



HAL
open science

Le bluff technologique de ChatGPT

Adrien Tallent

► **To cite this version:**

| Adrien Tallent. Le bluff technologique de ChatGPT. Revue Esprit, 2023. <hal-04609736>

HAL Id: hal-04609736

<https://hal.science/hal-04609736v1>

Submitted on 12 Jun 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Copyright - All rights reserved

Le bluff technologique de ChatGPT

par

[Adrien Tallent](#)

Paru dans la revue Esprit, mai 2023 < <https://esprit.presse.fr/actualites/adrien-tallent/le-bluff-technologique-de-chatgpt-44669> >

Le succès de l'intelligence artificielle ChatGPT risque d'accentuer la dilution de la vérité au profit d'un régime de « post-vérité ». De plus, au nom du « bluff technologique » théorisé par Jacques Ellul, cette expansion sans limites de la technologie transforme l'individu en automate obéissant.

Après DALL-E en 2021, l'entreprise américaine OpenAI a lancé l'intelligence artificielle conversationnelle ChatGPT en novembre 2022. L'engouement a été tel que le site est depuis régulièrement indisponible – à moins d'opter pour la version payante – pour permettre à tout un chacun d'essayer de converser avec cette intelligence artificielle. Depuis, les commentaires, tribunes et émissions spéciales ont envahi l'espace médiatique. Cet engouement provient en partie d'une fascination pour les robots, les algorithmes, l'intelligence artificielle et la technologie de manière générale. Nous voyons dans ces « objets » des êtres parfaitement infailibles, rationnels et désormais également des êtres créatifs. Si certains appellent à accueillir une révolution intellectuelle majeure, il convient de pointer certains dangers potentiels et de démystifier cet outil technique.

Déléguer aux machines

ChatGPT a fait preuve de capacités relativement impressionnantes dans de nombreux domaines touchant la création et la rédaction de contenus. La nouveauté majeure est la cible qu'il menacerait : les métiers de cols blancs. On pensait, jusqu'à présent, que les métiers affectés par le déploiement de l'intelligence artificielle seraient les métiers manuels, qui seraient les premiers – voire les seuls – à être automatisés. Le paradoxe de Moravec (du nom de Hans Moravec, chercheur en robotique) exprimait, dès les années 1980, l'idée que plus une tâche est simple à réaliser pour l'individu, plus elle est difficile à automatiser, notamment parce que le cerveau réalise une tâche qu'il maîtrise parfaitement sans en avoir conscience. Contre-intuitivement, automatiser une tâche demandant un raisonnement complexe est plus facile à faire.

Le développement de machines conversationnelles crée cependant un risque épistémique majeur. ChatGPT produit des textes grammaticalement parfaits et souvent pertinents, même sans avoir aucune conscience de ce qu'il produit... Dès lors, il devient de plus en plus difficile de différencier un discours produit par une machine d'un discours humain. Cette apparence du vrai peut alors être la source d'importants risques quant à la diffusion de *fake news* ou d'arnaques¹. Il est à parier que, dans le futur, nos interactions concerneront de plus en

¹ Voir le dossier "Artificial Speakers – Philosophical questions and implications", sous la dir. de Hendrik Kempt, Jacqueline Bellon et Sebastian Nähr-Wagener, *Minds and Machines*, vol. 31, no 4, décembre 2021.

plus d'agents artificiels. Il sera alors primordial de pouvoir différencier nos interlocuteurs afin d'adapter le niveau de confiance que nous leur accordons.

Une course est désormais engagée. Microsoft a décidé d'investir dix milliards de dollars dans l'entreprise mère de ChatGPT, OpenAI, afin d'intégrer ChatGPT dans son moteur de recherche Bing, tandis que Google et Facebook développent leur propre système... Or ChatGPT est un système fermé et les données ayant permis sa mise au point ne sont pas publiques. Si l'information devient un circuit fermé – la réponse de ChatGPT se veut unique, contrairement à un ensemble de liens proposés après une recherche Google –, quelle confiance pouvons-nous lui accorder ? ChatGPT reproduit facilement des discours erronés suivant les questions posées et pourrait à terme se transformer en « *super-propagateur* » de fausses informations². Comment assurer, dans ces conditions, un accès équitable et universel à une même information sourcée ? L'irruption et la diffusion de ces outils (on pourrait ajouter la technologie derrière les *deepfakes* ou les algorithmes de recommandations sur les réseaux sociaux) accentuent le phénomène de dilution de la vérité factuelle au profit du régime de « post-vérité ».

Le risque est qu'en se perfectionnant, de tels outils soient de plus en plus présents dans notre quotidien sans que nous n'interrogeons plus leur fonctionnement. Heidegger décrivait déjà comment l'élément technique, « l'outil », disparaissait dans la quotidienneté et finissait par aller « *de soi* »³. Les outils techniques deviennent omniprésents dans notre quotidien ; le mouvement se fait en douceur, sans même que l'individu s'en aperçoive, rendant ainsi possible son adhésion à la course en avant technique.

Ignorant leur fonctionnement, l'individu se retrouve dès lors à leur merci. Dans *L'Immortalité*, Milan Kundera fait de Goethe le point de bascule entre deux époques. Jusqu'au début du XIX^e siècle, il était ainsi possible pour un individu cultivé de comprendre les objets techniques qui l'entouraient. Dans la période suivante, qui se poursuit de nos jours, une telle compréhension est inenvisageable⁴. L'idée des Lumières était que la technologie serait vectrice de science, qu'elle serait pédagogique. Ainsi, *L'Encyclopédie* intégraient de nombreuses entrées, de nombreux schémas destinés à la compréhension des outils techniques de l'époque. Or, aujourd'hui, plus un objet technique est complexe dans sa conception scientifique et technique, moins nous en connaissons les principes. Au-delà du niveau de complexité des objets techniques qui a augmenté, c'est aussi le rythme de création qui, s'intensifiant, est devenu impossible à suivre pour une seule mémoire humaine.

Prendre du recul

Depuis que de plus en plus d'universités ont été confrontées à des cas de copies d'étudiants rédigées grâce à ChatGPT, une course à la surenchère s'est enclenchée entre des systèmes techniques qui proposent de détecter un texte généré par une intelligence artificielle et d'autres qui proposent de rendre ces textes indétectables par les premiers. Si ce jeu du chat et de la souris est une réponse compréhensible, il serait préférable de reconsidérer notre rapport à de tels outils et d'adapter notre manière d'éduquer. D'autant plus que, de l'autre côté de la

² Voir Jack Brewster, Lorenzo Arvanitis et McKenzie Sadeghi, « ChatGPT pourrait-il devenir un super propagateur d'infox ? » [en ligne], *Newsguard Misinformation Monitor*, janvier 2023.

³ Martin Heidegger, *Être et Temps* [1927], trad. par François Vezin, Paris, Gallimard, coll. « Bibliothèque de philosophie », 1986, § 12-18.

⁴ Milan Kundera, *L'Immortalité* [1990], trad. par Eva Bloch, préface de François Ricard, Paris, Gallimard, coll. « Folio », 2020.

salle de classe, des algorithmes sont utilisés pour noter les élèves depuis plusieurs années déjà.

La nature des relations entre les êtres humains et les machines est souvent réduite de manière simpliste à la capacité de ces dernières à s'exprimer de manière familière et claire pour nous. Si de telles innovations peuvent représenter des avancées importantes, une prise de recul est nécessaire pour approcher le phénomène.

Une fascination intacte

Penseur techno-critique, Jacques Ellul évoquait un homme moderne fasciné, hypnotisé par la technique⁵. Sa position de théologien l'avait conduit à une réflexion particulière : le monde occidental a, en réalité, toujours été obnubilé par la question du sacré. Dès lors, suite à la sécularisation des sociétés, l'homme moderne a placé le sacré là où il ne devrait pas être, notamment dans la technique. Depuis, la technique a été sacralisée et l'être humain est entré dans une course à l'immodération. On finit par faire, simplement parce qu'on peut. La question de la nécessité ou non d'une technologie n'est jamais posée, seule reste la question de l'adaptation des sociétés aux innovations dont elles ne peuvent qu'accueillir l'existence.

Aujourd'hui encore, tous ces outils semblent nous dépasser et nous éprouvons une fascination toujours intacte pour ChatGPT, qui semble nous répondre comme un être humain – tout en oubliant la myriade de « *travailleurs du clic* » au Kenya, payés deux dollars de l'heure afin de modérer la machine.

La mise en œuvre à grande échelle de ces technologies semble s'apparenter à une vague irrépressible. Cette tendance nourrit l'imaginaire de la rupture avec le passé. Tout est toujours une révolution. Ce « *bluff technologique* », déjà mis en avant par Ellul, permet la justification de cette expansion sans limites et transforme l'individu en « *automate obéissant*⁶ ». Mais que reste-t-il, une fois la fascination dépassée ? Si ChatGPT peut être une aide intéressante dans le cadre d'un travail de création pour lequel la notion de vérité n'est pas pertinente, un tel outil ne peut en aucun cas servir de caution à la vérité.

⁵ Voir Jacques Ellul, *Le Bluff technologique* [1988], préface de Jean-Luc Pourquet, Paris, Fayard, coll. « Pluriel », 2012.

⁶ Voir Patrick Chastenet, *Introduction à Jacques Ellul*, Paris, La Découverte, 2019, coll. « Repères », p. 21-46.