

Docteur IA et Mr. Hyde : Quand l'intelligence artificielle veut remplacer les chercheurs

Nabia : Bonjour Guillaume Gralé.

Guillaume : Bonjour Nabia.

Guillaume : Oui, en réalité c'est une organisation américaine qui a fait une déclaration un peu tonitruante le 5 novembre. Elle s'appelle Future House. Elle a présenté Cosmos qu'elle présente un peu comme un agent IA scientifique qui veut bousculer les frontières de la recherche, capable de lire jusqu'à 1500 articles scientifiques et d'écrire 42 000 lignes de code par session, pas loin de l'équivalent de 6 mois de travail humain en une seule journée. Donc on voit de grosses promesses avec la promesse là aussi d'avancer – j'allais dire de découverte, mais c'est plutôt des avancées dans le domaine des neurosciences, des nouveaux matériaux, l'alliage et les polymères auto-adaptatifs par exemple, ou encore la génétique clinique.

Alors il faut faire très attention parce qu'il y a seulement un taux de fiabilité de 79 %. 79 % seulement de ces résultats sont reproductibles. Et alors ce n'est pas une ONG, elle a lancé, elle a fait un lancement commercial sous le nom d'Edison, du nom du célèbre inventeur américain qui a créé un peu l'ancêtre des caméras notamment au cinéma. Mais ce qui est assez intéressant, c'est que cette invention va reposer sur un modèle du monde structuré. L'objectif, c'est de gagner en cohérence tous les jours et de diviser par 10, par exemple, des études précliniques.

On sait que l'IA peut permettre, lorsqu'elle est bien maniée, de faire de véritables progrès. On avait déjà parlé ensemble d'AlphaFold, vous savez, le modèle mis au point par DeepMind et qui permet la modélisation du repliement des protéines et c'est utilisé par plusieurs centaines de milliers de chercheurs dans le monde avec des progrès notables en agriculture ou encore en médecine. Il y a également au Japon Sakana qui repose sur un programme de recherche avec l'université d'Oxford et qui veut automatiser l'ensemble du cycle de recherche, même le peer review. Donc c'est assez prometteur, mais il faut faire encore attention, c'est-à-dire qu'il y a encore des erreurs.

Nabia : Ah justement, alors il y a cette bonne nouvelle, mais on va nuancer. Est-ce que l'homme ne risque pas d'être dépossédé quand même de ses fonctions les plus essentielles ?

Guillaume : Oui, c'est ça, c'est le véritable danger. C'est l'IA qui peut aussi tricher un tout petit peu avec les évaluations. Vous savez que les lectures, les peer reviews, c'est extrêmement important dans les sciences et on a vu plusieurs chercheurs comme le Français Guillaume Cabanac avertir, faire attention à cette IA qui peut nous tromper. Il y a un projet que je trouve plus intéressant qui s'appelle Gnosis qu'on doit à Anthropic où là plusieurs IA vont débattre de résultats avant publication.

Mais effectivement, vous avez raison, Nabia, on peut s'inquiéter légitimement d'un bouleversement entre guillemets à très court terme. Selon l'Organisation internationale du travail, 11 % des tâches humaines dans les services pourraient être automatisées d'ici à 2030 et jusqu'à 38 % dans les métiers de bureau. Les jeunes diplômés américains, on en a déjà parlé, ils mettent aujourd'hui 7,3 mois pour décrocher leur premier emploi contre 4,8 mois avant la vague de l'intelligence artificielle que l'on connaît.

Bien sûr, il y aura des nouveaux emplois, c'est ce qu'expliquent PricewaterhouseCoopers ou encore McKinsey, mais on a encore du mal à avoir une idée du solde. Et il y a des applications assez surprenantes. J'ai pu assister il y a quelques semaines à la présentation de Steve Napoléon. Steve Napoléon, c'est une entreprise américaine qui s'appelle AI Potential. C'est une spin-off à la fois d'Adecco et de Salesforce. Steve Napoléon veut être ce petit chatbot qui veut parler à l'oreille des grands patrons pour les conseiller en permanence sur les choix qu'ils doivent faire. C'est à la fois déstabilisant. Le PDG explique : « J'ai eu pour la première fois face à moi un logiciel capable de me contredire avec raison », mais en même temps on a envie de dire « surtout ne leur donnons pas le pouvoir de décision à ces machines parce que sinon ça peut être très dangereux ».

Nabia : Et donc ne pas laisser l'IA nous endormir intellectuellement.

Guillaume : Oui, il y a eu une étude très intéressante du MIT que l'on doit à la chercheuse Natalia Modestova qui a expliqué que ChatGPT peut réduire l'activité dans les zones du cerveau liées au raisonnement ou encore à la mémoire. Les utilisateurs de l'intelligence artificielle compulsive – j'ai envie de dire sans avoir à réfléchir – vont manifester une diminution d'activité dans le cortex préfrontal ou encore l'hippocampe. Ce phénomène d'endormissement cognitif traduit une délégation de la réflexion. C'est-à-dire qu'on va basculer en mode passif.

Et le philosophe Éric Sadin qui vient de publier un livre, **Le Dé de nous-même**, parle d'une déprise cognitive. Gérald Bronner, lui, dans **L'Apocalypse cognitive**, ça pose beaucoup de soucis. Là, j'ai voulu montrer un film qui a été mis au point et vulgarisé par Future of Life Institute qui montre un peu une dystopie lorsqu'on fait tout reposer, toutes les décisions à la machine. L'homme est complètement perdu.

Et sur un autre front, celui des robots humanoïdes, on voit des progrès encore fulgurants.

Nabia : Oui, c'est ça. C'est un peu l'autre annonce de cette semaine. C'est pour ça je voulais en parler avec vous parce qu'on va voir que ça rejoint un petit peu notre problématique de l'intelligence artificielle.

Guillaume : Là, on va vers la Chine. L'entreprise Xpeng a mis au point Iron. Alors vous allez voir, c'est un robot humanoïde de 1,73 m, un poids d'environ 70 kg. Mais il a une apparence extrêmement humaine lorsqu'on le voit bouger, lorsqu'on le voit se déplacer. On a vraiment l'impression d'avoir une femme devant soi. Et on voit bien la rivalité qu'il y a entre la Chine et les États-Unis. Ça, c'est un peu la réponse chinoise à toutes les dernières annonces dans les humanoïdes.

On néo que l'on a vu, mais à qui on pouvait reprocher à raison d'être complètement téléopéré. Mais on a aussi vu Figure 02 qui savait plier du linge, réaliser des gestes de précision, Optimus plus stable ou encore ITRI. Donc là en Chine, capable d'exécuter des saltos arrière. C'est intéressant, mais c'est extrêmement troublant.

Nabia : Oui, et là je reviens juste à l'image, ce n'est pas un hologramme. Moi, quand vous m'en avez parlé au départ, j'ai pensé que c'était un hologramme, mais c'est vrai, c'est une vraie structure bionique là qu'on voit.

Guillaume : Oui, c'est une vraie structure bionique. D'ailleurs, à la demande des personnes qui assistaient à la présentation, on a littéralement désossé ce corps pour voir si ce n'était pas un humain à l'intérieur. Mais en réalité, on va avoir avec ces robots humanoïdes un retour, c'est-à-dire doter la machine d'un sens commun physique. Les données sensorielles, par exemple visuelles, tactiles, auditives, vont permettre en réalité d'avoir une compréhension intuitive du monde physique. Un petit peu ce qui manque, c'est souvent ce qu'a dit Yann LeCun en disant qu'un enfant de 4 ans avec les petites expériences qui vont lui arriver a une meilleure compréhension du monde réel que n'importe quel robot, n'importe quel modèle de langue.

Mais là, si on a le retour sensoriel de ces robots, on va avoir une intelligence artificielle sans doute beaucoup plus puissante à l'avenir et c'est là où on peut être un tout petit peu déstabilisé effectivement.

Nabia : Et pourquoi faut-il raison garder et ne pas transformer ce qu'il y a en religion ?

Guillaume : Oui. On voit dans toutes les déclarations de géants de la Silicon Valley cette croyance dans l'AGI, vous savez, cette intelligence artificielle générale, la promesse de cette IA qui pourrait dépasser l'homme dans la totalité de ses capacités cognitives. Effectivement, il y a plusieurs chercheurs, chercheuses comme Timnit Gebru qui disent : « Il faut faire attention à cette technoreligion ». On ne prend plus – enfin plus beaucoup en tout cas qu'en petite dose – du LSD, mais on va prendre souvent dans la Silicon Valley de l'ayahuasca. C'est une boisson qui repose sur un rite amazonien qui veut dire c'est une transformation intérieure avec une nouvelle conscience.

C'est-à-dire qu'il faut faire quand même assez attention avec le fait de trop vouloir donner à la machine. J'ai été cette semaine à la présentation de codeurs informatiques qui, eux, décident de faire appel à l'intelligence artificielle comme Noam Meduni qui est ambassadeur chez D Technology, mais toujours en gardant le contrôle. Parce que sinon, si on ne garde pas la tête froide, si on écoute Eliezer Yudkowsky par exemple qui est le fondateur du Machine Intelligence Research Institute, il a une vision quasi apocalyptique. On veut tout donner à la machine. Je pense qu'il faut là à chaque fois savoir ce qu'on lui confie et ça serait, je pense, de bon augure de faire la même chose en réalité avec ce que nous promet ce nouveau chercheur, ces nouveaux scientifiques qui reposent sur l'intelligence artificielle.

Nabia : Merci beaucoup Guillaume Gralé. Je vous laisse pas filer avant. Alors vous n'alliez pas le faire, je le fais pour vous. Présentez votre ouvrage pionnier **Voyage aux frontières de*

l'intelligence artificielle*. Vous l'avez écrit vous-même, pas avec l'IA entre chaque chronique.

Guillaume : Voilà, c'est un ouvrage qui est disponible depuis 48 heures maintenant. Vous me l'apportez à peine, je vais m'y plonger. Je vous recommande de faire pareil si l'intelligence artificielle et ce voyage aux frontières de l'IA vous intéresse. Merci beaucoup Guillaume. À la semaine prochaine.

Nabia : Ouais, merci beaucoup Nabia.

Guillaume : Merci à vous. Du coup, on se retrouve.