Contre la surveillance en ligne, des internautes « empoisonnent » leurs données personnelles

Le Monde - publié le 29 mai 2022
https://www.lemonde.fr/pixels/article/2022/04/29/contre-la-surveillance-en-ligne-des-internautes-empoisonnent-leurs-donnees-personnelles_6124107_4408996.html

Avec l'« obfuscation » ou le « data poisoning », ils redoublent d'efforts pour empêcher les entreprises de les traquer en ligne. Mais ces modes de résistance, chronophages, parfois très complexes, peinent à se populariser.

Pour Amir (le prénom a été modifié), tout est parti d'un cadeau : une belle bouteille de bière reçue pour son anniversaire. Quelques jours plus tard, alors qu'il cherche à en savoir plus dessus, cet étudiant népalais d'une vingtaine d'années se rend compte qu'il a été catégorisé *« intéressé par les boissons alcoolisées »* par Facebook. Parce que la plate-forme a dorénavant associé son profil aux mots-clés recherchés, le jeune homme se retrouve assailli, sur son fil d'actualité, de publicités ciblées en lien avec la boisson. Un paradoxe, quand on sait qu'il ne boit, en réalité, pas d'alcool.

C'est à ce moment précis qu'Amir prend la décision « d'empoisonner » ses données. L'idée ? Tromper les algorithmes des plates-formes en se faisant passer pour ce que l'on n'est pas. *« Qu'y a-t-il de pire pour eux que de ne fournir aucune donnée ? En fournir de mauvaises ! La chose la plus simple à faire est de simuler des intérêts différents : quand vous suivez deux partis politiques opposés, par exemple, l'intelligence artificielle est désorientée et ne sait pas ce que vous aimez réellement. »*

Le « droit à mentir »

Doit-on continuer à céder ses informations personnelles, qui ne sont pas des marchandises comme les autres, à des entreprises privées ? Comment se prémunir de cette surveillance en ligne constante et particulièrement opaque sans pour autant s'interdire d'utiliser Internet ? Comme Amir, de plus en plus d'internautes s'interrogent sur la protection de leur vie privée et décident de reprendre le pouvoir en pratiquant ce qu'ils appellent l'*obfuscation* (offuscation ou obscurcissement) ou le *data poisoning*, l'empoisonnement des données, laissant derrière eux de fausses informations, censées brouiller les pistes sur leur profil numérique, et ainsi s'offrir une forme de protection.

> AdNauseam est une extension de navigateur Web. Lorsqu'elle est installée, elle simule de faux clics sur chaque publicité croisée en ligne

Cette stratégie n'est pas nouvelle, d'après Helen Nissenbaum, professeure au département des sciences de l'information de l'université Cornell Tech (Etat de New York). *« Pendant la seconde guerre mondiale, quand les Alliés survolaient l'Allemagne pour y larguer des bombes, leurs avions relâchaient des paillettes, des morceaux de papier avec de l'aluminium*, détaille la chercheuse. *Pendant ces courts laps de temps, les Allemands ne pouvaient déterminer lesquels des points qui apparaissaient sur leurs radars antiaériens étaient des avions et lesquels étaient ce que l'on pourrait appeler "des données empoisonnées". »*

Helen Nissenbaum a longuement étudié les formes de résistance en ligne et a justement participé à la création d'outils d'offuscation. Le dernier en date, créé en 2014 et intitulé AdNauseam, est une extension de navigateur Web. Lorsqu'elle est installée, elle simule de faux clics sur chaque publicité croisée en ligne par l'utilisateur, ce qui le rend particulièrement difficile à profiler et cibler.

C'est à partir de cette même logique que le Français Vincent Toubiana a contribué au développement de l'outil créé par Daniel C. Howe et Helen Nissenbaum, TrackMeNot, il y a quelques années. *« L'idée, c'est de ne pas jouer avec les règles des acteurs qui nous surveillent, d'avoir droit de leur mentir et ainsi de les faire douter sur nos requêtes et nos profils et, idéalement, sur les requêtes et les profils de tout le monde parce qu'ils savent que certains mentent »*, explique celui qui est à présent employé de la Commission nationale de l'informatique et des libertés (CNIL) et directeur de son laboratoire d'innovation numérique.

Piéger l'intelligence artificielle

Appelés *data poisoners* (« empoisonneurs de données »), certains utilisateurs vont plus loin que le simple brouillage de données en s'attaquant plus directement aux systèmes de collecte de données. On le sait, certaines entreprises comme YouTube, Netflix ou Spotify entraînent leurs modèles de *machine learning* sur des informations issues de profils et de comportements d'internautes. Or c'est précisément ces cibles que convoitent les *data poisoners* : ils créent de fausses données d'entraînement, censées être indétectables, et les font ingérer à ces systèmes afin de les biaiser et de les rendre inefficaces.

> Télécharger des centaines d'images de chiens et de les étiqueter comme étant des oiseaux

Ecouter à la fois un genre de musique que l'on aime et un genre que l'on déteste sur une plate-forme de streaming, par exemple, revient à « polluer » l'ensemble de ses données d'entraînement et perturber ses systèmes de recommandation. *« C'est la méthode la plus simple à utiliser. Mais il en existe de bien plus sophistiquées, qui assemblent de fausses séquences de données pour enseigner à la machine des choses qui ne sont pas vraies*, précise Nick Vincent, doctorant de l'université Northwestern (Illinois), qui a analysé ces pratiques. *Une attaque parfaite d'empoisonnement nécessite d'être sûr que les données que l'on injecte vont alimenter l'entraînement du modèle et donc de savoir comment ce dernier fonctionne. »* Plus complexes que l'ajout d'une extension de navigateur, ces pratiques requièrent des connaissances techniques mais peuvent se révéler très

efficaces, même à partir d'une faible quantité de données polluées, selon le chercheur.

Une manière simple d'en vérifier l'efficacité est d'utiliser AutoML, l'outil de Google qui sert à entraîner des modèles de *machine learning* personnalisés. Par exemple, si l'on veut s'attaquer à un système de classification d'images par ordinateur, il est possible de télécharger des centaines d'images de chiens et de les étiqueter comme étant des oiseaux. A terme, lorsque la machine devra reconnaître un chien, quel qu'il soit, elle estimera à tort que c'est un oiseau parce qu'elle aura été entraînée ainsi. A la différence de l'offuscation, qui passe aussi par la création de fausses données, cette technique repose sur le façonnement biaisé d'une intelligence artificielle, laquelle est censée reproduire des erreurs quand elle fonctionne.

Un sacerdoce qui demande de s'adapter

Mais le *data poisoning* n'est pas l'apanage de ceux qui souhaitent lutter contre la surveillance en ligne et est parfois utilisé à des fins malhonnêtes. Empoisonner les données d'un hébergeur d'e-mails, par exemple en utilisant de façon massive le bouton « Ne pas signaler comme spam » sur un jeu de données contenant des messages frauduleux, peut ainsi permettre à des personnes mal intentionnées de faire parvenir plus facilement leurs tentatives d'hameçonnage jusque dans les boîtes courriels de leurs victimes. Raison pour laquelle Elie Bursztein, chargé de la sécurité et de la lutte contre les abus chez Google, concédait, en 2018, dans un billet de blog, devoir être particulièrement attentif au *data poisoning*.

Résultat : les multinationales du numérique ont développé, au fil du temps et des attaques, des stratégies pour protéger leurs modèles, si bien que le combat des *data poisoners* simplement motivés par leur vie privée peut sembler perdu d'avance. Les plus déterminés doivent constamment s'adapter, au point que créer des jeux de données frelatées ou polluer ses propres informations personnelles devient une activité à temps plein, estime Vincent Toubiana. *« Taper des mots-clés qui ne nous intéressent pas, suivre des personnes sur les réseaux sociaux puis les masquer, ne pas donner de vraies informations quand elles ne sont pas nécessaires… tout cela est chronophage. »*

> Il est possible d'être radié de certains services car le « data poisoning » enfreint la plupart de leurs règles

Empoisonner ses données revient par ailleurs à sacrifier ses propres habitudes en ligne. Lutter contre un algorithme de recommandation, comme celui d'Amazon, par exemple, équivaut à se priver de fonctionnalités fondées sur la personnalisation de l'offre, comme des publicités ou des promotions ciblées. Pire, selon Nick Vincent, il est même possible d'être radié de certains services car le *data poisoning* enfreint la plupart de leurs règles : *« Si je crée un programme qui automatise ma fausse consommation de YouTube, il y a de très fortes chances que je sois repéré comme étant un "bot", et banni. »*

La méthode la plus sûre et la plus accessible reste alors l'extension de navigateur, qui connaît un succès grandissant. Selon Helen Nissenbaum, la dernière version d'AdNauseam a été téléchargée plus de 350 000 fois. L'extension sert par ailleurs de modèle à d'autres outils, comme l'« offuscateur » Cookie Factory, créé par l'Unesco et mis en ligne en novembre 2021. L'extension, trouvable sur le navigateur Chrome, permet de créer de faux profils d'utilisateurs par la manipulation de cookies.

L'enjeu : sensibiliser

Vincent Toubiana aimerait, lui aussi, développer des outils d'obfuscation, cette fois sous la houlette de la CNIL. Mais, là encore, il faut se frotter aux géants du numérique, comme Google qui a banni de Chrome l'extension AdNauseam, et trouver des investisseurs intéressés. Francis Hunger, artiste allemand derrière l'application Adversarial.io, ne trouve par exemple aucun repreneur pour son projet d'obfuscation d'images. *« Notre site ne circule pour l'instant que dans des festivals ou des*

conférences, indique-t-il au *Monde*. *Mon souhait est qu'une grande entreprise ajoute notre fonctionnalité à ses services pour que les gens soient plus avertis. »*

> La protection des données personnelles peine encore à se faire une place sérieuse dans le débat public

Là se situe en effet l'enjeu principal : sensibiliser. Certes, de nombreux scandales ont contraint, ces dernières années, les sphères politique et médiatique à s'emparer du sujet – en témoigne l'entrée en vigueur en Europe du Règlement général sur la protection des données (RGPD) et d'une nouvelle version du Data Protection Act, au Royaume-Uni, quelques semaines seulement après la publication de révélations sur la société britannique Cambridge Analytica et ses liens étroits avec Facebook. Mais, malgré cela, la protection des données personnelles peine encore à se faire une place sérieuse dans le débat public.

Dans une étude menée en 2018 pour la société ForgeRock auprès d'Américains et d'Européens – dont des Français –, 57 % des sondés estimaient en savoir *« un peu »* ou n'y connaître *« rien du tout »* sur leurs droits concernant l'utilisation de leurs données personnelles. *« Nous savons, depuis des années, que la plupart n'ont pas conscience de ces pratiques de collecte et d'utilisation de données parce qu'elles sont évidemment occultes et incroyablement complexes »*, analyse la chercheuse Helen Nissenbaum, qui voit en cette méconnaissance un frein manifeste à la lutte contre la surveillance en ligne.