



Grok : l'IA qui vous dénude sur le Net !

Podcast écrit et lu par Adèle Ndjaki

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

Grok déshabille des internautes, c'est un véritable problème et c'est le décryptage de la semaine dans *Vitamine Tech*.

[Fin du générique.]

2026 vient de débuter et pourtant une affaire secoue déjà le web. Grok, l'intelligence artificielle intégrée au réseau social X est au cœur d'une controverse majeure : des milliers de femmes et même des mineurs se sont retrouvés dénudés, sur Internet, sans leur consentement. Alors comment une simple photo peut-elle se transformer en image intime en ligne et surtout comment une IA, censée aider ses utilisateurs, a-t-elle pu être utilisée pour créer des contenus illégaux? Bonjour à toutes et à tous, je suis Adèle Ndjaki et cette semaine dans *Vitamine Tech*, on vous dit tout sur ce qu'il se passe avec Grok.

[Une musique électronique calme.]

On en a déjà parlé, et pourtant le phénomène continue de s'amplifier : les femmes sont aujourd'hui les premières victimes des deepfakes à caractère pornographique. Mais de quoi parle t-on exactement ? Les deepfake ou les hypertrucages en français, sont des contenus manipulés à l'aide de l'intelligence artificielle. Généralement ça renvoie surtout à des images et à des vidéos truquées très sophistiquées. Mais si l'on veut parler des deepfakes à caractère sexuelle, on peut dire que ce phénomène s'inscrit en réalité dans une histoire assez longue, celle des violences numériques et sexistes. Bien avant l'essor de l'IA générative, les femmes étaient déjà massivement ciblées en ligne. Dès 2014, une enquête de l'Agence des droits fondamentaux de l'Union européenne révélait, que dans l'UE, une femme sur dix, avait été victime de cyberharcèlement depuis l'âge de 15. Cependant, il faut souligner qu'à l'époque, il s'agissait surtout de messages insultants, de menaces, de rumeurs diffusées sur les réseaux sociaux, ou encore de photos intimes volées et partagées sans le consentement des principales intéressées. Conclusion : l'intelligence artificielle n'a pas créé les cyberviolences faites aux femmes, cependant on peut constater qu'elle les ont transformées et amplifiées. On l'a vu dans *Vitamine Tech*, certaines IA génératives sont capables d'aller très loin. Certaines peuvent générer des images à caractère sexuel à partir d'une simple photo trouvée en ligne. On parle alors de clichés qui mettent en scène des visages bien réels, de personnes célèbres ou anonymes, intégrés à des corps qui ne sont pas les leurs. Et ces images qui se retrouvent parfois sur des sites pour adultes sans le consentement des victimes. Le problème avec ce sujet c'est les chiffres ne cessent d'évoluer mais pas dans le bon sens : 82 % des victimes de cyberviolences sexistes et sexuelles sont des femmes ou des filles selon une enquête menée par les associations

Point de Contact et #StopFish. Et parmi ces violences, les deepfakes prennent une place de plus en plus importante, à mesure que les outils se démocratisent. En 2019, une étude menée par l'entreprise néerlandaise Sensity AI montre déjà que 96 % des deepfakes en circulation sur Internet sont à caractère pornographique et que dans 90 % des cas, ils représentent des femmes. On parle donc là de contenus qui peuvent très vite devenir des outils de pression, largement utilisés pour humilier, diffamer ou discréditer une personne. Les conséquences sont donc bien réelles et souvent durables. Face à cette explosion, les plateformes numériques et les législateurs tentent de réagir. Mais bien souvent, les réponses arrivent trop tard, ou restent insuffisantes face à la vitesse des innovations technologiques. De nouveaux outils continuent d'apparaître, toujours plus performants et toujours plus accessibles. Et c'est précisément là que l'actualité nous rattrape.

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

[*Une musique de hip-hop expérimental calme.*]

Depuis quelques jours, Grok, l'intelligence artificielle intégrée au réseau social X, est au cœur d'une controverse. Ce nouvel an, des milliers d'utilisatrices se sont vues dénudées, sur le réseau social, sans leur consentement. À la demande d'internautes mal intentionnés Grok a modifié des photos parfaitement normales de femmes et mineurs, les faisant apparaître en bikini, en lingerie ou partiellement dénudée. La situation est donc grave, l'affaire dépasse largement le simple débat sur la modération des contenus, des enquêtes judiciaires ont par conséquent été ouvertes, notamment en France. Mais essayons d'abord de comprendre comment on a pu en arriver là. Grok, c'est une intelligence artificielle dite multimodale, capable de lire du texte, mais aussi d'analyser et de générer des images. Quand un utilisateur lui fournit une photo, l'IA identifie des éléments visuels (un visage, une silhouette, des vêtements) en s'appuyant sur des modèles statistiques. Donc dans cette histoire, Grok ne retire pas réellement les vêtements présents sur la photo, mais propose une version alternative de la scène, qu'elle estime plausible d'après ses apprentissages. Cependant se retrouver dénudé du jour au lendemain sur le Net ça pose quand même un véritable problème, même s'il ne s'agit pas véritablement de son corps. Et ce qui peut paraître incompréhensible, c'est qu'en théorie, ce type de génération est encadré par des règles internes, conçues pour empêcher la nudité explicite ou les contenus illégaux. Mais vous pouvez le constater, dans les faits ces garde-fous se sont révélés insuffisants. Des requêtes très simples, comme "mets-la en bikini" ou "change sa tenue", suffisent à contourner les interdictions et passent sans être bloquées. Certains utilisateurs ont donc exploité des zones grises de l'IA. En tout cas, cette actu marque un avant et un après. Même si les deepfakes existent depuis plusieurs années, leur création demandait jusqu'ici des compétences techniques, des logiciels spécifiques, du temps et ces pratiques restaient relativement marginales. Aujourd'hui avec Grok, tout devient instantané et accessible. L'outil est intégré directement à X et est proposé à l'ensemble des utilisateurs. Donc, la barrière technique a désormais quasiment disparu. Mais cette affaire ne se résume pas à un simple problème technique : elle met en lumière des enjeux culturels et économiques plus profonds autour de l'IA. Dans le cas de Grok, l'intelligence artificielle développée par xAI, certains médias relèvent que l'outil a été conçu avec des garde-fous moins stricts que ceux de ses concurrents, ce qui a parfois permis la génération de contenus problématiques ou explicites qu'un système plus restrictif aurait bloqués. Et puis il faut aussi parler de la réaction d'Elon Musk. La situation est grave mais le ton est clairement moqueur. Face à la polémique autour d'images sexuelles générées par l'IA Grok, milliardaire choisit de tourner l'affaire en dérision

sur X, en minimisant les critiques et les inquiétudes. Pourtant, juridiquement, le sujet est très sérieux. En France, diffuser des images sexuelles sans consentement est un délit, encore plus grave lorsqu'il s'agit de mineurs. Le parquet de Paris a d'ailleurs ouvert une enquête, confiée à la cybercriminalité pour vérifier si des infractions ont été commises. Et ce n'est pas qu'un débat français. Dans d'autres régions du monde, notamment en Asie du Sud-Est, les autorités surveillent de près l'utilisation de cette IA sur le réseau X. Au fond, cette affaire relance une question plus large : comment encadrer l'intelligence artificielle ? L'Union européenne rappelle les règles avec le Digital Services Act, et l'ONU comme l'UNESCO alertent depuis longtemps sur la nécessité de protéger les droits humains, en particulier ceux des mineurs. En résumé, cette affaire montre surtout qu'il y a un décalage de plus en plus marqué entre la rapidité avec laquelle les technologies évoluent et le temps, beaucoup plus lent, du droit et de la régulation. D'où l'importance d'être particulièrement vigilant sur ce que vous publiez sur les réseaux sociaux.

[*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 tout nouvel épisode de Melissa Lepoureau vous parle du lapin d'Amami, un animal considéré comme étant un « fossile vivant » et qui est menacé d'extinction! 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.*]