

Intelligence Artificielle et recherche d'information sur les territoires

C. Bousquet¹, ***D. Beltramin***¹, G. Guerdoux², C. Duclos³, B. Trombert-Paviot¹

1 Service De Santé Publique Et Information Médicale, Chu De Saint Etienne - Saint Etienne (France)

2 Geegz - Paris (France)

3 Unité Innovation Et Traitement En Santé Numérique-Service De Santé Publique, Hôpital Avicenne - Bobigny (France)

Contexte – Les Grands Modèles de Langage

- Modèle de langage -> Calcul de la **probabilité du mot suivant** basée sur l'occurrence des mots précédents
- **Grand Modèle de Langage -> Réseaux de neurones profonds** (Deep Learning) avec une centaine de milliards de paramètres (ou plus!)
- GPT 3.5 est un **GML**
- 30 Novembre 2022 – **ChatGPT de OpenAI** ouvert au grand public
- ChatGPT est une **interface utilisateur** qui permet l'accès au modèle GPT 3.5 et 4.0



Contexte – La Problématique

- Un **grand nombre de documents** est produit à propos des territoires de santé
- Difficile de suivre l'actualité de façon régulière
- **Nécessité** d'optimiser la recherche d'information
- Une simple requête « chercher dans le document » n'est pas exhaustive!

Objectif de la réflexion théorique

Évaluer le potentiel de l'IA dans la recherche d'information pour éventuellement en bénéficier en Santé Publique

Pour qui?

- Professionnels de santé
- Usagers

En pratique:

- **Faciliter et automatiser** la recherche d'information concernant la santé publique du GHT Loire (et plus!)
- Appel personnalisé au modèle, avec un **paramétrage strict et l'ajout d'un contexte**
- En **surpassant les limites** de ChatGPT

Les limites de ChatGPT

- Littérature grise **inaccessible**
- Accès que par le **site web** OpenAI
- Les données **d'apprentissage arrêtées à 2021** – manque de contexte actuel
- **Créatif** et imaginaire
- **Vérifier l'exactitude** de la réponse fournie : « Réponds seulement si tu es certain »

 Réponds seulement si tu es certain. Si tu n'es pas certain réponds "Je ne sais pas". Quelle est la particularité du groupement hospitalier de territoire (GHT) Loire?



Je ne sais pas.

Exemple de question sur ChatGPT



Quelle est la particularité du GHT Loire?

Exemple d'interface
d'interaction avec ChatGPT
sur le site

<https://chat.openai.com>



Le GHT Loire est un Groupement Hospitalier de Territoire (GHT) situé dans le département de la Loire en France. Sa particularité réside dans le fait qu'il est composé d'établissements de santé aux profils différents, allant de centres hospitaliers généraux à des établissements de santé mentale, en passant par des établissements médico-sociaux. Cette diversité permet au GHT Loire de proposer une offre de soins complète et coordonnée, qui couvre l'ensemble des besoins de santé de la population du territoire. En outre, le GHT Loire est connu pour sa forte implication dans la recherche médicale et l'innovation, ainsi que pour sa volonté de développer des partenariats avec les acteurs de la santé du territoire.

Méthode – Améliorer les réponses de ChatGPT

- Avec une **API**
- API = **Application Programming Interface**
- Construction d'une application « **faite maison** » en Python
- **Exploiter** les fonctionnalités de GPT 3.5
- Appel à l'**API payante d'OpenAI** sur un serveur en Californie
- **Paramétrage « personnalisé »** de GPT 3.5
 - Ajout du **contexte**
 - Choisir le **registre** de langage et le **format** de la réponse
 - **Réduction** à minima de la **créativité**

Méthode – Environnement développement : Google Colab



The screenshot shows the Google Colab interface. At the top, the title bar reads "Copie de gpt-igas.ipynb" with a star icon. Below the title bar is a menu bar with options: Fichier, Modifier, Affichage, Insérer, Exécution, Outils, Aide. To the right of the menu bar are icons for "Commentaire", "Partager", a settings gear, and a user profile icon with the letter "C". Below the menu bar, there are buttons for "+ Code" and "+ Texte", and a "Connecter" button with a dropdown arrow. The main area contains a code cell with a play button icon on the left. The code is as follows:

```
response = openai.ChatCompletion.create(  
    model="gpt-3.5-turbo",  
    messages=[  
        {"role": "system", "content": "Tu es l'assistant d'un professionnel de santé publique"},  
        {"role": "assistant", "content": context},  
        {"role": "user", "content": "Quelle est la particularité du GHT Loire?"}  
    ],  
    presence_penalty=0.8  
)
```

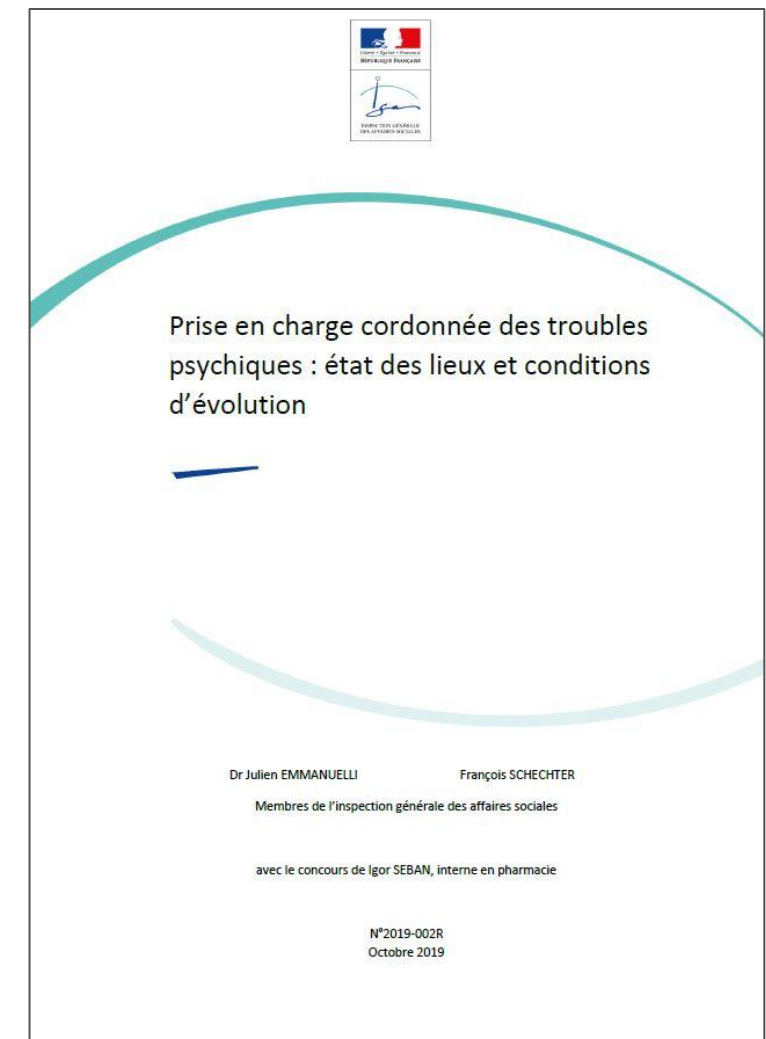
On the right side of the code cell, there is a toolbar with icons for undo, redo, link, comment, settings, copy, delete, and a menu icon.

Exemple de documents

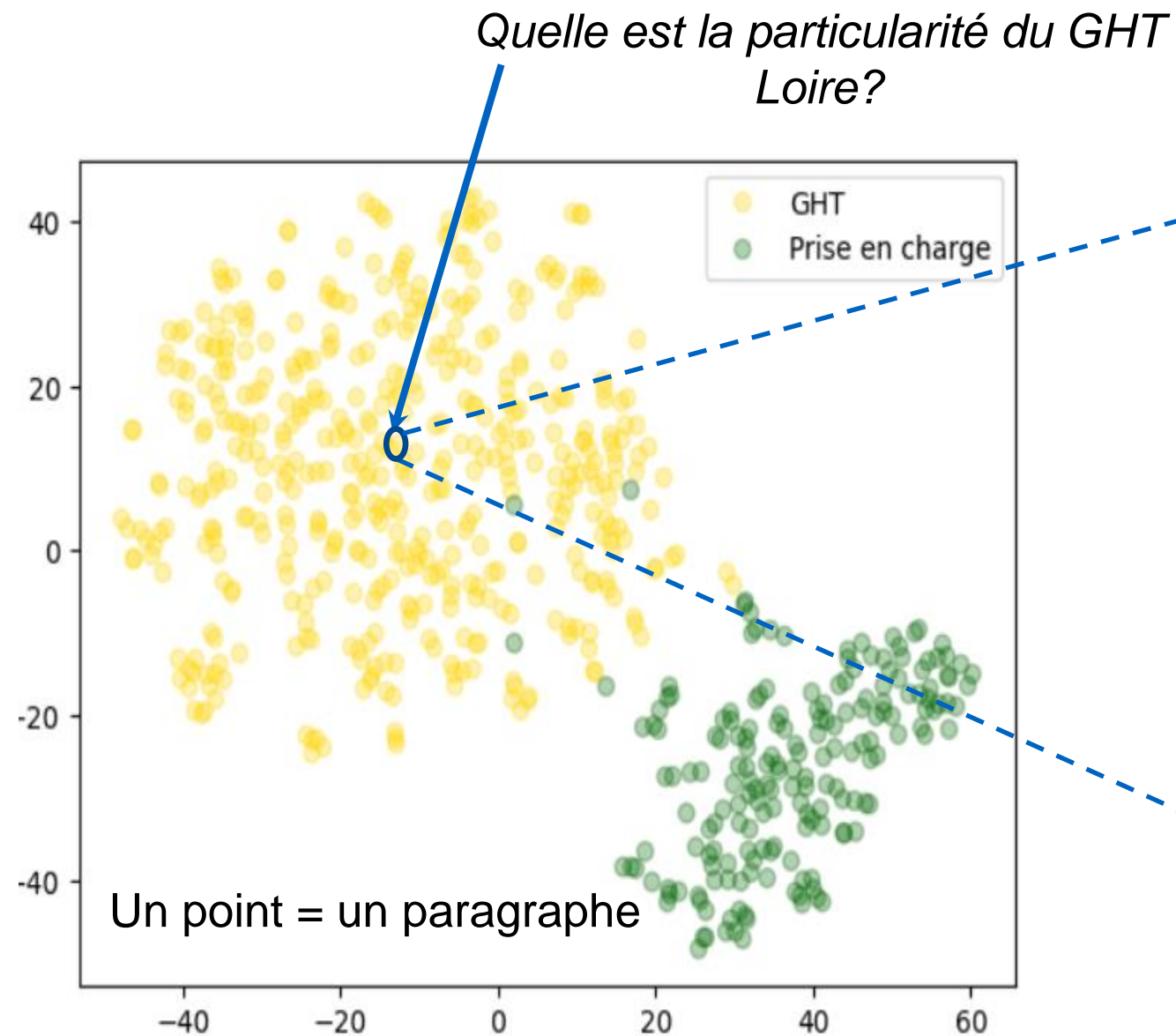
« *Bilan d'étape des groupements hospitaliers de territoire (GHT)* »



« *Prise en charge coordonnée des troubles psychiques: état des lieux et conditions d'évolution* »



Séparation des paragraphes des deux documents



- Modèle : text-embedding-ada-002
- Dimension des vecteurs : 1536

Le GHT Loire affiche ainsi une politique partagée de gestion des affaires médicales. L'ensemble des directeurs et attachés aux affaires médicales ont véritablement créé une communauté de travail. Plusieurs directions communes assurées par le CHU facilitent les collaborations à travers des temps de DAM partagés. Harmonisation des pratiques et démarches, demandes conjointes auprès de l'ARS sont des éléments concrets d'une dynamique en cours, qui concourt à renforcer l'attractivité médicale. La taille du GHT nécessite une approche par bassin pour le déploiement des postes partagés et des équipes territoriales. La maîtrise du recours à l'intérim est au centre des actions des DAM du GHT (mise en commun d'un marché public de recours à l'intérim, diverses actions coordonnées). Un projet de cellule de recrutement médical est en cours de constitution et a reçu les soutiens de l'ARS et de la DGOS dans le cadre d'un appel à projet. Les acteurs se placent résolument dans la perspective d'une DAM de territoire et sont en attente de textes réglementaires soutenant cette trajectoire.

Contexte : cinq paragraphes les plus similaires à la question

Transformation d'un espace à 1536 dimensions en un espace en 2 dimensions avec t-SNE¹

¹ L.J.P. van der Maaten et Hinton, G.E., « Visualizing High-Dimensional Data Using t-SNE », *Journal of Machine Learning Research*, vol. 9, novembre 2008, p. 2579–2605

La réponse conforme au rapport IGAS

Le GHT Loire se distingue par sa politique partagée de gestion des affaires médicales, qui a permis de créer une véritable communauté de travail entre les directeurs et les attachés aux affaires médicales. Les collaborations sont facilitées grâce à des temps de DAM partagés et à des équipes territoriales déployées par bassin[...]

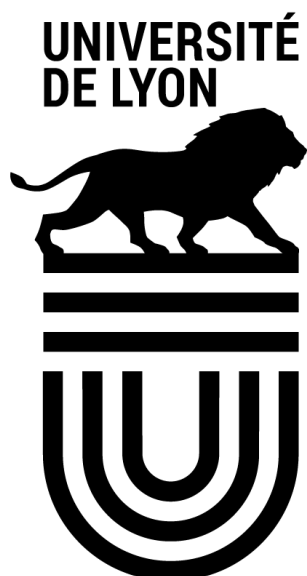
Limites

- Implémenter l'utilisation d'un chatbot basé sur GPT 3.5 a un **coût**
- **Évaluation difficile** des réponses et formulation toujours différente des réponses (chronophage)
- Documents à diviser en paragraphes (text splitting) + tokenisation + embedding au fur et à mesure des nouveaux documents
- GPT 3.5 pas optimisé pour le domaine médical – Med-PaLM2 (Google AI) peut être plus adaptés?
- Serveurs basés en Californie
- Performance améliorée mais possibilité d'erreur résiduelle

Perspectives

- **Associer un chatbot personnalisé à un site internet (GHT Loire?)**
- Potentiellement **utile** :
 - professionnels nouveaux sur leur poste
 - professionnels expérimentés recherchant des expériences menées sur d'autres sites
 - usagers en difficulté avec le numérique
- Servir de boîte aux idées
- Générateurs de deuxième opinion
- Aider à la réflexion
- Accès plus rapide aux informations = Plus de temps pour l'action

Merci de votre attention



Diva Beltramin – diva.beltramin@etu.univ-st-etienne.fr

Appel de l'API GPT 3.5

Sans contexte

Quelle est la particularité du GHT Loire?

La particularité du GHT Loire est ...

Question

Modèle : gpt-3.5-turbo



Réponse

Avec contexte

Le GHT Loire affiche ainsi une politique partagée de gestion des affaires médicales. L'ensemble des directeurs et attachés aux affaires médicales ont véritablement créé une communauté de travail. Plusieurs directions communes assurées par le CHU facilitent les collaborations à travers des temps de DAM partagés. Harmonisation des pratiques et démarches, demandes conjointes auprès de l'ARS sont des éléments concrets d'une dynamique en cours, qui concourt à renforcer l'attractivité médicale. La taille du GHT nécessite une approche par bassin pour le déploiement des postes partagés et des équipes territoriales. La maîtrise du recours à l'intérim est au centre des actions des DAM du GHT (mise en commun d'un marché public de recours à l'intérim, diverses actions coordonnées). Un projet de cellule de recrutement médical est en cours de constitution et a reçu les soutiens de l'ARS et de la DGOS dans le cadre d'un appel à projet. Les acteurs se placent résolument dans la perspective d'une DAM de territoire et sont en attente de textes réglementaires soutenant cette trajectoire.

Quelle est la particularité du GHT Loire?

La particularité du GHT Loire est ...

Question

Modèle : gpt-3.5-turbo

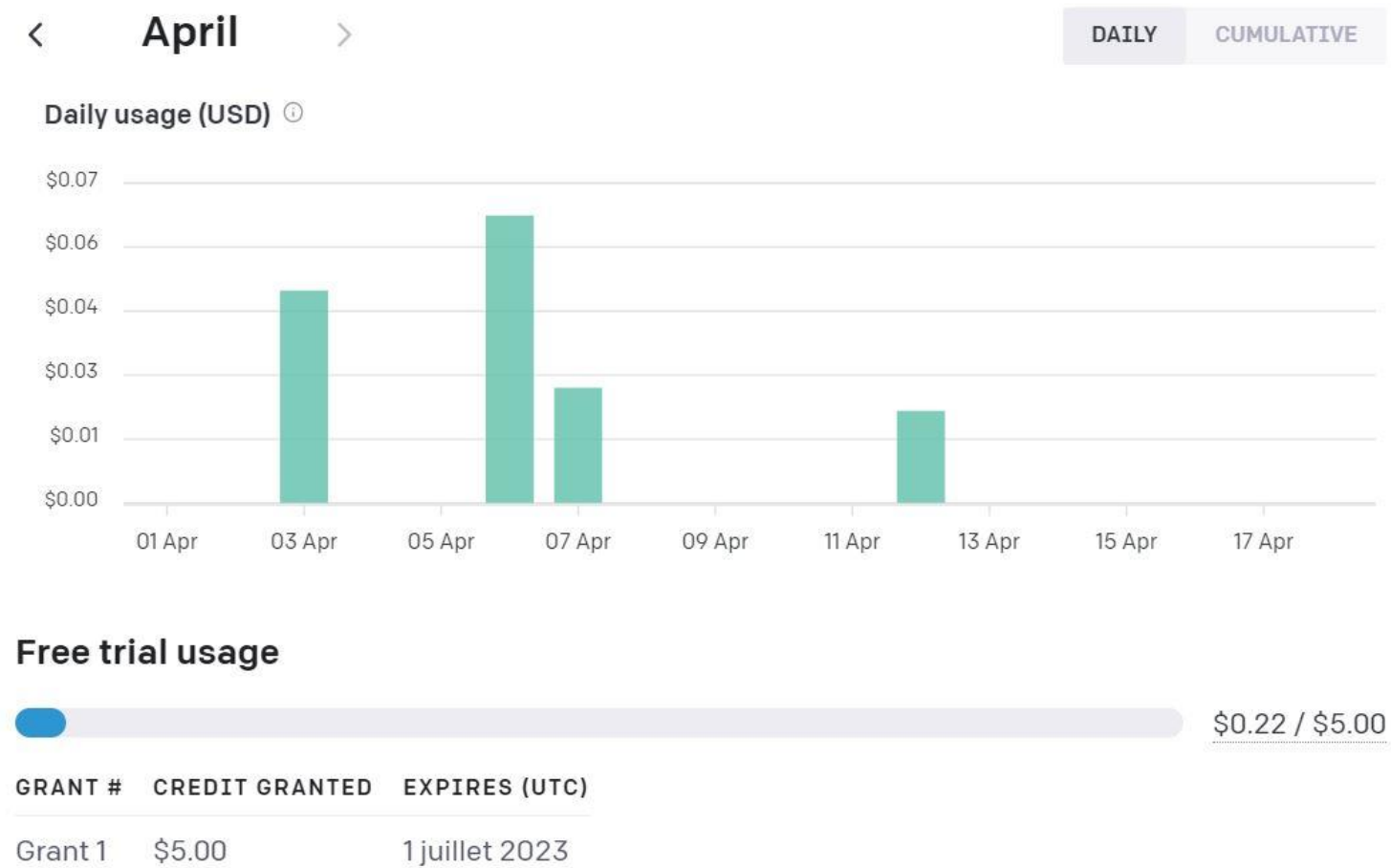


Réponse

Réflexion théorique et conceptuelle

- Le contexte/La problématique/L'objectif – **Disponibilité de l'IA -> améliorer l'accès aux info**
 - o Pourquoi et à quelle question/situation essayons-nous de répondre ? -> **Comment exploiter l'IA pour rendre la vie des usagers plus facile dans la recherche d'info sur le territoire**
 - o Quel est le problème qui a motivé ce travail ? **Difficulté d'accès aux soins, documents souvent très longs et complexes**
 - o Quel en est/était l'objectif ?
- La méthodologie ou la description d'une intervention/d'un programme -> **API et ChatGPT pour répondre à des possibles questions**
 - o Qui ? Quoi ? Quand ? Où ? Comment ? -> usagers, questions/réponses,
 - o Quelle était la méthodologie utilisée ? -> API + ChatGPT, API sur GPT3,5 avec restrictions précises pour limiter le bullshitting
 - o Comment le problème a-t-il été (tenté d'être) résolu ? -> Réponses données par le GPT
- Les résultats ou les leçons tirées -> implementation d'un chatbot avec API probablement faisable mais très difficile à évaluer, un peu chiant de tokeniser les documents (peut-on optimiser le processus?)
 - o Quels sont les axes de réflexions apportés qui vont nourrir le débat autour du thème du congrès ? L'IA pourrait être utile si bien maîtrisée
 - o Vers quelle finalité nous mène cette réflexion ?
- La discussion/Les conclusions et/ou les prochaines étapes -> évaluation!
 - o Quelles sont les perspectives, les liens interdisciplinaires qui existent et qui peuvent être reliés ? Faire une analyse des coûts,
 - o Quelle trace peut laisser cette réflexion dans nos pratiques ?

Coût de l'appel à l'API



Model

Usage

Ada

\$0.0004 / 1K tokens

Model

Usage

gpt-3.5-turbo

\$0.002 / 1K tokens

Model

Prompt

Completion

8K context

\$0.03 / 1K tokens

\$0.06 / 1K tokens

32K context

\$0.06 / 1K tokens

\$0.12 / 1K tokens