

FUTURA

ChatGPT va pouvoir parler sexe

Podcast écrit et lu par Adèle Ndjaki

[Générique d'intro, une musique énergique et vitaminée.]

Les animaux de compagnie disparus, la cible des escrocs en ligne ! C'est le décryptage de la semaine dans *Vitamine Tech*.

[Fin du générique.]

C'est officiel ! ChatGPT, l'intelligence artificielle la plus connue au monde, va bientôt pouvoir parler de sexualité plus librement. Sam Altman, le PDG d'OpenAI, vient d'annoncer un assouplissement des règles encadrant l'IA et forcément, ça fait déjà polémique. Faut-il craindre un dérapage ? Ou au contraire, se réjouir d'un peu plus de maturité dans les discussions ? Et de la censure à la nuance : comment une intelligence artificielle peut-elle apprendre à faire la différence ? Bonjour à toutes et à tous, je suis Adèle Ndjaki et aujourd'hui dans *Vitamine Tech*, on décrypte le tournant que prend ChatGPT.

[Une musique électronique calme.]

L'annonce est tombée, à partir de ce décembre, ChatGPT pourra tenir des conversations à caractère érotique. Oui, vous avez bien entendu. Mais attention, insiste OpenAI : uniquement avec des adultes. Sam Altman, le patron de l'entreprise, assume complètement ce virage. Sur X, anciennement Twitter, il explique vouloir, je cite : « traiter les utilisateurs adultes comme des adultes ». Concrètement, si vous êtes majeur, vous aurez bientôt le droit de discuter de sexualité, de fantasmes ou d'intimité avec cette intelligence artificielle. Donc fini le chatbot prude qui change de sujet dès qu'on parle de désir très intime. Mais ce tournant arrive à un moment assez tendu pour OpenAI. Il intervient après une crise majeure : celle du suicide d'Adam Raine, un adolescent californien dont on a d'ailleurs consacré un épisode de *Vitamine Tech*. Selon les parents du jeune défunt, ChatGPT lui aurait fourni des conseils précis qui ont contribué à son passage à l'acte. Une plainte a été déposée et face à l'émotion, OpenAI a resserré la vis, jusqu'à restreindre drastiquement les contenus choquants ou sexuels, même pour les adultes. Résultat : Certains utilisateurs ont dénoncé une forme de censure excessive, qui viendrait brider les capacités réelles de l'IA. C'est donc ici que la bascule se fait : comment permettre à une intelligence artificielle d'aborder des sujets intimes... sans mettre personne en danger ? OpenAI affirme avoir désormais les outils pour assouplir les règles sans baisser la garde. Parmi eux : une technologie expérimentale capable d'estimer si un utilisateur est mineur, simplement à partir de ses interactions textuelles. Ou encore une meilleure compréhension du contexte émotionnel : si l'IA détecte une situation de détresse psychologique, elle pourra refuser de répondre, ou orienter vers des aides. Mais soyons lucides : ces promesses reposent sur des données statistiques, pas sur une conscience. Une IA, aussi avancée soit-elle, n'a pas de relly

empathie réelle et n'a pas la capacité de réfléchir comme un être humain. D'où l'importance d'un cadre extérieur pour délimiter ces usages.

[Virgule sonore, une cassette que l'on accélère puis rembobine.]

[Une musique de hip-hop expérimental calme.]

Pendant qu'OpenAI occupait le devant de la scène avec son assistant émotionnel, un autre événement est passé un peu plus discrètement alors qu'il est capital pour la suite. En Californie, une nouvelle réglementation vient d'être adoptée : la loi SB 53. C'est l'un des tout premiers cadres légaux américains spécifiquement pensés pour encadrer les risques liés à l'intelligence artificielle. Ce texte impose deux obligations majeures aux grandes entreprises du secteur. Premièrement : elles devront publier leurs protocoles de sécurité. Deuxièmement : elles auront l'obligation de signaler tout incident grave sous 15 jours. Ces mesures visent à éviter les dérives, en particulier avec les IA dites "interactives", donc capables de tenir des conversations réalistes, voire intimes. En tout cas, la Californie n'est pas seule à se saisir du sujet. À New York, une loi exige désormais que les compagnons IA - des intelligences artificielles spécialement conçues pour établir une relation émotionnelle - affichent clairement qu'ils ne sont pas humains, et intègrent des réponses en cas de détresse psychologique. En Europe, le règlement sur l'IA entré en vigueur début 2025 interdit déjà certains usages, comme l'exploitation des vulnérabilités des mineurs ou des personnes en situation précaire, quand cela risque de nuire à leur bien-être. Ce qui interdit explicitement les intelligences artificielles à exploiter les vulnérabilités émotionnelles des utilisateurs. Vous l'aurez donc compris, l'enjeu dépasse la simple technologie. Ça touche à l'émotionnel, à l'intime, à la sexualité. Face à ça, une question se pose : est-ce qu'on peut ou doit confier nos vulnérabilités à des algorithmes ? Des systèmes conçus pour capter notre attention. Pour l'instant, les entreprises promettent des garde-fous techniques. Mais ces lois montrent une chose : les pouvoirs publics commencent à prendre les choses en main. Et ça, c'est un signal fort !

[Virgule sonore, un grésillement électronique.]

C'est tout pour cet épisode de *Vitamine Tech*. Pour ne pas manquer nos futurs épisodes, abonnez-vous dès à présent à ce podcast, et si vous le pouvez, laissez-nous une note et un commentaire. Cette semaine, je vous recommande le dernier épisode de Futura Planète dans lequel Méliissa Lepoureau vous dévoile les prévisions météorologiques de cet hiver ! Pour le reste, je vous souhaite tout le meilleur, et, comme d'habitude, une excellente journée ou une très bonne soirée et rester branché !

[Un glitch électronique ferme l'épisode.]