

Notes de la conférence de Véronique Mesguich sur l'Intelligence artificielle Réseau Prisme 12/01/2024

Sommaire

1 ^e partie. Qu'est-ce que l'Intelligence Artificielle (IA), comment fonctionne-t-elle	2
Introduction - définitions	2
Comment cela fonctionne ?	3
Ce que Chat GPT sait faire	3
Ce que Chat GPT ne fait pas ou pas bien.....	3
Intégration actuelle de Chat GPT dans les moteurs de recherche.....	4
Conseils.....	5
2 ^e partie. Aspects documentaires	5
Notion de Prompt, comment rédiger ses questions	5
Utiliser Chat GPT dans les phases de la veille documentaire.....	6
Limites et dangers	7
Risques concernant le sourcing.....	7
Conclusion	7
Questions / réponses (extraits)	8

1^e partie. Qu'est-ce que l'Intelligence Artificielle (IA), comment fonctionne-t-elle

Introduction - définitions

La conférence porte sur l'**IA générative** (l'IA est beaucoup plus vaste). Générative : permet de générer du texte.

Les IA sont une aide pour la recherche et le traitement des données, nous verrons comment, et en particulier pour les documentalistes. Seront aussi abordés : les dangers et limites de l'IA.

Chat GPT se définit lui-même comme un « modèle de langage développé par la startup Open AI, formé sur de grandes quantités de textes trouvés sur internet, et capable de générer des textes en se basant sur les patterns et les connaissances apprises pendant ses entraînements. Il est utilisé pour du chatbot (robot conversationnel), la génération de textes, la traduction automatique et bien d'autres choses encore. »

On peut demander par exemple à Chat GPT de se comparer à Google en générant un tableau.

Définition du **Prompt** : conversation textuelle avec les utilisateurs.

Les moteurs de recherche Bing et Google Bard utilisent déjà Chat GPT.

Les réponses données par Chat GPT peuvent être très justes... ou incorrectes !

GPT veut dire Generation Pretrained Transformer.

Chat GPT est sorti fin 2022, mais la technologie n'est pas nouvelle. Ce qui est nouveau c'est le « Chat ».

G = Génératif de texte

P = Préformation. Chat GPT est capable de prédire la probabilité que les mots ont de se trouver les uns à côté des autres, sur la base de milliards de documents numériques

T = Transformer = Architecture de réseau qui convertit les mots en vecteurs.

La Start up Open AI a été créée en 2015 par Sam Altman et Elon Musk (qui est parti en 2018), en partenariat avec Microsoft et Amazon web service.

// en 2019 Google a lancé son algorithme *Bert*.

Au printemps 2023, la version Chat GPT 4 (payante) est sortie (plus puissant, plus fiable).

En même temps, les moteurs de recherche intègrent les IA génératives.

// certains experts inquiets par rapport aux dangers potentiels demandent une pause de 6 mois dans la recherche sur l'IA, dans une lettre ouverte.

Novembre 2023 : lancement de GPT4 turbo et GPT builder, plus rapides, plus puissants permettant d'avoir un assistant spécialisé pour créer des GPT personnalisés.

Muriel Prost, notes de la conférence de Véronique Mesguich sur l'IA, janvier 2024

Comment cela fonctionne ?

Chat GPT utilise le procédé de machine learning, et des réseaux de neurones artificiels pour « apprendre » comment reproduire l'écriture humaine en plusieurs langues, avec deux méthodes : l'apprentissage supervisé (à partir d'exemples), et l'apprentissage par renforcement (à partir de commentaires humains sur les textes).

Il est assez « woke », paramétré pour être ni sexiste, ni raciste.

GPT dit lui-même qu'il n'est pas infallible et peut produire des résultats incorrects ou biaisés, il est important de vérifier ses réponses.

On appelle « hallucination » une réponse fausse ou trompeuse, et présentée comme des faits, apportée par une IA.

La version gratuite actuelle est Chat GPT 3.5. Ses sources sont le réservoir de pages web Common Crawl (les archives du web), les corpus de livres numérisés, Wikipedia (en croisant les sources).

Ce que Chat GPT sait faire

- Tenir une conversation structurée et contextualisée
- Trouver des idées sur un thème pour débroussailler en amont d'une recherche ou d'une veille (mais il ne peut pas faire la veille lui-même, il aide à trouver les mots-clés).
- Faire des résumés
- Rédiger des lettres ou des mails (le nouvel assistant Windows Copilot rédige)
- Faire des traductions
- Expliquer un concept en tenant compte de la cible
- Créer des histoires, poèmes, blagues...
- Répondre aux clients
- Analyser des sentiments, émotions exprimés dans un texte
- Générer des codes informatiques (pas parfaits) en C++, java ou python
- Réaliser la structure d'une carte mentale (mais pas la dessiner).

Ce que Chat GPT ne fait pas ou pas bien

- Les réponses liées à l'actualité car ses sources actuelles s'arrêtent en avril 2023 (pour GPT 4)
- Naviguer sur internet
- Comprendre les mots et contextualiser. L'IA ne fait que répéter et reformuler ce qu'elle a entendu, ça reste une machine !
- Produit des réponses incorrectes ou fausses... mais apprend par renforcement.

La conférence se poursuit par quelques exemples d'interrogation de Chat GPT 3.5 (version gratuite). On remarque que les réponses sont structurées par une série de points (1. 2. 3. 4. Etc). On peut demander à Chat GPT de développer un point précis.

Muriel Prost, notes de la conférence de Véronique Mesguich sur l'IA, janvier 2024

Remarques : **les sources ne sont pas citées. Il faut vérifier les informations. L'historique des conversations ne peut pas être effacé.**

A noter : le nouveau moteur Bing, basé sur l'IA, cite ses sources. Chat GPT ne le fait pas : il faut vérifier en faisant des recherches sur le web.

La conférencière utilise Chat GPT quand elle écrit ses ouvrages, mais a parfois eu des réponses fausses. Si on pose des questions très précises, Chat GPT a tendance à inventer des réponses ! (ex : donne moi des exemples de bibliothèques victimes de cyber attaques : les exemples donnés étaient faux. En reposant la question, il a dit « Je me suis trompé » !).

Intégration actuelle de Chat GPT dans les moteurs de recherche

- Google Bard (en France conforme au RGPD en juillet 2023)
- Bing, depuis mars 2023, dans le moteur et dans Teams, Edge (Microsoft)
- Copilot, depuis septembre 2023, assistant windows intégré dans Edge
- Le 6 décembre 2023, Google lance Gemini ajouté à Bard (pas encore en France).

Nouveau moteur **Bing.com** : il est basé sur Chat GPT 4, sans abonnement, avec des identifiants Windows. Aller dans Bing et cliquer sur « Conversation » = « copilot ». On peut choisir un style de conversation : « *précis* » recommandé. **Les réponses donnent des liens vers les sources.**

Google Bard est basé sur l'IA de Google, ajoute des images, mais pas de lien vers les sources. En France depuis juillet 2023.

Bing a une longueur d'avance sur Google.

Perplexity est un outil de recherche gratuit développé avec Chat GPT. www.perplexity.ai

L'interface est en anglais mais on peut converser en français. Les sources apparaissent. On peut ajouter une recherche images, vidéos...

On a des fonctionnalités supplémentaires en créant son compte (gratuit), par exemple la fonction attach pour charger un document et poser des questions sur ce document.

On peut chercher sur des sources « académiques », sous-ensembles du web.

Wolframalpha : fait des calculs.

Il y a un historique des questions, et on peut stocker des éléments.

You.com est aussi un robot basé sur Chat GPT

Claude est une IA générative qui ne dépend pas d'un géant du web, a une approche plus vertueuse, moins commerciale (startup Anthropic). Est peu accessible en Europe à cause du RGPD, mais on peut y accéder via le robot Poe.

Muriel Prost, notes de la conférence de Véronique Mesguich sur l'IA, janvier 2024

Autres...

Meta lance son Chatbot Meta AI « Llama2 »

Amazon développe le langage Titan avec Anthropic.

D'autres projets chez Apple ; annonce de Grok par Elon Musk, qui sera intégré à X ; Bloom est une IA generative basée sur la science ouverte Open source, développée par Hugging Face ; navigateur et moteur open source Brave avec l'AI Leo....

Cécile Dejoux propose un tableau comparatif sur son Mooc (été 2023) : Bard / Claude 2 / GPT....

Tout évolue à vitesse V !

Conseils

Pour bien se servir de l'IA il faut développer son habileté à bien rédiger ses Prompts (questions).

Se méfier des « hallucinations » de l'IA.

Utiliser un pseudo pour interroger car il y a des failles de sécurité.

Eviter de télécharger sur une IA des documents internes ou confidentiels... car on ne sait pas ce qu'en fait l'IA après.

Préférence de Mme Mesguich pour « Perplexity » car il cite ses sources et répond plutôt correctement.

2^e partie. Aspects documentaires

Démonstration de **Copilot**, que nous avons sous Windows, avec le navigateur Edge, ou sur le moteur Bing.com.

Notion de Prompt, comment rédiger ses questions

Conseils :

- Contextualiser sa demande
- Etre précis
- Pas trop d'instructions dans un même prompt, y aller par étapes
- On peut indiquer le format souhaité pour la réponse.

Le moteur est conversationnel, on peut donc relancer sur un point, reformuler...

Exemples de prompts : dresser une liste de pour et de contre par rapport à une mise en situation. Convertir des formules dans un autre système. Créer des quizz, des jeux pédagogiques. Créer une imitation, rédiger « à la façon de... » etc.

Muriel Prost, notes de la conférence de Véronique Mesguich sur l'IA, janvier 2024

Utiliser Chat GPT dans les phases de la veille documentaire

Chat GPT peut nous aider :

- A définir les besoins (trouver des mots-clés, le champ lexical, débroussailler le sujet, les sources, faire une synthèse de l'existant)
- A rechercher et collecter : aide à la création d'une requête
- A faire analyse et traitement : résumés, extraction d'entités nommées, synthèses
- Mise en forme et diffusion : aide à la génération de texte, aide à créer une mind map.

C'est un gain de temps sur différentes parties de la chaîne.

Phase préparatoire (débroussaillage)

Exemple de prompt : « Je suis chargée de veille dans tel domaine, quelles sont les tendances actuelles dans ce domaine en France, Chine, Etats-Unis ? »

« Analyse SWOT sous forme de tableau de tel secteur en France »

« Donne-moi le champ lexical de tel secteur »

Sourcing : « quels sont les salons professionnels ou les conférences dans le monde entier dans tel domaine ». « Présente moi sous forme de tableau une liste de medias classés par catégorie (presse quotidienne nationale, régionale ou professionnelle) dans tel pays ».

Mais attention : vérifier la pertinence et la fiabilité des sources !

On peut demander « Trouve moi des sources faisant autorité dans tel domaine »

Phase de traitement – reformatage des données

On peut demander à Chat GPT d'extraire d'un texte des données chiffrées et de les présenter sous forme de tableau (texte copié collé).

-toujours vérifier le résultat- !

L'IA génère du code Markdown pour créer une mind map (à coller dans un logiciel de map).

L'IA permet de résumer des documents. Exemples de prompts : « Résume les points principaux des textes » « Fais un bref résumé des principaux arguments présentés dans le texte » « Résume l'article en 3 à 4 phrases ».

Dans Perplexity, on peut uploader le texte, ou faire un copier coller dans chat GPT 3.5

Il existe des outils développés par Chat GPT pour traiter des documents (exemple : humata.ai) mais ils sont payants sinon limités.

Outil Getcody.ai : pour créer des bases de connaissances. Stockage de documents, constitution de dossiers, imports de sites web... A partir des dossier, on peut demander au chat « En te basant sur le dossier XYZ, peux-tu m'indiquer les tendances sur tel sujet... »

Christophe Deschamps a répertorié les outils créés via les modèles de langages, dans un tableau comparatif (le lien sera dans le support de la conférence).

Muriel Prost, notes de la conférence de Véronique Mesguich sur l'IA, janvier 2024

Il y a des dizaines d'outils fondés sur l'IA (ex : askyourpdf, dall-E qui convertit des textes en images) et il y a des IA génératives dans les logiciels de veille.

Limites et dangers

Tim Berners Lee (créateur du web) est hostile aux géants du numérique, mais il considère que les IA vont dans le sens de l'Histoire et dit qu'on aura des assistants de plus en plus personnalisés.

Risques concernant le sourcing

- Perdre de vue la notion de source
- Manque de transparence dans les sources, absence d'entraînement sur les sources de qualité
- Manque de représentativité dans les sources (dans le cas d'un sujet polémique)
- Risques de biais, hallucinations, erreurs
- Dérive de plagiat, d'appropriation artistique
- Prolifération... ou oubli
- Atteintes à la vie privée

D'où l'importance de professionnels avertis !

Les outils de détection du plagiat détectent les textes écrits par les IA (compilatio, zerogpt...) (dans les travaux étudiants, on reconnaît facilement le « style » Chat GPT).

Pour détecter les images générées par les IA : regarder les détails... mais ce n'est pas garanti !

Il existe des IA au service du fact checking : buster.ai, factinsect, newsback... c'est encore embryonnaire.

Conclusion

Le cerveau n'est pas remplacé, vérifier la fiabilité c'est l'avenir des journalistes, documentalistes...

Le temps gagné d'un côté peut être perdu en vérifications...

Que reste-t-il aux professionnels humains ?

- La contextualisation de l'information
- Savoir personnaliser finement
- Confronter les points de vue divergents
- Animer un réseau
- Accompagner les utilisateurs
- Former aux enjeux et à l'utilisation de l'IA

Dans Libération spécial IA de juin 2023, Cédric Villani a dit « La plasticité de notre cerveau, nos façons d'apprendre, les capacités de notre imagination vont pour l'instant bien au-delà de ce que font les programmes. »

Muriel Prost, notes de la conférence de Véronique Mesguich sur l'IA, janvier 2024

Questions / réponses (extraits)

La valeur ajoutée des documentalistes ? Comme quand Google est apparu, on maîtrise les fonctions avancées.

La question de la Propriété intellectuelle ? C'est un gros problème. Les choses se mettent en place, cf IA Act de l'Union européenne.