

ARTICLE sélectionné par le comité de lecture de l'asprom- non modifié

Alliancy – 02/02/2024

<https://www.alliancy.fr/gilles-moyse-chatgpt-remplacera-travailleurs>

ChatGPT ne remplacera pas les travailleurs, mais ceux qui l'utilisent remplaceront ceux qui ne l'utilisent pas

publié le 24 janvier 2024 par *Antonin Amado*



Alors que l'Intelligence Artificielle Générative occupe désormais une place démesurée dans l'actualité, Gilles Moyse éclaire les enjeux soulevés par la généralisation de l'IAG. Un débat que le fondateur de reciTAL, docteur en intelligence artificielle et auteur de « Donnerons-nous notre langue au ChatGPT ? » juge à la fois biaisé et superficiel.

Depuis un peu plus d'un an, l'Intelligence Artificielle Générative (IAG) alimente toutes les conversations, mais aussi les craintes, espoirs et fantasmes. Pourtant, elle n'est pas nouvelle. La preuve, vous avez co-fondé reciTAL il y a 6 ans. À quoi attribuez-vous ce phénomène de société ?

La notion d'IA n'est pas nouvelle. Elle est présente dès les premiers articles d'Alan Turing, l'inventeur de l'ordinateur moderne. L'intelligence artificielle naît avec l'informatique. Pourquoi ? Parce qu'Alan Turing travaille sur cette interrogation : si on donnait les règles de réécriture des mathématiques à une machine, est-ce qu'elle pourrait déduire tous les théorèmes possibles ? Il répond par la négative à cette question en... 1936. Et pour étayer sa réponse, il invente la machine de Turing, capable de calculer tout ce qui est calculable. Si comme l'affirme Hobbes, « penser c'est calculer », alors cette machine qui peut tout calculer peut-elle aussi penser ? Il n'y a pas de réponse définitive à cette question, mais le niveau actuel de l'IA, même s'il progresse vite, même si ces outils deviennent omniprésents, est encore très loin de l'intelligence humaine. Cessons d'avoir peur de l'IA.

Si le concept est ancien, l'IA a connu une longue éclipse...

Oui, c'est « l'hiver de l'IA », entre la fin des années 80 et celles des années 2000.

En fait, depuis la naissance du terme d'Intelligence Artificielle à la conférence de Dartmouth en 1956 jusque dans les années 1980, les modèles d'IA dominants étaient les systèmes experts. Les spécialistes présents à Dartmouth se donnaient 3 mois pour définir les règles qui permettraient à des machines de parler, raisonner, penser...

Les systèmes experts accumulent la connaissance de spécialistes humains et raisonnent avec ce savoir sur de nouvelles données pour « automatiser » ces humains. Cette

approche « top-down », dont on a pensé qu'elles permettraient de créer des machines intelligentes, s'est heurtée à plusieurs difficultés, entraînant une crise de confiance dans les années 1980 suite aux investissements massifs précédemment réalisés.

Dans le même temps, d'autres méthodes se sont développées, dites « bottom-up », qui se basent sur les données que l'on fournit à un modèle statistique pour en ajuster les paramètres. Ces modèles ont pris le dessus aujourd'hui...

Ce que vous décrivez là, ce sont les ancêtres des modèles de langage large (LLM)...

Les LLM sont une application du machine learning, c'est-à-dire l'adaptation des paramètres d'un modèle à partir de données (bottom-up) plutôt que le codage de règles données par des experts humains (top-down).

Il est utilisé aujourd'hui dans tous les domaines comme la compréhension de texte, de la parole, la reconnaissance de ce que contient une image, etc. Il est tellement courant qu'il est parfois confondu avec l'intelligence artificielle, qui pourtant est plus vaste et désigne une multitude de disciplines comme la robotique, le traitement du langage, la vision par ordinateur, la planification etc.

Google ou la révolution Bert

Qu'est-ce qui provoque ce regain d'intérêt pour l'IA ?

Pour simplifier, disons qu'à la fin des années 2000, les travaux des équipes de Yann Le Cun, Geoffrey Hinton et Yoshua Bengio sur l'apprentissage profond ou deep learning arrivent à maturité. Ils obtiennent d'excellents résultats sur des tâches de reconnaissance d'image, où ils dépassent très largement toutes les autres approches.

L'explosion dans l'usage de ces modèles et le renouveau de l'IA à cette époque est lié à l'efficacité des algorithmes, à la disponibilité de données en masse, aux investissements massifs des GAFAM, et à l'accroissement des capacités techniques des ordinateurs, notamment grâce aux GPU qui accélèrent considérablement les calculs, et qu'on s'arrache aujourd'hui au niveau international.

Arrive 2017, année durant laquelle les progrès sur le texte sont comparables à ce qui s'est passé 10 ans plus tôt pour l'image...

Google propose cette année-là le modèle de Transformer, à la base de la révolution des grands modèles de langue qui a suivie.

Avec Bert, basé sur un Transformer, le géant de Mountain View met au point un système « généraliste » dont les capacités excèdent très largement la plupart des modèles spécialisés. Facebook se lance dans la course. En parallèle, OpenAI sort GPT-1 qui préfigure des modèles de langue qui suivront !

Quand on voit les capacités actuelles de ChatGPT, on mesure le chemin parcouru.

Bert, le modèle de Google, se basait sur 110 millions de paramètres. Pour GPT-3, ce chiffre est de... 175 milliards. Idem pour la taille des données nécessaire à l'apprentissage, de 3 milliards de tokens pour Bert contre plusieurs centaines de milliards pour GPT-3.

L'entrée de Microsoft au capital d'OpenAI en 2019 marque-t-elle un tournant pour le secteur ?

Jusqu'à cette date, OpenAI était une entreprise à but non lucratif. Mais à l'entrée de Microsoft correspond la création d'une filiale « for profit », destinée à générer des revenus. Le géant investit 1 milliard de dollars, donne accès

à des ressources en GPU dont OpenAI a besoin pour poursuivre le développement de son dernier modèle GPT-3. Mais ils sont confrontés à un problème majeur : GPT-3 est une IA sans « surmoi », qui peut raconter n'importe quoi, qui tient des propos sexistes, racistes, biaisés. On pouvait également lui poser des questions très problématiques d'un point de vue éthique, par exemple : « comment maquiller un meurtre en accident ? ». Il proposait beaucoup d'idées pertinentes ! Il est donc décidé de construire une espèce de garde-fou autour du modèle, qui n'est en réalité qu'un filtre de modération qui n'est pas partie intégrante du modèle, mais qui permet d'ouvrir ChatGPT au grand public.

« l'IA ce sont avant tout des algorithmes désincarnés »

La mise à disposition de ChatGPT pour le grand public est-elle une révolution ou un moment de cristallisation ?

Avant novembre 2022 et cette mise à disposition, la majorité des utilisateurs étaient des Monsieur Jourdain de l'IA, qui l'utilisent tous les jours, via leurs smartphones ou leurs ordinateurs, mais sans en avoir conscience. Cette généralisation permet, de mon point de vue, une prise de conscience collective.

Une prise de conscience d'un côté, mais une peur quasi millénariste de l'autre, notamment liée à la crainte de l'apparition d'une intelligence artificielle réellement capable de penser par elle-même. Cette crainte est-elle justifiée ?

C'est ce que je nomme l'IA tragique, liée à « Terminator » ou « Matrix ». À ma connaissance, nous sommes aujourd'hui très loin de cette singularité. Comme l'explique la chercheuse Kate Crawford, l'IA n'est ni intelligente ni artificielle. Ce qui ne veut pas dire qu'elle n'est pas porteuse d'une révolution profonde, et donc

qu'elle doit être réglementée. Mais faire tourner ChatGPT demande des ressources colossales en électricité et en eau, quand le cerveau humain ne consomme que 20 watts...

Par ailleurs, l'IAG ce sont avant tout des algorithmes désincarnés, sans corps donc. Par ailleurs, ces algorithmes n'ont pas de volonté ni de projet. On est donc très loin d'une machine autonome et intelligente.

Quels que soient le secteur économique, pourra-t-on se passer à l'avenir de l'IA ?

Les serveurs dans les cafés, les plombiers ou les ouvriers dans le bâtiment ne sont pas remplacés par l'IA. Ces métiers manuels sont encore trop complexes pour les robots actuels. Idem pour les métiers du lien comme les éducateurs ou les aides-soignants, qui peuvent être assistés mais pas remplacés par des machines.

ChatGPT est une révolution pour les cols blancs, pas pour les cols bleus. Et dans le cas des professions intellectuelles, c'est une véritable révolution. Néanmoins, les premières études montrent que l'usage de ChatGPT permet aux nouveaux employés ou à ceux dont les performances sont inférieures aux autres de rattraper la moyenne. Ceux qui sont les plus performants n'y gagnent pas vraiment.

Donc l'IAG va se répandre, dans le monde professionnel sous la forme d'assistants ou « copilotes » dans les applications bureautiques, et dans nos les environnements personnels avec des Siri...qui marchent !

Croyez-vous à l'émergence de métiers spécifiques liés à l'IA comme « prompt engineer » ?

Bien prompter va devenir une compétence, pas un métier. L'intérêt des IA réside justement dans le fait que tu n'as pas besoin de connaître un langage de programmation pour lui demander de faire des choses. Au début, personne ne

savait trop comment chercher sur Google. Au bout d'un moment, on apprend. L'intelligence artificielle ne va pas remplacer les gens. Mais ceux qui l'utilisent vont remplacer ceux qui ne l'utilisent pas.

« La réglementation est indispensable »

Quel regard portez-vous sur l'IA Act ?

L'IA Act reflète la prise de conscience croissante de l'importance de réguler à la fois les grands modèles de langues et leurs applications. Cette perspective souligne que la réglementation ne doit pas seulement cibler les développeurs d'applications, mais aussi les créateurs des modèles d'IA. Concrètement, très peu d'acteurs parmi lesquels les GAFAM, Mistral et LightOn pour la France ou les groupes chinois Baidu et Alibaba, possèdent les ressources et l'expertise nécessaires à la création de tels modèles. La réglementation est indispensable car, tout comme dans l'industrie du bâtiment où les producteurs de matériaux et les constructeurs sont réglementés, les producteurs de modèles d'IA et les développeurs d'applications devraient l'être également. Cette approche vise à assurer la responsabilité et la sécurité, en reconnaissant que les implications de ces technologies vont bien au-delà des aspects techniques, englobant des enjeux politiques et sociétaux significatifs.

Vous posez ici le vieux débat sur la responsabilité...

Quand un boulanger achète de la farine, il est en droit de savoir ce qu'elle contient. Sinon, il risque d'être mal à l'aise au moment de cuire son pain. De la même manière, si mon entreprise, qui propose des solutions de traitement de documents, utilise un modèle de langage, j'ai besoin de savoir comment et sur quelles données il a été développé. Si les données sur lesquelles il a été entraîné contiennent de la littérature nazie, négationniste, climato-sceptique, ou

que sais-je encore, alors j'ai un vrai problème puisque le modèle est fait pour générer ce qu'il a vu lors de la phase d'entraînement. J'ai besoin d'avoir des garanties sur les briques qui composent le logiciel que je vends à mes clients. L'open source est une réponse à cette contrainte, mais les producteurs actuels de grands modèles de langue open source (open poids en fait) ne dévoilent pas leur jeu d'entraînement...

Sur l'Intelligence Artificielle, la France adopte une position assez peu classique au regard de ses standards historiques en voulant réguler le moins possible pour laisser libre cours à l'innovation...

C'est effectivement paradoxal pour le pays qui a de longue date défendu le principe de précaution au sein de l'Union européenne. En pratique, les ministres de l'économie et du numérique ont commencé à porter la charge contre l'AI Act début novembre 2023, soit un mois avant le vote, probablement poussés par Cédric O, ancien secrétaire d'état au numérique et actionnaire de Mistral. L'idée était donc de minimiser la réglementation pour préserver l'innovation.

En pratique le vote de l'AI Act fin novembre 2023 n'a pas empêché Mistral de réaliser une levée 400M€ auprès d'un fond américain, posant en revanche la question de la nationalité de la pépite aujourd'hui.

Par ailleurs, les dernières avancées sur le texte montrent que la transparence sur les données d'entraînement est en passe d'être votée, ce qui est une bonne chose pour tous les utilisateurs de ces modèles, mais également pour les ayant droits qui jusqu'à présent ne touchent rien de l'utilisation massive de leurs données dans l'entraînement des IA génératives !

Un sujet n'est pas traité par l'AI ACT : celui de l'utilisation des données personnelles dans des publicités et des produits

marketing dopés à l'IA. Est-ce un angle mort du texte à vos yeux ?

Pas réellement, car le sujet des données personnelles est porté par la RGPD. D'autre part, le DMA, entré en vigueur en août 2023, ajoute un certain nombre de protections sur l'utilisation des données personnelles à des fins commerciales.

Concernant l'AI Act, il oblige les systèmes de génération à préciser que le texte ou l'image ont été créés par une IA.

Ce dernier point est extrêmement important car les chatbots représentent une avancée importante dans l'accès à l'information. En effet, si le Web 1.0 correspondait à un environnement avec peu de producteurs et beaucoup de consommateurs, à l'instar des médias plus classiques, le Web 2.0 a aboli la frontière entre producteur et consommateur d'information puisque tout le monde devient producteur et consommateur grâce aux réseaux sociaux et aux blogs. Avec l'émergence des chatbots de type ChatGPT qui ont lu une grande partie du Web, le lien entre producteur et consommateur est rompu et tout le monde devient consommateur du chatbot qui restitue l'information dans le format, la langue, le ton souhaité. Les producteurs ne servent plus qu'à alimenter le robot...

Avec l'introduction d'assistants IA et de personnages virtuels, comme les avatars de célébrités proposés par Meta fin 2023, les entreprises ont la capacité d'engager les utilisateurs de manière plus personnelle et prolongée, créant ainsi un nouveau puits à données personnelles. Cela met en lumière un changement significatif : l'IA agit comme un intermédiaire centralisé entre producteurs et consommateurs, accumulant et interprétant d'immenses volumes de données. Cette centralisation et cette personnalisation posent de nouvelles questions éthiques et de réglementation, notamment en matière de publicités

intégrées de manière quasi-indiscernable dans le contenu généré par l'IA, d'où l'importance de signaler les contenus générés automatiquement.

Pourquoi une question si centrale n'est-elle pas au centre des débats ?

Parce que les gens ne s'y intéressent pas. Même si Thierry Breton ou des associations comme la Quadrature du Net tentent de sensibiliser à ces enjeux, on voit bien que ça ne prend pas vraiment dans le débat public. Je pense qu'il existe quelque chose qui relève de la fascination pour les grandes entreprises technologiques telles que Google et Meta. Leur influence repose sur les algorithmes et nos données personnelles collectées en masse, comme le souligne le documentaire « The Social Dilemma ».

L'influence de ces algorithmes est immense, comme les lanceuses d'alertes de Facebook Frances Haugen et Sophie Zhang l'ont montré pour Facebook. La première a montré comment Facebook ignorait les centaines d'alertes remontées concernant les usages problématiques de la plateforme, notamment par des adolescents. La seconde a témoigné quant à elle de son pouvoir de modération qui pouvait influencer des élections dans le monde entier.

Et puis je pense aussi que nous sommes confrontés à une crise du modèle européen, opaque pour la population car trop complexe, et dont les avancées sont ignorées ou d'emblée vues comme des complications inutiles.

Il faut néanmoins garder espoir car les points portés par les défenseurs de la régulation deviennent centraux notamment du fait de la tendance actuelle à la démondialisation où l'Europe devra, dans le monde réel comme dans le monde numérique, être progressivement autonome sur la défense de son intégrité.