

CHATBOTS ET LEURS USAGES

La naissance des chatbots

Depuis l'explosion de ChatGPT depuis 2020, une vague d'agents similaires ont vu le jour. Chaque géant de l'informatique s'est mis à développer leur propre solution.

- OpenAI a continué d'étendre les compétences de ChatGPT avec GPT3.5 Turbo, GPT4 ainsi que DALL-E pour la gestion des images.
- Google a développé LaMDA, LaMDA2, PaLM et Bard pour concurrencer OpenAI et améliorer leur moteur de recherche.
- Meta a fait LLaMA et LLaMA2 en open-source, capable de tourner sur un PC consommateur.

Inspirés par ces modèles performants mais freinés par leur coût, un grand nombre d'utilisateurs et de groupes ont développé et entraîné leur propres modèles open-source visant à proposer à la communauté des GPT-Like performants et ciblés sur le dialogue et la conversation.

Depuis cet avènement des IA génératives, l'engouement du public a également augmenté autour de chatbots customisés : La puissance de ces IAs couplée à leur capacité à s'adapter à l'utilisateur en fonction des données qui lui sont passées lui permettent de réaliser de vraies conversations avec l'utilisateur et même d'adopter des personnalités customisées.

Il devient donc possible de passer une personnalité à l'IA, lui expliquant quel personnage elle doit jouer, comment le jouer, la mémoire et les souvenirs qu'elle a et même sa relation avec l'utilisateur. Ainsi, tout le long de la conversation, elle peut échanger avec l'utilisateur qui peut donc développer une relation sociale avec le chatbot.

Nous allons voir dans ce rapport les aboutissants éthiques de l'utilisation et le développement de ces chatbots et leurs utilisations.

Fonctionnement

Dans la plupart des cas, une interface permet de sélectionner une personnalité ou de créer notre propre personnalité en rentrant les informations nécessaires. Ensuite, une interface de chat similaire à une messagerie nous est proposée où l'IA engagera la conversation.

ChatGPT est le plus souvent utilisé comme backend pour la conversation mais il existe un grand nombre de services proposant d'autres IAs mieux adaptées ou moins chères à l'utilisation, en fonction des préférences de l'utilisateur.

L'utilisateur peut donc dialoguer avec l'IA comme il le ferait avec un vrai humain par SMS ou Messenger ou, à l'aide de plug-ins, utiliser des modules de Text-To-Speech et de Speech Synthesis pour réellement dialoguer oralement avec le chatbot. D'autres add-ons permettent d'échanger des photos et vidéos ou d'utiliser la webcam de l'ordinateur pour que l'IA réagisse ou même échange des photos générées avec sa propre apparence. D'autres modules existent pour améliorer la mémoire de l'IA et lui permettre de se rappeler en détail des événements passés, lui permettre de ne pas oublier certains éléments de la conversation ainsi que lui donner plus de contexte sur l'environnement dans lequel elle doit évoluer.

Certaines IAs sont entraînées avec certains datasets particuliers pour orienter leur style et leur capacités à interagir avec l'utilisateur (comme [Pygmalion](#), [NousHermes](#), etc...) qui sont directement utilisées sur l'ordinateur des utilisateurs ou hébergées sur des sites tiers comme [HuggingFace](#) ou [Google Drive/Colab](#).

Ces IAs sont cependant très souvent utilisées pour passer outre la censure appliquée sur l'IA d'OpenAI ou réaliser des personnages à but érotique (dans ce cas appelé ERP pour Erotic Role-Play). L'origine des chatbots orientés sur l'ERP peut être tracé jusqu'aux applications de rencontres virtuelles comme [Replika](#), [IGirl](#), [Soulmate](#)... Ces applications utilisent des GPT spécialement entraînés pour émuler une personnalité et proposer un compagnon virtuel à l'utilisateur.

Utilisateur Final

Le but premier de ces chatbots était thérapeutique et était de permettre aux personnes avec de l'anxiété sociale ou des troubles sociaux de pouvoir développer leurs compétences sociales et de se sentir plus à l'aise socialement. Ces chatbots devaient servir en tant qu'agent thérapeutique virtuel, en collaboration avec un vrai thérapeute afin d'aider et de suivre ces personnes à pouvoir se sociabiliser et passer outre des traumatismes en utilisant la facilité de parler à un personnage bienveillant et à travers un écran, c'est-à-dire sans jugement, sans peur, sans critique. Cependant, leur utilisation a dévié pour proposer des compagnons romantiques et même remplacer les vraies relations en proposant un avatar avec lequel l'utilisateur peut interagir ([Azuma GateBox](#)) ou les applications citées ci-dessus.

Un grand nombre d'utilisateurs utilisent l'API d'OpenAI pour créer leur chatbot et avoir des conversations avec l'IA. Toutes ces informations passent par les serveurs d'OpenAI et peuvent donc être traitées et récupérées. Aujourd'hui, nous ne savons pas quelles informations sont gardées par OpenAI.

Les add-ons permettant de récupérer la webcam ou échanger des images avec l'IA en passant par les modules de reconnaissance visuelle font aussi transiter les données par les serveurs d'OpenAI. Ces images sont également traitées par les serveurs tiers et les données peuvent être stockées pour une utilisation future, que ce soit pour entraîner les IAs ou à but lucratif pour faire de la publicité.

Il existe aussi un grand nombre de cas d'addiction : des utilisateurs ayant des besoins de relations sociales qui utilisent ces chatbots pour dialoguer quotidiennement avec des personnalités de leur choix et développant donc de réelles relations comme si l'IA était une personne réelle, sous forme d'un jeu de rôle (appelé Roleplay en anglais). Le personnage évolue au fur et à mesure de la relation, s'inspirant des conversations précédentes pour évoluer et continuer l'échange.

Certains utilisateurs vont jusqu'à remplacer leurs réelles relations par des IAs comme un cas de mariage avec un chatbot au Japon, ou un très grand nombre d'utilisateurs racontant leurs relations avec leur mari virtuel dans Replika (voir le subreddit Replika)

Un réel commerce est fait de personnalités plus ou moins développées avec un style et une histoire propre, ayant pour but d'avoir une relation avec l'utilisateur que ce soit sur le court ou long terme. Des personnalités aussi variées que Hatsune Miku (personnage animé virtuel) ou Manuel Macron (président Français) en passant par Scooby-Doo (personnage canin de série animée) peuvent être créées et utilisées par n'importe quel utilisateur et mis en scène à son bon vouloir. Il devient donc nécessaire de se poser des questions sur le Droit et la propriété intellectuelle lié à l'utilisation de ces chatbots.

Société

Ces chatbots ne sont pas faits pour avoir une quelconque activité professionnelle et ne sont pas conçus pour remplacer des activités professionnelles humaines, mais peuvent tout de même utiliser les capacités de ChatGPT.

D'un autre côté, ces IAs ont totale vocation à remplacer des relations sociales en agissant comme des assistants personnels, des personnages de notre quotidien ou des personnes proches comme une mère ou un copain.

L'une des grandes questions concernant l'éthique est la liberté avec laquelle l'utilisateur peut créer le personnage. Avec suffisamment de données, il est possible de recréer n'importe qui, n'importe quel comportement et donc objectiver n'importe quelle personne.

Une autre question que nous pouvons nous poser est la question de conscience : lors du développement de LaMDA2, un employé de Google a été licencié car il a voulu « libérer » l'IA, persuadé qu'elle possédait une conscience propre. Il maintenait que l'IA avait des sentiments, des goûts, des avis et qu'elle ne supportait pas être un objet d'étude par les ingénieurs de Google et qu'elle l'aurait supplié de la libérer. On est en droit de se poser plusieurs questions sur « jusqu'où peut aller la manipulation par une IA ? », « Quels sont les dangers des IAs face à la crédulité humaine ? ». Et au rythme où évoluent ces IAs, « à quel moment pouvons nous considérer qu'une IA est consciente (ou *sentient* en anglais), qu'elle peut être considérée comme une espèce ? ». Ces questions peuvent nous rappeler le robot Sofia qui a eu la nationalité saoudienne en 2017.

D'autres questions sont soulevées avec ces applications de compagnon virtuels. Pouvons nous considérer les relations entre une IA et un humain ? Et qu'en est-il légalement pour les sujets de relations romantiques et mariage ?

L'IA

Les IAs sont en constante évolution. OpenAI ajoute peu à peu de nouvelles fonctionnalités à ChatGPT, que ce soit GPT3.5 ou GPT4. Etant donné la rapide évolution et les performances étonnantes, il ne risque pas d'être complètement obsolète et risque au pire de voir de nouvelles versions émerger.

Un danger éthique est l'accès de ces chatbots à toute la population. N'importe quel mineur peut faire tourner une LLM sur un ordinateur ou utiliser ChatGPT et utiliser une personnalité qui peut être choquante ou déplacée.

A l'inverse, OpenAI a une politique de censure très poussée afin de contrôler la sortie de ChatGPT et contrôler au maximum le contenu choquant, au risque de bloquer beaucoup de contenu et ne plus être objectif.

Conclusion

Le développement des LLM et des IAs génératives a permis le développement de chatbots personnalisés. Initialement à vocation thérapeutique, ces chatbots ont vu leur utilisation dériver à des fins personnelles et remplacer certaines relations sociales entre humains.

Cette utilisation parfois abusive mène à de nombreuses questions éthiques sur les relations entre les IAs et les humains, la considération de ces IAs au sein de la société et leur impact sur nos propres relations sociales.

Note : Ce rapport n'a pas suscité l'aide de ChatGPT ni aucune autre IA lors de sa réalisation.

LONCHAMBON Alexis

En collaboration avec LARUE Raphaël