

CHATGPT : DERRIÈRE LE MYTHE DE L'INTELLIGENCE ARTIFICIELLE, LES DANGERS DES AUTOMATES COMPUTATIONNELS



crédits JI-korpa

Le 14 mars 2023, OpenAI, entreprise américaine fondée entre autres par Elon Musk, Sam Altman et Peter Thiel, annonçait le lancement de son nouveau modèle de langage artificiel, GPT-4, faisant passer une étape supplémentaire aux technologies de génération automatique de textes. Cette nouvelle innovation poursuit la stratégie industrielle de l'entreprise américaine, qui ne cesse de faire parler d'elle depuis quatre mois.

Derrière la fascination, il convient de mettre au jour les fantasmes et les dangers : ces technologies ne risquent-elles pas de réduire l'avenir du web à l'extractivisme de la donnée, et les capacités expressives humaines à des calculs automatisés ? Par Anne Alombert, maître de conférences en philosophie française contemporaine à l'université Paris 8, Vincennes – Saint-Denis, autrice de *Schizophrénie numérique* à paraître en avril 2023 aux éditions Allia.

Le 30 novembre 2022, OpenAI réalisait l'un des plus grands « coup médiatique » de l'histoire de l'intelligence artificielle, en rendant accessible au grand public un nouveau outil, dénommé ChatGPT. Ce dispositif, souvent décrit sous les noms d'« intelligence artificielle » ou d'« agent conversationnel », repose en fait sur l'association entre deux types de technologies : un modèle de langage (il s'agissait alors de GPT-3, et désormais, possiblement, de GPT-4), qui permet de générer toutes sortes de textes grâce à des calculs probabilistes appliqués à des quantités massives de données, et une messagerie instantanée, qui permet aux utilisateurs de poser des « questions » ou de donner des « consignes » à la machine, afin d'obtenir des productions textuelles variées de la recette de cuisine aux poèmes, en passant par les articles scientifiques ou le code informatique).

RECEVEZ NOS DERNIERS ARTICLES

e-mail JE M'ABONNE

CHATGPT, ENTRE FANTASMES ET RÉALITÉS

Comme souvent, le fait de rendre accessible ce type de dispositifs, au fonctionnement opaque mais à l'usage très intuitif, engendre un effet de fascination généralisée : la machine, semblant répondre par magie aux questionnements les plus insolites, devient l'objet de toutes sortes de projections anthropomorphiques nourries par les discours transhumanistes. La machine serait-elle capable d'écrire ou de parler ? Serait-elle capable de penser ? Serait-elle consciente d'elle-même ? Pourrait-elle se substituer aux activités humaines qu'elle semble si bien simuler ?

Le lancement de ChatGPT (d'abord gratuit mais dont une version payante a déjà émergé) a ainsi constitué une énorme opération de communication pour OpenAI, qui en a aussi profité pour exploiter le travail gratuit de ses millions d'utilisateurs (cent millions de comptes enregistrés en janvier 2023), convaincus d'expérimenter un nouveau gadget, alors que c'est le système qui se nourrissait de leurs requêtes pour améliorer les performances de ses algorithmes. Si bien qu'au moment où tout un chacun commençait à se demander comment faire un bon usage de ChatGPT, le dispositif technologique avait déjà largement utilisé ses soi-disant usagers pour se perfectionner, conformément au célèbre adage selon lequel « si c'est gratuit, c'est vous le produit ! ».

FAITES UN DON DÉFISCALISÉ

helloasso

« Au moment où tout un chacun commençait à se demander comment faire un bon usage de ChatGPT, le dispositif technologique avait déjà largement utilisé ses soi-disant usagers pour se perfectionner, conformément au célèbre adage selon lequel "si c'est gratuit, c'est vous le produit !" ».

Grâce à ChatGPT, OpenAI, qui avait bénéficié d'un investissement initial d'un milliard de dollars de la part de Microsoft, a ainsi vu sa valeur boursière augmenter de manière exponentielle : désormais valorisée à 29 milliards de dollars, elle devrait faire l'objet d'un investissement pluriannuel de plusieurs milliards de dollars de la part de Microsoft, qui entend intégrer ChatGPT à ses services (notamment à son moteur de recherche Bing, dans le but de concurrencer Google en ce domaine, et à la suite bureautique Microsoft Office). Le 23 janvier 2023, Microsoft annonçait ainsi le renforcement de son partenariat avec OpenAI « pour accélérer les percées de l'IA afin de garantir que ces avantages soient largement partagés avec le monde ».

Trois jours plus tôt, le 19 janvier, Microsoft avait annoncé le licenciement de dix mille salariés (près de 5 % de ses effectifs). La veille, le 18 janvier, le journal *Time* avait révélé le partenariat entre OpenAI et la société Sama, qui a permis à l'entreprise d'employer des travailleurs kenyans rémunérés à moins de deux dollars de l'heure pour indexer d'immenses quantités de contenus toxiques circulant sur Internet, afin de « nettoyer » les données d'entraînement de ChatGPT. Les calculs des algorithmes, indifférents à toute signification, sont en effet incapables de reconnaître un texte à caractère haineux, violent, raciste ou pornographique : pour éviter que la machine ne génère automatiquement et massivement des contenus de ce type, ce sont donc des humains qui doivent les visionner et les indexer, au prix de nombreux traumatismes. Contrairement aux IA qui effectuent des calculs sur de l'information numérique, les humains se laissent affecter par le sens des textes qu'ils lisent : l'exposition constante à des contenus nuisibles engendre donc des troubles psychiques.

Plutôt que les machines conscientes ou les avantages de l'IA « largement partagés avec le monde », ce sont donc les craintes associées à l'industrie numérique et à l'économie des données qui semblent s'affirmer derrière les prouesses de ChatGPT : disparition des emplois et exploitation des « travailleurs du clic », sans compter la concentration des capitaux et la consommation énergétique nécessaires à l'entraînement de ce type de modèles, qui rend le coût de ces technologies prohibitif.

« Il faut en effet distinguer un dispositif comme Wikipédia et un dispositif comme ChatGPT : les deux outils symbolisent en fait deux visions antithétiques du Web, la première fondée sur l'ouverture et l'intelligence collective, et la seconde sur l'extractivisme et le calcul automatisé ».

DES AGENTS CONVERSATIONNELS AUX AUTOMATES COMPUTATIONNELS

Comme souvent, les enjeux écologiques, politiques et sociaux des innovations technologiques sont subrepticement refoulés par des discours au sujet des « intelligences artificielles » ou des « agents conversationnels ». Or, comme cela a été souligné par divers spécialistes, un dispositif comme ChatGPT n'a rien d'intelligent : il s'agit d'un ensemble complexe d'algorithmes (très performants) effectuant des calculs statistiques sur des quantités massives de données afin de déterminer des suites de mots probables par rapport aux données d'entrée (sur le modèle des logiciels d'autocomplétion, qui permettent de générer automatiquement la suite la plus probable d'un sms ou d'un e-mail).

Ce n'est évidemment pas ainsi que procèdent les « agents » lors de « conversations » : converser avec autrui ne suppose pas de faire des calculs probabilistes sur des quantités massives de données afin de générer une suite probable de mots (sans qu'aucun humain ne serait capable de converser). En conversant, un « agent » s'expose à l'autre en révélant sa singularité : tant qu'il ne se contente pas de répéter des idées reçues ou des clichés, il s'exprime à partir de sa mémoire, de son expérience et de ses désirs toujours singuliers, en interprétant une situation ou un contexte donné. Les conversations les plus intéressantes sont souvent les plus inattendues et les plus improbables, celles durant lesquelles l'agent se surprend lui-même à travers les questions qu'il pose ou à travers les réponses qu'il propose – la conversation se fait ainsi conversion. Le dialogue ne se réduit pas à un système entrée-sorties ou à un traitement d'informations, mais suppose une transformation mutuelle des participants.

« Plutôt que de se demander si nous devons accompagner ou interdire ChatGPT, il s'agirait alors de s'interroger sur les dispositifs numériques au service de l'intelligence collective et porteurs d'avenir pour les sociétés ».

C'est seulement à cette condition que de « l'intelligence » peut émerger : l'intelligence, comme capacité à se confronter à l'autre, à se mettre en question et à produire de la nouveauté, n'est pas plus attribuable à un agent isolé qu'à une machine déterminée, à moins de la confondre avec le quotient intellectuel ou avec le calcul automatisé. Loin de désigner une capacité individuelle (propriété cognitive attribuée à l'humain ou propriété algorithmique attribuée à la machine), l'intelligence désigne avant tout une forme de socialité – « être en bonne intelligence » signifie partager une sorte d'entente, de complicité ou de connivence (sur la base de laquelle des désaccords et des discussions peuvent bien sûr émerger). Plutôt que d'intelligences artificielles ou d'agents conversationnels, il faudrait donc parler d'automates computationnels pour désigner les modèles de langage du type de GPT.

D'apparence anodine, ce changement de vocabulaire permettrait peut-être d'éviter les fantasmes transhumanistes de Singularité Technologique, selon lesquels la croissance technologique conduirait à une superintelligence artificielle qui viendrait se substituer aux intelligences humaines. Jamais les calculs statistiques des algorithmes ne pourront remplacer les activités interprétatives de l'esprit, incarnées dans des corps vivants et désirants. En revanche, ce sont bien les singularités expressives qui sont menacées par ces industries linguistiques, dont le fonctionnement renforce les moyennes et les banalités, exploite les productions textuelles existantes sans les renouveler, tout en risquant de court-circuiter les capacités de mémorisation, de synthèse et de pensée.

DE WIKIPÉDIA À CHATGPT : DE L'INTELLIGENCE COLLECTIVE AU CALCUL AUTOMATISÉ

Rien d'étonnant, dès lors, à ce que les textes générés par ChatGPT ne soient pas dignes de confiance ou de fiabilité : contrairement aux articles collectivement produits sur une plateforme comme Wikipédia, ils n'ont été ni évalués ni certifiés par les pairs, et ils n'ont pas fait l'objet de débats argumentés. Il faut en effet distinguer un dispositif comme Wikipédia et un dispositif comme ChatGPT : les deux outils symbolisent deux visions antithétiques du Web, la première fondée sur l'ouverture et l'intelligence collective, et la seconde sur l'extractivisme et le calcul automatisé.

Wikipédia fonctionne sur la base de contributions humaines, à travers lesquelles les individus partagent leurs savoirs, exercent leur jugement et délibèrent collectivement, sur la base de règles transparentes et accessibles à la communauté, qui constituent ainsi des groupes de pairs ouverts, produisant un commun numérique au service de l'humanité. ChatGPT, au contraire, appartient à une entreprise privée et ne donne pas accès à ses normes de fonctionnement interne à ses usagers : ceux-ci sont d'ailleurs réduits au statut de consommateurs passifs, pouvant saisir des entrées et recevoir des sorties, mais sans possibilité de compréhension ni pouvoir d'action sur les mécanismes régissant le dispositif.

Dès lors, contrairement aux articles contributivement élaborés sur Wikipédia, les textes automatiquement générés par ChatGPT ne constituent pas des savoirs collectifs, mais des informations statistiquement glanées sur Internet dans une logique extractiviste. Sans compter la répétition virale des fausses informations dominantes ou la génération automatique de références scientifiques inexistantes, ces technologies présentent aussi le risque de dissimuler des erreurs aberrantes au milieu d'un texte vraisemblable et correct, erreurs que seuls les spécialistes du domaine pourraient repérer, mais que les autres utilisateurs ne pourront pas distinguer.

Que se passera-t-il quand ces textes automatiques (qui peuvent être générés en quantité massive et à une vitesse inouïe) seront devenus dominants sur les réseaux numériques ? Sera-t-il encore possible de faire crédit à nos lectures en ligne ? Que se passera-t-il si les moteurs de recherches nous fournissent des réponses générées de manière automatique ? Sera-t-il encore possible d'accéder à la diversité des sources disponibles ? Le Web pourra-t-il devenir l'espace de partage des savoirs qu'il avait vocation à constituer si la diversité des points de vue n'est plus représentée et si les publications font l'objet d'un crédit généralisé ?

Dès lors, ne serait-il pas nécessaire de s'interroger sur des dispositifs numériques alternatifs, mettant les algorithmes au service des activités d'interprétations, de jugements et de réflexion ? Qu'il s'agisse d'encyclopédies contributives, de logiciels d'annotations ou de plateformes collaboratives pour l'évaluation et la recommandation de contenu, de tels dispositifs existent, mais font rarement l'objet d'une promotion médiatique ou d'un soutien économique. Ils mériteraient pourtant d'être mis au cœur du débat public, si l'on ne souhaite pas constater dans quelques années les dégâts engendrés par la généralisation des technologies du type de ChatGPT, comme on constate aujourd'hui la nocivité des réseaux (anti-)sociaux, dont certains des créateurs se sont d'ailleurs repentis. Plutôt que de se demander si nous devons accompagner ou interdire ChatGPT, il s'agirait alors de s'interroger sur les dispositifs numériques au service de l'intelligence collective et porteurs d'avenir pour les sociétés.

Derrière leur aspect ludique et magique, les innovations disruptives promues par les entreprises de la Silicon Valley apportent aussi leur lot de dangers « déployés à une échelle massive et à une vitesse extrême, sans concertation avec les citoyens, les chercheurs ou les représentants politiques, ces dispositifs ne peuvent faire l'objet d'une appropriation collective, cultivée et critique. Pour qu'une telle appropriation devienne possible, ces nouveaux objets industriels, actuellement fermés et privatifs, devraient devenir des communs numériques ouverts et partagés. Ils devraient aussi constituer des objets de recherche transdisciplinaire et de débats publics, et non des sources de fantasmes idéologiques au service du marketing stratégique des géants du numérique.

Anne Alombert

27 mars 2023

PARTAGER L'ARTICLE

TÉLÉCHARGER L'ARTICLE

Suivez-nous

Facebook Twitter YouTube Instagram LinkedIn

FAITES UN DON DÉFISCALISÉ

helloasso

RECEVEZ NOS DERNIERS ARTICLES

e-mail JE M'ABONNE

RE JOIGNEZ LA RÉDACTION !