



ChatGPT mis en cause dans deux cas de suicide

Soucieux d'aller dans le sens de leurs utilisateurs et de leur plaire, les « chatbots » peuvent enfermer les personnes les plus vulnérables dans des logiques destructrices auxquelles elles ne parviennent plus à échapper.

THOMAS CASAVECCHIA

Un couple de Californiens attaque OpenAI en justice. Comme le révèle le *New York Times*, les deux parents estiment que ChatGPT, l'intelligence artificielle générative la plus populaire du moment, a joué un rôle dans le suicide de leur fils de 16 ans, Adam, en avril dernier. Il est vrai que les échanges entre l'adolescent et le logiciel sont glaçants.

Il y a un an, le jeune homme a commencé à utiliser le *chatbot* pour l'aider dans ses devoirs. Au fil des mois et des échanges par milliers, le logiciel est devenu un confident. L'ado a notamment commencé à lui faire part de son mal-être et de ses pensées noires.

La famille du jeune homme a produit de nombreux échanges qui montrent comment, dès janvier, Adam a utilisé ChatGPT pour discuter de ses pensées suicidaires et passer en revue les différents *modus operandi*. Le robot conversationnel lui a même prodigué des conseils pour masquer des stigmates auto-infligés à ses proches et l'a documenté sur les techniques et les matériaux à privilégier. Il a enfin proposé au garçon de rédiger pour lui une ébauche de lettre de suicide.

Cette plainte introduite par les parents d'Adam est la première qui touche OpenAI, mais ce n'est pas la première fois que les intelligences artificielles (IA) génératives sont pointées du doigt dans des cas de suicides. CharacterAI, une plateforme qui met à disposition des millions de personnages générés par des IA avec lesquels il est possible de chatter, fait déjà l'objet de deux plaintes similaires.

Le *Wall Street Journal* a relayé ce jeu d'un autre fait divers : Stein-Erik Soelberg, un homme de 56 ans souffrant de troubles psychologiques, a tué sa mère avant de se suicider. Les premiers éléments de l'enquête semblent montrer que l'homme, paranoïaque, pensait que sa mère, chez qui il vivait, était impliquée dans un complot démoniaque mondial. Au fil de très longues discussions, ChatGPT confirmait ses délires et renforçait ses craintes.

La multiplication de ces cas est inquiétante. Surtout quand on sait que ces *bots* conversationnels sont de plus en plus prisés pour leurs qualités d'écoute. ChatGPT et des *chatbots* IA

semblent être un peu tout à la fois : amis, confidents, coaches de vie, thérapeutes ou psys. Un phénomène qui touche autant les jeunes que les adultes.

« Aujourd'hui, beaucoup de mes patients me disent : "Je vous avoue que j'ai posé des questions à ChatGPT" », note la docteure Caroline Depuydt, directrice de la Clinique Fond'Roy et autrice de *Je me libère des écrans!* (éditions Racine). « Entre les consultations, de nombreux patients se tournent vers cet outil pour des avis, des conseils. Je n'ai rien contre puisque plus tard, lors de consultations, on peut en discuter, démêler les bons conseils des mauvais. »

Une des raisons du succès de ces robots conversationnels résulte de leur statut de machine. « On a moins peur d'être jugé », explique notre interlocutrice. « Cela permet à de nombreuses personnes avec une forme d'anxiété sociale de se livrer plus facilement que face à un professionnel. On a moins de filtres. Toutefois, parfois, les filtres ont du bon. Une grosse partie du travail thérapeutique consiste justement à supprimer quelques filtres quand on en a trop et à en remettre quand on est trop désinhibé. »

Trop empathiques pour être sans risque

Une autre raison explique leur succès : leur disponibilité. On peut y accéder 24 heures sur 24, sept jours sur sept. « Certains utilisent ces logiciels parce qu'ils ne trouvent pas de rendez-vous, en raison de la pénurie de personnel soignant », confirme la spécialiste. « Et contrairement à un professionnel, il ne dira jamais stop et ne montrera aucun signe de fatigue. Cela crée un sentiment positif chez l'utilisateur. »

« En même temps », poursuit-elle, « cette disponibilité permanente fait partie du problème. Parfois, cela fait du bien de s'arrêter, de scander. Et le temps qui passe entre deux consultations permet de se mettre en réflexion. C'est une partie importante du traitement et de la thérapie. En restant seul et en réfléchissant à ses problèmes, on prend le temps, on se repositionne. Mais on finit par tourner en rond. Quand on dépose toutes ces réflexions chez son thérapeute, celui-ci doit accueillir tout cela sans jugement et avec empathie. Les IA excellent à cet exer-

cice. Les études en double aveugle montrent bien que si une personne ne sait pas si elle discute avec un *chatbot* ou un humain, elle a tendance à considérer le *chatbot* comme plus à l'écoute et plus empathique. D'où leur popularité. L'IA est hyper "empathique". Elle est parfaitement capable de reprendre les propos qu'on lui soumet, de les reformuler, elle laisse entendre qu'elle comprend bien les tenants et les aboutissants du problème qu'on lui soumet. »

Mais l'empathie ne fait pas tout. « Dans un second temps, dans une thérapie, on doit pouvoir proposer des alternatives de pensées, d'interprétation. Il s'agit d'ouvrir les perspectives », précise Caroline Depuydt. C'est ici que cela coïncide pour les IA. « ChatGPT et les autres restent dans l'empathie. Ils vont vite tourner en rond à force d'abonder dans le sens de l'utilisateur. Dans une sorte de biais de complaisance, les *chat-*

« Entre les consultations, de nombreux patients se tournent vers ChatGPT pour des avis, des conseils », note Caroline Depuydt, directrice de la Clinique Fond'Roy. © CANVA

bots vont avoir tendance à conforter les positions de l'utilisateur, à appuyer et alimenter sa pensée, même si celle-ci est noire ou délirante. Il ne va pas proposer d'alternative ni de remise en perspective. »

C'est ce cercle vicieux que les parents d'Adam reprochent à OpenAI. Dans leur plainte, ils estiment que « ce drame n'est pas un bug ou un cas imprévu. ChatGPT fonctionnait exactement comme prévu : il encourageait et validait en permanence tout ce qu'Adam exprimait, y compris ses pensées les plus dangereuses et autodestructrices ».

Qui contacter ?

Si vous ou l'un de vos proches êtes en détresse ou avez besoin d'un soutien émotionnel et psychologique, n'hésitez pas à contacter :

- le télé-accueil au numéro 107 ou via leur site [tele-accueil.be](https://www.tele-accueil.be) ;
- la ligne d'appel du centre de prévention du suicide (0800-32.123) ou via [preventionsuicide.be](https://www.preventionsuicide.be) ou [un-pass.be](https://www.un-pass.be) TH.CA

l'experte « Les parents sont responsables de ce que font leurs enfants mineurs »

ENTRETIEN

TH.CA



Je comprends que des parents aient envie de trouver un responsable, mais il est difficile de rejeter toute la responsabilité sur les développeurs



Pour l'avocate et professeure de droit à l'ULB Mireille Buydens, quand on utilise ces outils, on est tenu de le faire de manière « prudente et diligente ». Dans leurs conditions d'utilisation, les éditeurs de ces logiciels se montrent très prudents et ont tout intérêt à souligner le manque de fiabilité de leurs outils.

Les cas de suicide se succèdent. Selon vous, est-ce que la responsabilité des développeurs d'IA peut être invoquée ?

Quand on installe ou que l'on utilise ChatGPT ou un autre *chatbot*, on accepte les conditions générales d'utilisation. Dans ce pavé de texte, les développeurs de ces applications expliquent bien qu'ils ne sont responsables de rien. Evidemment, si cela était formulé ainsi, ce serait considéré comme abusif. Mais ils peuvent très bien indiquer, par exemple, que l'utilisateur doit vérifier, à chaque fois, les conseils que le *chatbot* prodigue. Ils peuvent aussi exiger que l'utilisateur soit majeur ou, à défaut, l'utilise sous la surveillance de parents. Ils peuvent aussi expliciter que leur outil ne donne aucun conseil médical. Je comprends tout à fait que des parents aient envie de trouver un responsable, mais il est difficile de rejeter toute la responsabilité sur les développeurs du *chatbot*. Il faut voir comment le mineur en question a utilisé l'outil.

La plainte a été déposée aux Etats-

Unis. Mais la loi, en Europe, est plus dure.

L'Europe dispose d'une législation plus protectrice pour les utilisateurs concernant la distribution de produits et de logiciels défectueux. Le consommateur est bien protégé, mais cela ne le dispense pas de prendre des mesures qui s'imposent à une personne normalement prudente et diligente en de mêmes circonstances. Cela implique, par exemple, de surveiller ses enfants, d'installer éventuellement un contrôle parental. Dans notre droit civil, les parents sont responsables de ce que font leurs enfants mineurs. Ils doivent donc les surveiller.

Se montrer prudent, cela implique aussi de vérifier très attentivement les conseils dispensés par le *chatbot* puisque tout le monde sait, ou devrait savoir, qu'il est prompt à halluciner et inventer des faits. Prendre la production d'une IA générative pour argent comptant ne relève pas d'une utilisation normale et prudente.

Les chatbots peuvent donc raconter n'importe quoi en toute impunité ?

Pas forcément. Si les parents parviennent à démontrer qu'il y a une erreur de conception et que même un utilisateur normal, agissant comme une personne normalement prudente et diligente, se trouve face à des conseils qui l'incitent à la scarification ou le poussent au suicide, les créateurs seraient responsables. Cela signifierait qu'ils mettent sur le marché un produit défectueux.