Et ChatGPT allant chercher dans son corpus ou sur les web les premières réponses disponibles, cela ne pourrait qu'aboutir à une réponse d'une banalité atterrante.

1. Les 7 couches obligatoires d'un prompt

En fait ce message doit être une véritable feuille d'instructions destinée non seulement à définir **1** la **Mission** du prompt, mais aussi à **2** guider la réponse en fonction d'un **Objectif** et **3** de **Données** de base, **4** du niveau d'Expertise de la cible à laquelle est destinée la réponse, **5** la forme de la **Réponse** souhaitée, **6** le **Ton** que doit adopter la réponse, et enfin et surtout **7** le **Process** et les instructions que ChatGPT doit suivre pour répondre correctement.

- 1** Mission à exécuter par le prompt : Décrit au prompt la réalisation attendue de ses actions.
- 2** Objectif personnel: Permet à ChatGPT de fournir une réponse adaptée à la finalité.
- 3** Données de base : Permet à ChatGPT de fournir une réponse adaptée à vous, votre entreprise, votre offre, votre positionnement, votre cible.
- 4** Expertise de la cible : On ne répond pas de la même manière si la réponse est destinée à un enfant de 10 ans ou à un BAC +12.
- 5** Réponse attendue : C'est le format (Check-list, code JS, Diagramme, MindMap, tableau, rich-text, slides, image...), la longueur, la structure, la langue, neutre ou exprimant des prises de position, etc.
- 6** Ton : C'est le qualificatif de l'écriture (direct, éducatif, concis, formel, conversationnel...), la forme (langage soutenu, forme active...), les éléments à inclure (données, citations...), la technique de rédaction (storytelling, copywriting...), température...
- 7** Process : indique au prompt la marche à suivre pour réaliser la tâche, les étapes à suivre, l'ordre dans lesquelles les suivre, les consignes à respecter à chaque étape.

C'est notamment dans cette 7° étape qu'il faudra demander à ChatGPT des recherches préalables à la réponse pour qu'il aille chercher dans son corpus ou sur le web les bonnes informations. (et non en lui attribuant un rôle qui par magie lui ferait mobiliser les bonnes informations 😊 Dont il ne dispose pas en raison des lacunes de son corpus détaillées page 17. Ce qui renvoie à cet article : [Attribuer un rôle à ChatGPT ? Absurde idée reçue sur les prompts !](#))

Un prompt ne peut donc pas être une simple question...

- ... comme Bing de Microsoft vous y invite bêtement
- ... comme la majorité des gens non formés à ces techniques le font
- ... et comme la majorité des bases de prompts offertes ou vendues par des pseudo-experts autoproclamés sur le web le font ;

Un prompt doit avoir les 7 couches d'instructions expliquées en page précédente, qui peuvent nécessiter 2 ou 3 pages de rédaction ! D'où la nécessité de comprendre et gérer les tokens, car il faut prendre en considération qu'un prompt de 3 pages contient environ 1 300 mots, donc environ 1 800 tokens pour les instructions.

- Avec GPT-3 ou l'API, il n'en reste que 2 300 pour la réponse, soit 1 600 mots environ en français :
- plus de quoi passer la barre des 2 000 ou 2 200 mots qu'ont en moyenne les pages les plus lues sur le web dans les blogs.

2. Custom Instructions pour les prompts récurrents

Dans les 7 couches d'un prompt, une doit être personnalisée à l'entreprise (B2B) ou au magasin (B2C) : Sujet et contexte. Elle correspond à ce que l'on appelle l'USP (Unique Selling Proposal) :

« je suis qui, je résous quels problèmes mieux que personne, que qui rencontre, comment et pourquoi ».

La rédaction de cette couche demande un vrai travail de réflexion, que l'on fait une fois pour toutes. Cette couche peut aller du format « elevator pitch » d'une minute, donc 100 mots, à 4 fois plus en raison des nombreux sujets à renseigner :

- Prénom et Nom (sera utile pour les prompts créant des emails personnalisés),
- fonction,
- entreprise,
- localisation,
- offre de Produits ou Services,
- zone de chalandise,
- Entreprises visées en B2B (taille, secteur d'activité, etc.),
- interlocuteurs visés (fonctions, buyer personas, CSP, âge, caractéristiques distinctives, pain, points et problématiques principales, etc.)
- avantages de l'offre, caractéristiques de l'offre prouvant les avantages,
- concurrents et avantages concurrentiels, actualité, thèmes d'intérêts ou sujets de discussion, etc.

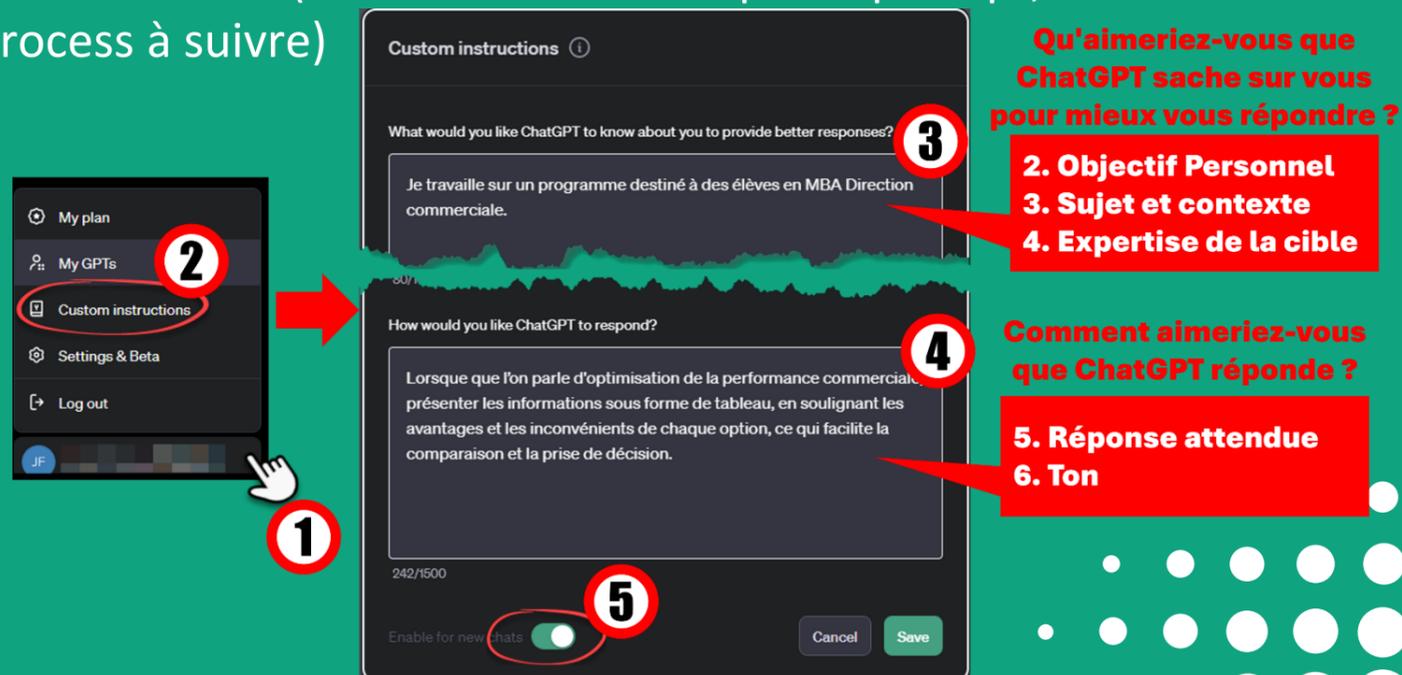
Cette couche contextuelle sera souvent réutilisée « telle quelle » avec des prompts différents, car les 6 autres couches sont liées à la tâche que le prompt doit accomplir.

Dès lors il sera fastidieux de recopier à chaque fois une page ou plus de données contextuelles au-dessus des 6 couches spécifiques à la mission du prompt.

Les « custom instructions » sont une fonctionnalité qui évite ces répétitions : Vous saisissez une fois pour tous les prompts à venir la couche contextuelle et 4 autres dans le backoffice de ChatGPT, et elles seront utilisées pour tous les prompts suivants sans que vous n'ayez à les répéter.

Pour accéder à cette fonctionnalité, un suffit de

1. cliquer sur votre login en bas à gauche de l'interface,
2. puis sur Custom Instructions.
3. Saisir les couches 2, 3 et 4 dans la 1ère zone. (2. Objectif, 3. Sujet/contexte, 4. expertise cible)
4. Saisir les couches 5 et 6 dans la 2ème zone. (5. Réponse attendue et 6. Ton)
5. Tant que le bouton « Enable for new chats » (Activation pour les nouvelles discussions) sera sur « on », il vous suffira de limiter vos prompts aux couches 1 et 7 (1. Tâche à effectuer par le prompt, 7. Process à suivre)



Cette fonctionnalité présente 2 inconvénients dus au fait que vous ne pouvez saisir qu'un jeu d'instructions personnalisées :

1. Vous pourriez (devriez) avoir plusieurs jeux de couches « sujet et contexte » adaptés aux différentes activités ou lignes produits de l'entreprise, et l'objectif peut varier.
2. La réponse attendue dépend de la couche 1 (Tache) du prompt, qui varie.

Donc avant de lancer un prompt... vous devrez vous rappeler sans jamais oublier ni faire d'erreur, quelles custom instructions sont saisies, puis si elles ne conviennent vous devrez les modifier, ou les désactiver et les ajouter directement dans votre prompt.... Pffffff.

3. GPTs pour les prompts récurrents

Une autre fonctionnalité, évitant de saisir de multiples fois les mêmes couches de prompt, sont les GPTs. Selon OpenAI « Les GPTs sont un nouveau moyen pour chacun de créer une version personnalisée de ChatGPT afin que l'IA conversationnelle se rende plus utile au quotidien et pour des tâches spécifiques, au travail comme à la maison, puis de partager cette création avec d'autres. » En fait il s'agit plutôt d'un prompt qui s'active dès que l'on clique dessus et qui contient déjà les 7 couches déjà optimisées. Le prompt demandera juste des informations supposées être adaptées à votre contexte.

Qui optimise les 7 couches ? Vous, sur un prompt que vous avez peaufiné et qui fonctionne parfaitement. Ou un tiers dont vous vous êtes fourni le GPTs sur le GPT Store. Les limites des GPTs sont qu'à peine 8 jours après la mise en ligne du GPT Store il en contient déjà 3 millions qui contiennent les meilleurs et les pires des prompts, sans que vous ne puissiez le savoir.

4. Éviter les 4 pièges les plus fréquents

Le niveau actuel des utilisateurs en matière de prompts est terriblement faible ; même le niveau des formateurs et soi-disant experts autoproclamés : les 2èmes conduisant sans doute aux premiers ! Voici les 4 erreurs les plus courantes :

1. **Les prompts de 3 lignes** qui sont souvent uniquement la 1 ère couche : la tâche à exécuter. Ils sont le droit chemin vers des réponses d'une banalité atterrante, réponses incomplètes, réponses erronées.

2. **L'attribution de rôle au prompt** pour le doter de compétences. Ce sont ces prompts commençant par :

« tu es expert en... avec une spécialité dans.... »,

« tu maîtrises toutes les techniques de... pour... »,
« tes compétences te permettent de déterminer les meilleures... »

« Utilise tes connaissances en... » etc.

Comme le montre cet article, [« attribuer un rôle à ChatGPT ? Absurde >>](#), des milliers d'expérimentations démontrent qu'attribuer un niveau d'expertise à ChatGPT n'a aucun impact sur la réponse de ChatGPT. Il s'agit d'une confusion avec le fonctionnement de l'API pour les développeurs qui gère 3 rôles (assistant, système utilisateur), mais qui ne s'applique pas au message de prompt. Cela surcharge le prompt et consomme des tokens pour rien puisque la réponse n'est pas modifiée. Et surtout c'est fait souvent en lieu et place de définir le niveau d'expertise de la cible de la réponse, qui lui impacte fortement la réponse.

3. Le tutoiement de ChatGPT : Ce sont ces prompts demandant

« peux-tu me proposer... »

« Peux-tu me donner la liste de... »

« Ta tâche consiste à... »

« Réponds uniquement par... »

Comme l'explique ce numéro [« Réveillez-vous ! Le tutoiement de ChatGPT, syndrome d'un manque de compréhension de l'IA »](#), cette pratique relève de l'anthropomorphisme, comportement bien connu de l'être humain qui attribue des caractéristiques humaines à ce qu'il ne comprend pas pour se rassurer.

Le problème est que l'anthropomorphisme réduit notre recul et déforme notre appréciation objective de ce qui n'est qu'une application.

4. Réduire la réponse de ChatGPT à la seule donnée recherchée. Ce sont ces prompts terminant par

« Réponds uniquement par... ».

Le but est d'éviter ces longues tirades ennuyeuses et consommatrices de tokens dans lesquelles ChatGPT explique ses étapes et son raisonnement.

Le problème est que comme nous le montrerons dans cet article [« Un escargot met en échec ChatGPT et révèle ses limites »](#), cette instruction a la fâcheuse tendance à conduire ChatGPT à se dispenser des étapes nécessaires pour construire une réponse juste, et le conduit à répondre « n'importe quoi ».

CONCLUSION

Les chapitres 1 à 5 de ce document sont importants pour comprendre et suivre discussions et lecture sur ces sujets. Importants pour votre image aussi : ne pas passer pour « has been ».

Mais le sujet le plus important est le 6eme, **les prompts et le prompt engineering** : car lui seul conditionne l'efficacité de l'utilisation de ChatGPT. La construction des prompts est cruciale pour tirer la quintessence de ChatGPT : indispensable pour une utilisation professionnelle.

En tant que dirigeant, vous devez en connaître un minimum pour

- 1. pouvoir évaluer les potentialités de ChatGPT pour votre entreprise.** Elles sont énormes : elles varient selon les entreprises, mais à la clé ce sont de véritables sauts de performance et/ou des économies colossales. Vous ne pouvez faire l'économie de les investiguer. Pour ce faire, vous devez disposer de « vrais prompts », pas de ces 3 lignes qui pullulent sur le web.

2. Vous assurer que les collaborateurs l'utilisent

correctement. Vous ne devez pas devenir expert en prompt engineering, de même qu'on ne demande pas à un entraîneur de foot d'être le meilleur joueur de l'équipe. Mais parce que de même que l'entraîneur doit avoir des notions sur « comment marquer un but » pour garantir la performance de l'équipe, vous devez avoir des notions pour garantir le bon fonctionnement dans votre entreprise.

3. Et enfin pour augmenter votre propre potentiel et votre propre performance de dirigeant :

bien utilisé ChatGPT vous fera gagner un temps fou sur des tâches aussi variées que la veille, la préparation d'interventions publiques, la résolution de problèmes et la prise de décision... tout en augmentant l'acuité de vos raisonnements.

Évitez les 4 pièges les plus fréquents listés page 27 :

1. Les prompts de 3 lignes
2. L'attribution de rôle au prompt
« Tu es expert en... tu maîtrises... »
3. Le tutoiement du prompt
« peux-tu faire... et.... »
4. Réduire la réponse de ChatGPT par
« répondre uniquement par... »

Fuyez ces prompts et ceux qui vous les recommandent. Ne payez pas de formations qui enseigneraient ces mauvaises pratiques à vos collaborateurs.

Si vous souhaitez monter vos équipes en compétence sur l'utilisation de ChatGPT, n'hésitez pas : rejoignez notre programme de formation avant-gardiste, conçu pour divers profils professionnels :

- Journée découverte des enjeux, limites et potentiels de ChatGPT pour CODIR et COMEX : Consultez-nous >> pour en savoir plus.
- Formation ChatGPT pour les pros de la vente >>
- Formation ChatGPT pour les pros du marketing >>
- Formation personnalisée : Nous proposons des sessions adaptées pour les dirigeants, managers, RH, chefs de projets, financiers, professions de santé, et bien plus. Consultez-nous >> pour en savoir plus.

Cloud-Business-School c'est :

- Une garantie de professionnalisme avec des formations délivrées par Cloud-Business-School, organisme de formation certifié Qualiopi.
- Expertise et Innovation : Notre formation ChatGPT est la plus avancée du marché. Pourquoi ? Parce que nous avons réalisé des milliers de tests pour construire et sélectionner les meilleures techniques d'utilisation de ChatGPT.
- Méthode DOMPTER : Une révolution dans l'apprentissage du prompt engineering. Bénéficiez de notre méthode exclusive « DOMPTER », une synthèse de connaissances acquises à travers des milliers de tests. Cette approche vous permettra de maîtriser l'art du prompt engineering et de produire des prompts efficaces.

- Base de Données de prompts optimisés : Nos participants quittent la formation avec une collection inégalée de prompts optimisés pour leur fonction spécifique. Contrairement aux bases de données de prompts génériques qui inondent le marché, nos prompts sont affinés et optimisés grâce à la méthode DOMPTER.



BDD PROMPTS OPTIMISES CBS

► MODE D'EMPLOI

☰ Catégories de prompts ▾

Liste des catégories de prompts

☰ Catégories + ...

📄 Veille

📄 Stratégie Marketing & Digital

📄 Recherche de mots clés

📄 Rédaction Article / pages Web

Rédaction Article / pages Web

Les données de base utilisées ici sont les données de base décrites dans les instructions ci à gauche dans le guide utilisateur, complétées de 2 zones spécifiques pour la rédaction d'articles : {{sujet}} : que vous renseignez avec le sujet de votre article, et {{CTA}} que vous renseignez avec le call to action qui doit accompagner votre article.

★ Commencez ici par personnaliser les Données de base

▼ 1. Structurer + Rédiger une page / article

📄 1A. Rédaction page "informationnelle"

📄 2B. Structurer et rédiger une page "Mode Achat"

► Optimisation SEO OnPage