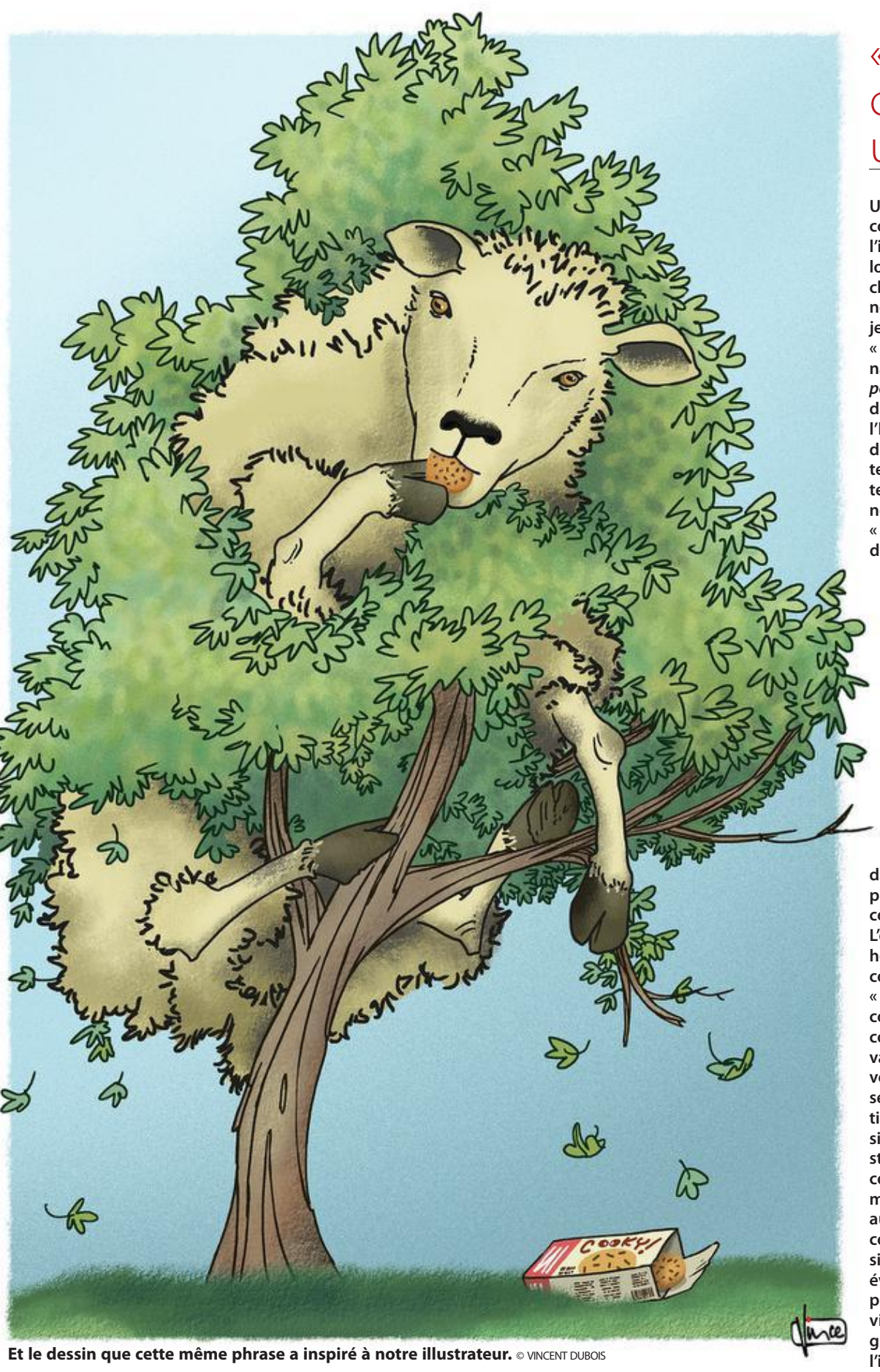
e « fake »



Et le dessin que cette même phrase a inspiré à notre illustrateur. © VINCENT DUBOIS

« J'ai pensé comme un ordi »

© DR

Un mouton mange un cookie dans un arbre : l'instruction donnée au logiciel Midjourney était claire. Vincent Dubois, notre illustrateur, a joué le jeu, le doigt sur la couture. « Un exercice passionnant », commente-t-il *a posteriori*. « Le paradoxe dans cet exercice entre l'humain et l'ordinateur, c'est de voir deux modèles de composition qui tentent de converger. Les deux vont tenter de créer une image. De l'imaginer, donc de la "penser". »

« Le dessinateur a été amené à rester dans les clous d'une phrase simple et imposée », poursuit-il, « sans la contextualiser comme on le fait habituellement pour illustrer un article : arbre, mouton, cookie, point-barre ! Il ne s'est pas posé la question de savoir pourquoi le mouton est dans l'arbre. Pourquoi un cookie ? Deux questions qui amènent à réfléchir. A penser. Le dessinateur a dû "freiner" sa sémantique habituelle : pas de second niveau de langage, qui amènerait, par exemple, à dessiner le mouton non pas comme un animal, mais comme un individu, qui se noie dans le groupe social. L'ordinateur va puiser dans son stock d'images, comme l'humain dans sa connaissance. L'ordi va tenter de "penser comme un homme", le dessinateur va penser comme l'ordinateur. »

« Nous sommes donc sur un pont », conclut-il. « A gauche, le dessinateur va composer une image basique, et l'ordi va tenter une image "humaine". Ils vont avancer sur ce pont et tenter de se rejoindre. La différence de composition se fera au niveau du mouton dessiné : il a une bonne bouille, plutôt style "cartoon". Comme il mange un cookie, il a une pose un peu anthropomorphique. Il y a un paquet de biscuits au sol, car il faut bien contextualiser le cookie. Il y a un ciel simple, car le dessinateur a pensé comme l'ordi et a évité le fond blanc habituel, qui "remplit" tout, (car l'ordinateur a horreur du vide). Comme des Legos de notre imagination, nous avons construit l'image. »

hackers russes montrant Volodymyr Zelensky prêt à déposer les armes. L'impact sur les opinions publiques est ravageur.

La menace va grandissante. Grâce aux techniques de création d'images, les « chatbots » (typiquement, celui du service après-vente d'une chaîne de magasins avec lequel l'internaute discute via des messages textuels) sont parfaitement capables de feindre une allure humaine. Et, comme c'est déjà le cas, d'interagir personnellement avec les internautes en détectant leurs émotions et leurs intentions. Demain, ils pourront aussi exploiter ses biais cognitifs et les manipuler de façon très efficace, que ce soit pour acheter un produit ou pour voter pour un parti politique. La vérification des « intentions des machines » est donc devenue un enjeu aussi crucial que la vérification des infos relayées sur les réseaux sociaux.

Le potentiel d'influence de l'intelligence artificielle n'a d'ailleurs pas échappé à la propagande chinoise, laquelle multiplie les vidéos de faux « influenceurs ». Non loin de là, les Japonais n'en croyaient pas leurs yeux le jour où ils ont découvert que derrière la belle et jeune influenceuse @azusagakuyuki se cachait un informaticien de 50 ans. La supercherie n'a pas refroidi les fans. Au contraire, son compte est passé de 19.000 à 25.000 followers.

Digne de confiance

Sur les applications de rencontre, l'imposture est devenue monnaie courante. Un témoin, actif dans le secteur médical, nous raconte aussi ce moment où il a réalisé que les personnes qui, la veille, avaient tenté de lui vendre une solution informatique en visioconférence n'étaient en réalité que des avatars hyperréalistes. « Demain, il faudra vérifier l'identité d'un interlocuteur sur Teams, par exemple par des moyens biométriques », prévient-il.

La vraie vie devient un conte de fake, préfigurant notre vie d'avatar dans le métavers. Sera-t-elle plus belle ? Des chercheurs de l'université de Lancaster ont pu démontrer que les visages générés par l'intelligence artificielle sont perçus, en moyenne, comme 7,7 % plus dignes de confiance que ceux humains. Un phénomène perturbant qui s'expliquerait par le fait que les visages de synthèse sont assez lambda. Cette apparente normalité inspirerait un sentiment de confiance chez la plupart des individus.

C'est précisément ce facteur « confiance » (« je crois ce que je vois ») qui est sournoisement exploité par des entreprises ou des Etats, mais aussi, fatalement, par des pirates informatiques. Selon le spécialiste en cybersécurité VMWare, le nombre d'attaques opérées via des « deep fake » a augmenté de 13 % en 2021. L'usurpation d'identité est devenue leur nouveau cheval de Troie. Le 28 juin dernier, le FBI publiait une note alertant sur un nombre grandissant de plaintes pour des usages de *deep fakes* lors d'entretiens d'embauche effectués par visioconférence. Une fois engagés, ces faux profils n'ont guère de mal à accéder à des données sensibles. Avant le FBI, c'est Europol, en avril, qui affirmait que le « deep fake » était devenu un outil du crime organisé,

Les « deep fakes » répondent aux mêmes règles de viralité que les « fake news », disons, « classiques ». Cela signifie, comme l'avait démontré une étude du MIT, qu'elles se propagent six fois plus vite que les vraies informations. La lutte contre les *deep fakes* est devenue une course contre la montre. Et contre la technologie.

Qui sont les cerveaux de l'intelligence artificielle ?

PH.L.

L'intelligence artificielle (IA) porte mal son nom. A l'inverse de son modèle, l'intelligence humaine, elle n'a pas conscience d'elle-même. Il s'agit donc, juste, d'une technologie protéiforme, à la croisée des mathématiques et de l'informatique, visant à reproduire ou simuler l'intelligence humaine. Elle a connu un essor phénoménal grâce à l'avènement du « big data » et des capacités de calculs vertigineux des microprocesseurs. Les possibilités offertes par ces nouvelles IA donnent le tournis, notamment en termes de création d'images. Le type de rendu dépend des bases de données d'images auxquelles il a eu accès et avec lesquelles il s'est entraîné (des peintures de maîtres, des mangas, de la BD, des photos…)

Dall-E est l'acteur majeur de ce créneau. Mot-valise entre *Wall-e* (le film d'animation) et Savaldor Dali, ce programme est développé par la société OpenAI, fondée par Elon Musk en 2015 et soutenue par Microsoft, qui a injecté 1 milliard de dollars dans l'affaire. Il a pour but de concevoir des intelligences artificielles semblables au cerveau humain. Dall-E est capable de comprendre un texte écrit en langage naturel – comme une recherche sur Google – pour composer une image inédite. Dall-E ne colle pas ensemble des images qu'il a trouvées sur internet. Il a, au contraire, appris, en absorbant d'immenses bases de données visuelles, à quoi ressemblait un labrador, la couleur rouge, l'action de conduire un véhicule ou bien le style pop art. OpenAI a, lui, mis en place des garde-fous. Ses bases de données ont été nettoyées pour éliminer les images violentes, racistes ou sexualisées. La plateforme bloque des mots-clés, interdit tout usage de l'image des personnalités politiques et dit aussi se soucier des préjugés ou stéréotypes.

Sans garde-fous

Dall-E n'est, par ailleurs, réservé qu'à une petite élite, sur invitation uniquement. Mais ce n'est pas le cas de son Dall-E2, rebaptisée Craiyon, arrivé en avril dernier, dont les images ont vite inondé les réseaux sociaux. Sans garde-fous, cette fois. On ne sait pas si DALL-E sera un jour entièrement accessible au public.

OpenAI n'est pas seul à travailler sur la création automatisée d'images. Google, par exemple, a un projet similaire nommé Imagen. L'idée de ne laisser que les Gafam agir sur ce terrain sensible a aussi fait germer des projets « open source ».

C'est le cas d'un deuxième acteur majeur, lequel est aujourd'hui au cœur de nombreuses polémiques. Car qui dit « open source » dit forcément « ouvert » à quiconque et à n'importe quelle fin (son code est accessible gratuitement sur Github). Stable Diffusion est capable de créer d'incroyables œuvres d'art photoréalistes et artistiques qui n'ont rien à envier aux modèles d'OpenAI ou de Google. Il a été créé et développé par une start-up londonienne du nom de Stability AI. Fondée en octobre 2020 par Emad Mostaque, un gestionnaire de fonds spéculatifs, l'entreprise souhaite, selon ses propres mots, « construire des outils d'IA libres afin de poser les bases pour éveiller le potentiel de l'humanité ».