

Le « slop », ce **raz-de-marée** de contenu par l'intelligence artificielle qui menace

Les IA génératives ont facilité la production automatisée et sans vérification d'images, de vidéos ou de sites web cherchant à capter le maximum d'audience sur Facebook, YouTube ou Google. Un appât du gain qui trouble notre rapport entre le vrai et le faux.

LE FIGARO

ENQUÊTE

CHLOÉ WOITIER

Artiste dans le jeu vidéo, Rachel Dufossé conçoit des décors pour les productions du studio rennais Wild Wits Games. Ce travail l'amène à faire des recherches pour trouver des références visuelles. « Pour des questions de gain de temps, Internet est une ressource naturelle dans notre métier, en plus de notre bagage artistique et culturel », nous explique-t-elle.

Mais depuis quelque temps, son quotidien se complique. « J'ai récemment dû créer un décor de monastère », relate-t-elle. « En faisant mes recherches, je suis tombée sur des images qui ne m'ont pas choquée à première vue. Mais en y regardant de plus près, plein de choses clochaient... Les informations de ces visuels, qui étaient en réalité générés par l'intelligence artificielle, étaient inexactes. » De tels incidents deviennent de plus en plus fréquents dans sa vie professionnelle. « Les référentiels d'images se dégradent de jour en jour sur Google et

sur Pinterest », alerte-t-elle. « En tant qu'artiste, j'ai peur de ne plus avoir accès à des ressources fiables sur Internet. Et, plus largement, que nos imaginaires collectifs se dégradent. »

« Pinterest ou Behance sont aujourd'hui pollués par des amas de visuels générés par l'intelligence artificielle » (IA), abonde Chloé Chanot, directrice de la création chez l'agence publicitaire 87 Seconds. « Si vous cherchez, par exemple, des visages de femmes âgées, 70 % des contenus mis en avant n'auront rien de naturel. Ces plateformes,

qui étaient nos moteurs de recherche, ne sont plus des références. Nos usages vont devoir changer face à cette perte de qualité, cette perte de temps, mais aussi cette perte de confiance liée à ce nouveau type de spam. »

« Des contenus jetables et vides de substance »

Les métiers de l'image ne sont pas les seuls touchés par cet afflux soudain de contenus synthétiques de mauvaise qualité. « Un de mes étudiants m'a récemment conseillé une chaîne YouTube qui parle de ma matière. J'ai regardé : les vidéos, au script clairement généré par l'IA, racontaient n'importe quoi », témoigne Marie Molins, artiste et chercheuse en design des interfaces utilisateurs (UX/UI). Son travail est aussi affecté par l'émergence sur le web de textes douteux : « Je dois désormais vérifier par deux fois si les références que j'utilise pour mes cours sont exactes, et si leurs auteurs existent réellement. Cela touche même la littérature scientifique. »

Images, vidéos, sites web... Produits

en masse à partir d'outils tels que ChatGPT, Grok, Bing Image Creator ou Invideo AI, ces contenus synthétiques de mauvaise qualité s'infiltrent à bas bruit partout sur Internet, des résultats de recherche Google aux vidéos YouTube, en passant par les groupes Facebook ou les livres autoédités sur Amazon. Un terme est en train de s'imposer pour les désigner : le « AI Slop », *slop* étant dans la langue anglaise un terme désignant la bouillie industrielle donnée aux cochons.

« Tous les contenus générés par l'IA ne sont pas du slop », précise Jean-Noël Buisson, responsable veille et analyse d'Image 7. « Sont du slop les contenus IA de basse qualité, qui sont le plus souvent erronés, et avec derrière une intention de profit. »

Aucune vérification avant publication

« Le slop, ce sont des contenus jetables, vite faits, et vite oubliés, vides de substance, et dont l'objectif est de nourrir les algorithmes des plateformes », ajoute Chloé Chanot. Esteban Giner, enseignant-chercheur en sciences de la communication à l'école Gobelins Paris, abonde : « Ces contenus ne sont pas simplement chaotiques. Ils prennent en compte les logiques des algorithmes, dans une recherche de viralité et d'engagement. »

Leurs effets sur le monde réel commencent à devenir tangibles. Le 31 octobre dernier, des milliers de personnes ont afflué dans les rues de Dublin pour assister à une grande parade d'Halloween. Problème : un tel évènement n'a jamais été prévu. Il avait pourtant été annoncé par Myspirthalloween.com,

un site alors bien référencé par Google. Ce dernier a été conçu à partir d'IA générative par une entreprise pakistanaise cherchant à capter le maximum de trafic, et donc de revenus publicitaires, sans prendre la peine de vérifier ce qu'elle publie... Même incident, et autant de personnes déçues, à Birmingham le 31 décembre dernier, après qu'une IA a inventé un feu d'artifice pour le Nouvel An.

« Ces acteurs peu scrupuleux sont en recherche d'audience à tout prix en exploitant les mots-clés tendance sur Google. Le phénomène existait bien avant l'IA, mais il est aujourd'hui facilité », explique Valentin Simony, conseiller en stratégie SEO. On retrouve la même logique sur YouTube ou bien sur TikTok, avec un déluge de vidéos, dont le script puis les images auront été générés par l'IA, qui exploitent les tendances de recherche du moment. Comme elles répondent à une demande du public, sont bien titrées et illustrées, et sont « likées » et commentées par des faux profils, ces vidéos sont mises en avant par les algorithmes. Mais, là encore, leurs auteurs, appâtés par les revenus publicitaires, n'auront pas vérifié que ce qu'elles racontent est correct... « La capacité de générer en masse de tels contenus est un désastre pour notre système informationnel », prévient Ioana Manolescu, chercheuse en informatique à l'Institut national de recherche en sciences et technologies du numérique et professeure à l'école Polytechnique.

La start-up américaine NewsGuard a recensé à ce jour 1.550 sites de soi-disant information utilisant l'IA générative. « Ils se concentrent souvent sur

des thématiques à clic, comme la crypto ou les célébrités », explique Chine Labbé, rédactrice en chef Europe chez NewsGuard. « Certains traduisent dans d'autres langues des articles de presse, parfois sous paywall, ou bien les résumant automatiquement, quitte à faire des erreurs. Et ce sont autant de revenus publicitaires qui échappent aux médias. » Sur ce panel, 10 % diffusent de la désinformation. « Ces pratiques n'ont rien de nouveau, mais avec l'IA, elles sont sous stéroïdes », poursuit-elle. « Une seule personne peut contrôler tout un réseau de sites. On change d'échelle, donc on change la nature du problème. »

Certaines pages Facebook illustrent ce phénomène jusqu'à l'absurde. Elles publient une masse d'images générées par l'IA cherchant à susciter une émotion, qui se transcrit par un like, un commentaire ou un partage. Un petit garçon moqué par ses camarades ; un jeune vétéran américain amputé qui fête son anniversaire ; des photos d'un couple à cinquante ans d'écart ; un mendiant qui fabrique une moto avec des cannettes de soda...



La capacité de générer en masse de tels contenus est un désastre pour notre système informationnel

Ioana Manolescu

Chercheuse en informatique

”

Les « pièges à baby boomers »

Ces pages sont qualifiées de « pièges à baby boomers » (« boomers trap »), car elles cherchent à exploiter la crédulité d'internautes âgés, dont l'œil n'est pas exercé à repérer les images IA, en appuyant sur des concepts qui peuvent leur parler. « Les images de personnes pauvres, c'est génial », clame une formation payante en ligne indienne qu'a pu suivre le site 404 Media. « Tout ce qui touche aux sentiments, les bébés, les animaux, ça suscite beaucoup d'engagement. Vous pouvez mettre en ligne cinquante à cent images par jour et vous faire pas mal d'argent. »

Ces internautes, vivants en Inde ou en Asie, cherchent à profiter du programme Performance Bonus de Facebook, qui rémunère les créateurs à hauteur de dizaines de dollars par image ayant « percé » – un chiffre qui peut s'envoler en cas de forte viralité. Cette promesse d'argent facile a engendré sur le réseau social une avalanche de slop depuis plusieurs mois. Se retrouvent ainsi propulsés sur les fils d'actualité ces visuels factices et attrape-clics. Certains sont clairement étranges, comme un Jésus ayant fusionné avec des crevettes, une image devenue symbole des dérives du slop...

Cette mécanique ne semble pas combattue par les plateformes. « Les internautes publient de moins en moins sur les réseaux sociaux, qu'ils consomment désormais passivement. Le slop aide à nourrir la bête », décrypte Jean-Noël Buisson. Le groupe Meta voit même plus loin. « Nous pensons que des IA vont avoir leurs comptes sur nos réseaux sociaux, avec leurs photos, et seront capables de générer et de partager leurs propres contenus », a clamé en janvier, dans le *Financial Times*, Connor Hayes, vice-président chargé des produits IA.

« Par défaut, tout devient faux »

Nos fils Instagram seront-ils bientôt remplis d'images et de vidéos créées par des robots ? L'ensemble des experts que nous avons interrogés pointent un danger à l'horizon si rien n'est fait pour bloquer cette submersion. « Nous avons été éduqués à avoir une grande confiance envers les images », prévient Marc de Vaugiraud, *data scientist* chez Datacraft. « Or, il sera bientôt impossible de faire la distinction entre une



Près de 130.000 réactions pour cette photo IA de couple « avant-après ». © DR.



Ce type d'images IA, où une fausse personne montre quelque chose soi-disant réalisé de ses mains et réclame des encouragements, est devenu un classique du slop. © DR.

