

Fake news Quesaco? 2/

Les avancées technologiques vont rendre la réalité de plus en plus difficile à distinguer

Source: Courrier International Août 2018



Depuis des mois, les fake news s'affichent en une : leurs dégâts sont connus et des réponses légales font leur chemin. Mais c'est sans compter sur les avancées technologiques qui vont rendre la réalité de plus en plus difficile à distinguer, notamment dans le domaine de la vidéo. La résistance prend forme avec des penseurs et des informaticiens qui, tel [Aviv Ovadya](#), estiment que la lutte contre la désinformation devrait devenir une priorité dans nos démocraties.

Sur certains sites sordides du Net, les actrices de *Game of Thrones* ou de *Harry Potter* se livrent à toutes sortes d'actes sexuels. Du moins, les femmes de ces vidéos ressemblent à ces actrices, et les visages sont bien les leurs; mais tout ce qui se trouve au-dessous du cou appartient à d'autres. Une intelligence artificielle a glissé ces visages familiers dans des scènes pornographiques, a remplacé leurs traits par d'autres. C'est là l'une des formes les plus cruelles, les plus invasives, de vol d'identité qui aient été inventées à l'ère d'Internet. Et c'est la précision de la technique qui rend la chose possible : la supercherie ne se remarque que de très près.

Ces vidéos sont l'œuvre d'un programmeur dont le nom de pirate est "[Deepfakes](#)". Et ce n'est là qu'une version bêta d'un projet bien plus ambitieux. L'un des acolytes de Deepfakes a confié sur le site *Motherboard* en janvier que celui-ci avait l'intention de démocratiser son œuvre. Il compte affiner le processus, l'automatiser, ce qui permettrait à n'importe qui de coller la tête d'un béguin, d'un ex ou d'un collègue dans un clip pornographique en quelques manipulations. Aucune connaissance technique ne serait requise. Et comme les labos de recherche universitaires et commerciaux développent des outils encore plus sophistiqués à des fins non pornographiques – des algorithmes qui cartographient les expressions du visage et imitent la voix avec précision –, ces fakes sordides présenteront bientôt une vraisemblance encore plus grande.

Internet contient depuis toujours les graines de l'enfer postmoderne. La manipulation massive est courante sur ce média, des pièges à clics pour logiciels russes aux ruses addictives qui gouvernent le fil d'actualité de Facebook. Internet a toujours été un lieu où l'identité est fuyante, où l'anonymat engendre grossièretés et confusion, où les escrocs peuvent voler les éléments mêmes de l'individualité. À cet égard, les *deepfakes* marquent l'apothéose de son histoire – et ne sont probablement qu'une version gentille de ce qui nous attend.

[L'écrivain américain] Vladimir Nabokov a écrit un jour que le mot "réalité" était l'un des rares qui ne signifiaient rien sans guillemets. Il évoquait la relativité de la perception : quand vous et moi regardons le même objet, comment pouvez-vous savoir que nous voyons la même chose? Pourtant, les institutions (médias, gouvernement, recherche) fédèrent les gens autour d'un consensus – enraciné dans la foi en la raison et l'empirisme – sur la façon de décrire le monde, même s'il révèle sa fragilité depuis quelques années. Les réseaux sociaux ont contribué à l'avènement d'une nouvelle ère où chacun peut être en contact avec des informations qui confirment ses préjugés et occultent les faits qui les infirment. Le président Donald Trump accélère encore l'arrivée d'un monde au-delà de la vérité et donne l'imprimatur de la fonction suprême au mensonge et à la manipulation.

Tout cela semblera pourtant bientôt une époque d'innocence. Nous vivons dans un monde où nos yeux nous tromperont régulièrement. En d'autres termes, nous ne sommes pas loin de l'effondrement de la réalité.

Aujourd'hui, nous nous accrochons à la réalité, nous la désirons même ardemment. Nous vivons encore dans le monde d'Abraham Zapruder. La vidéo des huit minutes de l'assassinat de John Fitzgerald Kennedy filmées par hasard par la caméra personnelle de cet habitant de Dallas est l'exemple même de ce que nous vénérons. La vidéo brute, non travaillée, possède dorénavant une autorité gigantesque dans notre culture. Car les gens font désormais preuve d'un cynisme aveugle, irrationnel envers les reportages et autres matériaux que les médias manipulent et traitent – une réaction excessive dans un siècle de publicité, de propagande et d'informations télévisées hyperboliques.

Les comportements scandaleux suscitent plus sûrement l'indignation quand ils sont filmés. Ce type de vidéo a joué un rôle décisif pendant les deux dernières élections présidentielles américaines. En 2012, un barman qui officiait lors d'un événement destiné à collecter des fonds pour Mitt Romney en Floride a appuyé sur la touche "enregistrer" de sa caméra pendant que le candidat dénonçait les "47 %" d'Américains qui soutenaient Obama et constituaient selon lui un ramassis d'assistés dépendant du gouvernement fédéral. Il est plus que probable que cette vidéo capturée furtivement a ruiné ses chances de devenir président. Ces propos n'auraient certainement pas fait autant de bruit s'ils avaient été notés puis publiés par un journaliste. Les images vidéo – avec leur angle de prise de vues indirect, le tintement des couverts et les serveurs qui passent avec des serviettes pliées – étaient bien plus puissantes. Tous les détails témoignaient de son origine de façon incontestable.

Quand vous et moi regardons le même objet, comment pouvez-vous savoir que nous voyons la même chose?

Certes, Donald Trump s'est remis de la vidéo [diffusée par l'émission] *Access Hollywood* dans laquelle il se vantait d'avoir agressé des femmes sexuellement. Mais ce film a davantage excité les passions et la conscience de l'opinion que tout autre sujet lors de la campagne présidentielle de 2016. De même, la vidéo a été par procuration l'élément déclencheur de nombre d'autres conflagrations sociales récemment. Il a fallu les images des caméras de surveillance montrant le joueur de football américain Ray Rice traînant sa femme inconsciente hors de l'ascenseur d'un hôtel [en 2014] pour que la NFL se décide enfin à réagir sérieusement au problème de la violence conjugale. Puis il y a eu la vidéo du meurtre de Philando Castile diffusée en direct sur Facebook par son amie en 2016. Tous les reportages du monde, toutes les statistiques affolantes et les anecdotes bouleversantes sur les violences policières n'avaient pas provoqué autant d'indignation. Les images de la mort de Philando, abattu comme un chien dans son Oldsmobile par un policier dans le Minnesota [en juillet 2016], ont bouleversé l'opinion et poussé les personnalités politiques, y compris quelques commentateurs conservateurs purs et durs, à reconnaître enfin ce genre d'abus.

Ce qui nous amène au cœur du problème. Il est naturel de faire confiance à ses sens, de croire ce qu'on voit. C'est une tendance profondément enracinée que la future ère de la manipulation vidéo va exploiter. Considérez les vidéos qui ont récemment provoqué ce qu'Aviv Ovadya, de l'université du Michigan, appelle une "infocalypse": les choses auraient pu être bien pires avec un film manipulé. Prenez le Pizzagate [une théorie conspirationniste qui mettait l'ancien conseiller d'Hillary Clinton au cœur d'un réseau pédophile], et ajoutez-y des images truquées de John Podesta en train de lorgner un enfant avec concupiscence, ou pire. Le mensonge acquerra soudain une nouvelle intensité émotionnelle.

Les vidéos truquées créeront une suspicion légitime à propos de tout ce que nous regardons.

Le problème, ce n'est cependant pas simplement la prolifération du mensonge. Les vidéos truquées créeront une suspicion légitime à propos de tout ce que nous regardons. Les politiques et les journalistes exploiteront ces doutes. Un coupable filmé sur le fait déclarera simplement que la preuve visuelle est un montage destiné à lui nuire.

Le président semble déjà avoir adopté cette tactique : s'il a commencé par reconnaître l'authenticité de la vidéo d'*Access Hollywood*, il doute en privé que la voix qu'on entend soit bien la sienne.

En d'autres termes, les vidéos manipulées détruiront toute foi en notre lien avec la réalité ordinaire. Pour reprendre les propos tenus par Ian Goodfellow, un scientifique de Google, à la *MIT Technology Review* : "C'est quand même un coup de chance extraordinaire historiquement parlant qu'on puisse s'appuyer sur des vidéos pour prouver que quelque chose s'est vraiment passé."

Réalité virtuelle. L'effondrement de la réalité n'est pas une conséquence involontaire de l'intelligence artificielle. C'est depuis longtemps un des objectifs – ou du moins une aspiration – de certains des architectes les plus illustres des nouvelles technologies. Le parcours de la Silicon Valley commence à plus d'un égard au début des années 1960 avec la Fondation internationale pour les études avancées. La fondation se spécialisait dans les expériences sur le LSD. Ses développeurs voulaient créer des machines capables de modifier la conscience comme le font les drogues : les ordinateurs allaient eux aussi creuser un trou dans la réalité et arracher l'humanité à la banalité gris flanelle du quotidien de l'Amérique de [la série télévisée familiale] *Leave It to Beaver* pour l'amener à un état d'esprit plus cool et plus holistique. Steve Jobs a confié que le LSD avait été *“l'une des deux ou trois expériences les plus importantes de sa vie”*.

Contrairement à ce que les adeptes des nouvelles technologies voudraient nous faire croire, les vidéos réalistes mais truquées ne marquent pas le terminus de la fuite hors de la réalité. L'apothéose, c'est la réalité virtuelle [*virtual reality* en anglais, VR]. Celle-ci a pour objectif de créer l'illusion complète d'être ailleurs. Avec son casque et ses gants, elle cherche à tromper nos sens et à pervertir nos perceptions. Les jeux vidéo ont commencé à transporter les joueurs dans d'autres mondes. Or si les jeux peuvent être addictifs, ils ne sont pas encore pleinement immersifs. La VR est capable de vous transporter ailleurs plus complètement – nous voyons ce que voient nos avatars et sentons ce qu'ils sentent. Il y a plusieurs décennies, Timothy Leary, le chantre du psychédélisme, déclara, après avoir essayé cette technologie naissante, que c'était le *“nouveau LSD”*.

Maintenant que la réalité virtuelle sort de la petite enfance, la vie va peut-être y être plus intéressante.

Cependant, quand toute l'excitation sera retombée, elle se transformera en énorme industrie – comme l'ordinateur personnel ou les réseaux sociaux –, s'efforcera de rendre les consommateurs dépendants pour son propre bénéfice et sera peut-être dominée par une ou deux entreprises extraordinairement puissantes. Le fait que Facebook investisse dans la VR, en acquérant la société Oculus, n'est pas vraiment rassurant.

La capacité de manipulation augmentera parce que la VR crée par définition une confusion sur la réalité. Certains concepteurs de VR racontent que les utilisateurs ont parfois des réactions émotionnelles tellement fortes à des stimuli terrifiants qu'ils arrachent le casque pour s'enfuir. D'après des chercheurs allemands qui se sont efforcés de codifier l'éthique pour la VR, sa *“nature globale”* offre des *“possibilités de nouvelles formes de manipulation mentale et comportementale particulièrement puissantes, en particulier si la création et l'entretien des mondes virtuels sont le fait d'intérêts commerciaux, politiques, religieux ou gouvernementaux”*. Comme l'écrit Jaron Lanier, le pionnier de la VR, dans ses mémoires publiés récemment [*Dawn of the New Everything*, inédit en français] : *“Jamais un média n'a été autant à même de produire de la beauté et aussi vulnérable à la perversion. La réalité virtuelle va nous mettre à l'épreuve. Elle va amplifier notre caractère plus que les autres médias.”*

Confiance. La société trouvera peut-être le moyen de faire face à ces changements. Peut-être apprendrons-nous le scepticisme nécessaire pour nous y frayer un chemin. Jusqu'à présent cependant, les êtres humains ont manifesté une vulnérabilité pratiquement infinie à la tromperie et à l'escroquerie – ils se retrouvent aisément dans des mondes qui correspondent à leurs convictions ou à l'image qu'ils ont d'eux-mêmes, et peu importe que ces convictions soient excentriques voire carrément erronées. Les gouvernements ont mis du temps à réagir aux défis sociaux posés par les nouvelles technologies et préféreront peut-être éviter celui-ci. Décider de ce qui constitue la réalité n'est pas seulement une question épistémologique mais c'est aussi une question politique et elle impliquerait de déclarer spécieuses certaines convictions profondes.

Rares sont les individus qui auront le temps voire la capacité de distinguer une fabulation soigneusement élaborée de la vérité. Notre meilleur espoir, c'est peut-être de sous-traiter le problème, de redonner l'autorité culturelle à des institutions de confiance possédant la formation et les connaissances nécessaires pour validation : journaux, universités. Peut-être les grandes sociétés du secteur comprendront-elles le problème et joueront-elles aussi ce rôle. Comme elles contrôlent les points d'accès les plus importants à l'information, elles pourraient très facilement écraser les vidéos manipulées. Pour jouer ce rôle, il faudrait toutefois qu'elles acceptent certaines responsabilités qu'elles se sont jusqu'à présent largement refusées à prendre.

En 2016, pendant que la Russie utilisait Facebook pour influencer sur l'élection présidentielle américaine, Elon Musk confiait sa vision de la vie humaine lors d'une conférence. Selon la théorie qu'il a adoptée, qui est dérivée d'un philosophe d'Oxford, nous vivons en fait dans une simulation informatique, nous sommes déjà les personnages d'un film de science-fiction ou d'un jeu vidéo. Il a déclaré : *“La probabilité que nous nous trouvions dans la ‘réalité de base’ est d'une sur des milliards.”* Si les patrons de l'industrie qui président à notre information et espèrent définir notre avenir ne sont même pas capables de reconnaître l'existence de la réalité, il y a peu d'espoir de la sauver.

DÉTOURNEMENTS INTERDITS

Jusqu'à sa fermeture, en février 2018, une section de la célèbre plateforme de discussion Reddit, consacrée aux [*deepfakes*](#), comptait 91000 utilisateurs inscrits. Elle permettait de remplacer le visage d'une personne

par une autre, et était utilisée principalement pour des vidéos pornographiques. Face à l'attrait croissant du logiciel, la plateforme a fermé le fil de discussion et changé ses conditions d'utilisation. Il est désormais interdit de publier des images ou des vidéos de personnes nues sans leur permission.

La boîte à outils du fabricant de fake news

1. Le programme [DeepWarp](#) permet de modifier la direction des yeux d'une personne sur une photo.
2. Des chercheurs de la société américaine Nvidia ont conçu un système capable de manipuler les conditions climatiques sur une vidéo déjà enregistrée.
3. Le logiciel [Face2Face](#) a la capacité de changer les expressions du visage d'un individu sur une vidéo.
4. Une équipe de l'université de Washington a reproduit le visage de Barack Obama en 3D puis a manipulé ses lèvres pour les synchroniser avec un montage sonore.
5. La société américaine Adobe a présenté en 2016 le [prototype Voco](#), capable de transformer des extraits audio en ajoutant des mots qui n'ont pas été prononcés.
6. À l'aide de l'intelligence artificielle, la société Nvidia a créé un système capable de créer de faux visages inspirés de célébrités.
7. Des chercheurs de Google ont mis au point [Tacotron 2](#), qui peut produire une voix artificielle ressemblant à s'y méprendre à celle d'un humain.
8. Le logiciel [Ultimate 12](#) facilite grandement l'incrustation sur fond vert d'éléments qui n'étaient pas présents sur le plateau de tournage.

Suite: [Internet, le Progrès et les Fake News](#)