

<https://larcenciel.be/spip.php?article1458>



# Le chatbot Eliza a brisé une vie : il est temps d'agir

- MATIÈRE À PENSER - ENJEUX DU NUMÉRIQUE - ENJEUX DE L'INTELLIGENCE ARTIFICIELLE -

Date de mise en ligne : jeudi 30 mars 2023

---

Copyright © LARCENCIEL - site de Michel Simonis - Tous droits réservés

---

**Le suicide, en Belgique, d'un homme suite aux incitations d'un chatbot souligne le risque de la manipulation émotionnelle.**

**Une carte blanche de 50 académiques, juristes, ingénieurs, philosophes, médecins, philosophes, sociologues, criminologues et éthiciens tire le signal d'alarme.**

**Développeurs et fournisseurs d'Intelligence Artificielle (IA) mais aussi les gouvernements doivent adopter au plus vite un cadre juridique protecteur.**

**Extraits de la Carte blanche (publiée le 29-03-2023 [sur le site de La Libre avec la liste des 50 signataires, ici.](#))**

Compte tenu des implications éthiques, juridiques et sociales de l'IA, la question de sa désirabilité – en particulier de sa forme, de sa fonction, de ses capacités, de ses exigences techniques et de ses garanties – est de plus en plus pressante. Il a déjà été démontré, par exemple, que les chatbots, comme d'autres systèmes basés sur l'IA, souffrent de biais capables de générer ou d'accentuer certaines formes de discriminations. Ils sont également susceptibles d'"halluciner", d'affirmer des "vérités" sans fondements factuels, de tenir des propos haineux ou encore de propager des informations erronées. Leur fonctionnement opaque et leur évolution imprévisible ajoutent à ces dangers.

Le suicide récent, en Belgique, d'un homme suite aux incitations d'un chatbot cristallise des craintes restées jusqu'ici théoriques, et met en lumière un risque bien présent : celui de la manipulation. Ce drame illustre l'une des conséquences les plus extrêmes de ce risque, mais cette manipulation émotionnelle peut également se manifester sous des formes plus subtiles.

Dès lors que les personnes ont le sentiment d'interagir avec une entité subjective, elles entretiennent avec ce "partenaire", parfois inconsciemment, un lien qui les expose à ce risque et peut éroder leur autonomie. Il ne s'agit donc pas d'un incident isolé. D'autres utilisateurs de chatbot ont en effet décrit ses effets manipulateurs.

### **Manque de compréhension, néanmoins trompeuse**

Suite à l'enquête d'un journaliste américain le mois dernier concernant le Bing AI Bot de Microsoft, dont l'utilisation a généré des textes similaires à ceux de la victime belge (à savoir, des déclarations d'amour et des exhortations à quitter sa femme), Microsoft a pris des mesures fortes destinées à limiter le nombre de chats que l'on peut échanger à la fois. Mais des bots similaires existent sur de nombreux sites web sans aucune restriction.

De plus, de nombreux sites Web ont comme but précis la mise à disposition de chatbots dotés d'une "personnalité", ce qui augmente encore le risque de manipulation émotionnelle.

Bien que la plupart des utilisateurs soient conscients que le bot avec lequel ils discutent n'est pas une personne douée de conscience, mais un algorithme capable de prédire la combinaison de mots la plus plausible sur la base d'une analyse poussée de données, il est toutefois dans notre nature de réagir de manière émotionnelle lorsque nous sommes confrontés à des situations impliquant des entités dont les caractéristiques évoquent un comportement humain. Exiger des entreprises qu'elles indiquent clairement la nature "artificielle" du chatbot n'est donc pas suffisant.

### **Tout le monde est vulnérable**

Nous inquiète tout particulièrement, le fait que des enfants puissent interagir sans effort avec des chatbots qui, d'abord, gagnent leur confiance, pour, ensuite, préférer des propos haineux, conspirationnistes ou pornographiques, voire même les inciter à se suicider.

Les personnes isolées qui ne disposent pas d'un réseau social solide, ou les plus déprimées, sont également à risque.

En effet, la réponse émotionnelle que ces derniers déclenchent se produit sans que nous en ayons réellement conscience.

Si, dans l'incident belge, c'était un humain qui avait interagi avec la personne, nous qualifierions certainement cela d'incitation au suicide ou d'omission coupable – deux infractions pénales passibles d'une peine d'emprisonnement.

### **Le progrès à tout prix – pour le reste, on verra plus tard**

Comment se fait-il alors que ces systèmes d'IA soient librement accessibles ? Le recours à la réglementation est souvent critiqué dans ce contexte, car celle-ci "ne devrait pas entraver l'innovation".

Si une société pharmaceutique désire commercialiser un nouveau médicament contre une maladie X, elle ne peut pas simplement ignorer ses effets secondaires, aussi révolutionnaire que soit le traitement. De même, le concepteur d'une nouvelle voiture devra soumettre son produit à des tests approfondis pour détecter toutes sortes d'incidents et prouver qu'il est sûr avant que la voiture ne soit lancée sur le marché. Est-il déraisonnable d'en attendre autant des développeurs de systèmes d'IA ?

Aussi divertissants que puissent être les chatbots, ils sont plus qu'un simple jouet.

### **Et maintenant ?**

Au vu de tout ce qui précède, nous appelons à une campagne de sensibilisation du public aux risques associés aux systèmes d'IA et encourageons les développeurs à prendre leurs responsabilités.

Un changement d'état d'esprit est nécessaire, afin que les gens identifient, testent et traitent d'abord les risques de ces systèmes avant de les rendre accessibles au plus grand nombre. À cet égard, l'éducation joue également un rôle important, à tous les niveaux.

Soyons clairs : nous sommes tous fascinés par les capacités offertes par ces systèmes. Il est cependant crucial de s'assurer que leur développement et leur utilisation s'inscrivent dans le respect des Droits de l'Homme. La responsabilité incombe non seulement aux développeurs et aux fournisseurs d'IA, mais aussi à nos gouvernements aux niveaux national, européen et international. Ils se doivent d'adopter, au plus vite, un cadre juridique protecteur assorti de garanties solides et de vérifications préalables. Cela nécessite également des organes (consultatifs) politiques qui anticipent les risques en suivant une approche multidisciplinaire incluant toutes les parties concernées.

Que cela soit un signal d'alarme pour tous.

La récré est terminée : il est temps de tirer des leçons et d'assumer ses responsabilités.

La liste des co-signataires se trouve sur le site de La Libre : "*Il est temps d'agir face à l'IA manipulatrice*", à consulter si on veut lire le texte de la Carte blanche en entier.

*Post-scriptum :*

*Cette carte blanche fait suite à deux articles dont j'ai repris des extraits ICI.*