LIEN ARTICLE. Infocalypse now

Après l'infobésité, l'infocalypse.
Un article qui invite à penser ce qui pourrait arriver avec l'amplification de fake news produites par des logiciels de plus en plus perfectionnés du côté de l'audio-vidéo. Nous verrons des discours qui n'ont jamais eu lieu, nous réagirons à des événements qui n'ont jamais eu lieu… jusqu'au moment où dégoutés, *incapables de savoir ce qui a lieu ou pas*, de plus en plus de personnes délaisseront l'information, ne penseront plus avec un monde qui leur échappe, et laisseront donc faire… comme si nous vivions dans une grande fiction.

" Aviv Ovadya avait prédit la crise des Fake News de 2016. Aujourd'hui il annonce une apocalypse de l'information. L'article complet est paru sur le site buzzfeed.

Des extraits ci-dessous.

« Que se passe-t-il quand n'importe qui peut faire croire que n'importe quoi est arrivé, que ce soit vrai ou pas ? »

Publié le 16 Février, 2018 à 4:03 p.m.

"À la mi-2016, Aviv Ovadya (…) a exposé ses inquiétudes à des professionnels du secteur des nouvelles technologies de la région de la Baie de San Francisco et averti de l'imminence d'une crise de la désinformation lors d'une présentation qu'il a intitulée « Infocalypse. »

Dans cet exposé, Aviv Ovadya expliquait que le web, et tout l'écosystème de l'information qui s'est développé autour, sont terriblement malsains. Les motivations qui régissent ses plus grandes plateformes sont calibrées pour **récompenser des informations souvent trompeuses, polarisantes ou les deux.** Des plateformes comme Facebook, Twitter et Google donnent la priorité aux clics, aux partages, aux publicités et à l'argent au détriment de la qualité de l'information, et Aviv Ovadya ne parvient pas à se débarrasser de la sensation que tout cela va déboucher sur quelque chose de néfaste : un genre de **seuil critique de désinformation addictive et toxique**. (…)

Stephen Lam pour BuzzFeed News : (…) Aviv Ovadya a vu très tôt ce que beaucoup — y compris des législateurs, des journalistes et des PDG des géants technologiques — allaient mettre des mois à comprendre : notre monde de plateformes, optimisé par des algorithmes, est vulnérable. Perméable à la propagande, à la désinformation, à la publicité ciblée malintentionnée de gouvernements étrangers. Au point qu'il **menace de saper une des pierres angulaires du discours humain : la crédibilité des faits.**
Mais c'est ce qu'il a anticipé qui donne vraiment des sueurs froides.

« L'alarmisme peut être une bonne chose, il faut être alarmiste pour ces trucs-là », nous explique Aviv Ovadya une après-midi de janvier, avant d'exposer calmement les grandes lignes d'une **projection extrêmement dérangeante pour les vingt prochaines années, faite de fake news, de campagnes de désinformation assistées par l'intelligence artificielle (IA) et de propagande.** (…)

Cet avenir, selon Aviv Ovadya, adviendra accompagné de tout un tas d'outils technologiques sophistiqués, faciles à utiliser et, au final, indécelables qui serviront à **manipuler la perception et à falsifier la réalité**, et pour lequel des expressions ont déjà été inventés : « apathie face à la réalité », « phishing laser automatisé » et « marionnettes humaines. »

Il est devenu clair à ses yeux que si quelqu'un devait exploiter l'économie de l'attention et utiliser les plateformes qui la soutiennent pour déformer la vérité, il n'existait aucun vrai moyen de contrôle ou de contre-pouvoir pour l'arrêter. « Je me suis rendu compte que si ces systèmes échappaient à tout contrôle, il n'y aurait rien pour les réfréner et que ça tournerait mal, très rapidement » explique-t-il. (…)

Pour Aviv Ovadya — aujourd'hui responsable nouvelles technologies du Center for Social Media Responsibility de l'Université du Michigan et membre du comité d'innovation Knight News du Tow Center for Digital Journalism de Columbia — le choc et l'inquiétude qu'ont généré les publicités russes sur Facebook et les bots sur Twitter font pâle figure au regard d'une menace bien plus inquiétante : **les technologies susceptibles d'être utilisées pour augmenter et déformer la réalité évoluent plus vite que notre capacité à les comprendre, à les contrôler ou à en atténuer les effets.**

Les enjeux et les éventuelles conséquences sont plus catastrophiques qu'une ingérence étrangère dans une élection. Il s'agit de la sape ou du bouleversement d'institutions au cœur de notre civilisation : d'une « infocalypse ». Et Aviv Ovadya affirme que celle-là est tout aussi plausible que la précédente — qu'elle serait pire encore.

Pire à cause de nos prouesses informatiques toujours plus étendues, pire à cause des avancées actuelles dans les domaines de l'intelligence artificielle et de l'apprentissage automatique qui peuvent **brouiller les lignes entre les faits et la fiction**, pire parce que ces choses pourraient faire naître un avenir où, comme l'observe Aviv Ovadya, n'importe qui pourrait « faire croire que quelque chose, n'importe quoi, est arrivé, que ce soit vrai ou pas ».

(…) Aviv Ovadya prévient que les outils qui se développent rapidement et fonctionnent grâce à l'intelligence artificielle, à l'apprentissage automatique et aux technologies de réalité augmentée pourraient être piratés et utilisés par des agents mal intentionnés pour imiter les humains et mener une guerre de l'information. (…) **Certains outils disponibles pour manipuler des vidéos et des enregistrements audio, sont dans le domaine des fake news, de la même envergure que la bombe atomique dans le domaine militaire.** Dans les recoins ténébreux d'internet, des gens se sont mis à utiliser des algorithmes d'apprentissage automatique et des logiciels open source pour créer facilement des vidéos pornographiques qui superposent de façon tout à fait réaliste les visages de célébrités — ou de n'importe qui d'ailleurs— et les corps d'acteurs pornos. Dans des institutions comme Stanford, des spécialistes des nouvelles technologies ont mis au point des programmes capables de manipuler des vidéos en associant et en mélangeant des séquences filmées et des techniques de suivi de visage. Dans le même genre, à l'université de Washington des informaticiens ont réussir à élaborer un programme apte à « transformer des séquences audio en vidéo réaliste et synchronisée de la personne qui parle ». Pour prouver la validité du concept, les deux équipes ont manipulé des vidéos existantes où ils font dire à des chefs d'État des paroles qu'ils n'ont jamais prononcées.

Ces outils se démocratisent et deviennent de plus en plus répandus, ce qui pousse Aviv Ovadya à craindre le pire avec des scénarios qui pourraient s'avérer extrêmement déstabilisant.

Il cite par exemple **la « manipulation diplomatique », où un agent malveillant utilise des technologies avancées pour « faire croire qu'un événement s'est produit » afin d'influencer la géopolitique.**

Imaginez par exemple un algorithme d'apprentissage automatique (qui analyse des masses de données pour *s'enseigner à lui-même* à accomplir une fonction particulière) alimenté par des centaines d'heures d'enregistrements de Donald Trump ou du dictateur nord-coréen Kim Jong Un, et qui serait ensuite capable de recracher un fichier audio ou vidéo quasi-parfait — et pratiquement impossible à distinguer d'un vrai — du dirigeant déclarant la guerre nucléaire ou biologique. « Cela n'a pas besoin d'être parfait, juste assez bien pour que l'ennemi pense qu'il s'est passé quelque chose afin de **déclencher une réponse-réflexe irréfléchie** en représailles. »

(…)

Au-delà de tout cela, il existe d'autres scénarios cauchemardesques à long terme qu'Aviv Ovadya estime « tirés par les cheveux », mais pas au point qu'il envisage de les écarter totalement. Et ils font peur. Les « marionnettes humaines » par exemple : une version clandestine d'un marché des réseaux sociaux, avec des gens à la place des bots. « Fondamentalement, il s'agit d'un futur marché transnational développé pour humains manipulables », explique-t-il.

Les prémonitions d'Aviv Ovadya sont particulièrement terrifiantes étant donné la facilité avec laquelle notre démocratie a déjà été manipulée par les techniques de désinformation brutales et rudimentaires. Les arnaques, les supercheries et le brouillage de piste à venir n'ont rien d'une nouveauté. Ils sont juste plus sophistiqués, plus difficiles à détecter et travaillent main dans la main avec d'autres forces technologiques qui ne sont pas inconnues aujourd'hui mais qui sont susceptibles d'être imprévisibles. (…)

Dans certains cas, la technologie est tellement performante qu'elle va jusqu'à surprendre ses propres créateurs. Ian Goodfellow, un scientifique et chercheur Google Brain qui a contribué à coder le premier « réseau contradictoire générateur » (GAN), **un réseau neuronal capable d'apprendre sans supervision humaine**, a averti que l'IA pourrait renvoyer la consommation d'informations environ 100 ans en arrière. Lors d'une conférence de la MIT Technology Review en novembre dernier, il a expliqué au public que les GAN faisaient à la fois preuve « d'imagination et d'introspection » et « pouvaient prendre connaissance de l'état du générateur sans avoir besoin de feedback humain. » Et que si les possibilités créatives sont illimitées pour les machines, l'innovation appliquée à la manière dont nous consommons l'information « fermerait sans doute certaines des portes que notre génération a eu l'habitude de voir ouvertes ».

(…) Les scénarios pessimistes d'Aviv Ovadya semblent tout à fait plausibles dans le champ politique notamment. Cet été, plus d'un million de faux comptes bots ont inondé le système de commentaires de la Federal Communications Commission, l'équivalent du CSA américain, pour «amplifier l'appel à abroger les protections de la neutralité du net. » Les chercheurs ont conclu que les **commentaires automatisés** — qui grâce à un traitement automatique reproduisent un style d'écriture naturelle pour avoir l'air réels — cachaient les commentaires authentiques et minaient la légitimité de tout le système de commentaires. Pour Aviv Ovadya, l'exemple de la FCC ainsi que la récente campagne #releasethememo sur le Twitter américain, amplifiée par les bots, est une version grossière de ce qui nous attend. « Ça peut être tellement pire », explique-t-il.

Cette érosion de l'authenticité et de l'intégrité du discours officiel est la plus sinistre et la plus inquiétante des

menaces à venir. « Que ce soit l'IA, de drôles de manipulations Amazon ou du faux activisme politique, ces fondations technologiques (conduisent) à une érosion croissante de la confiance », explique Renee DiResta, chercheuse experte en propagande informatique, au sujet de la future menace. « Cela rend possible de jeter le doute sur l'authenticité de vidéos ou de plaidoyers. » (…)

« Vous n'avez pas besoin de créer une fausse vidéo pour que cette technologie ait un impact sérieux. Il suffit de signaler le fait que la technologie existe pour mettre en doute l'intégrité de ce qui est réel. »

(…) « Au cours des deux, trois, quatre prochaines années, nous allons devoir gérer des propagandistes amateurs capables de gagner des fortunes en créant des simulations photos réalistes, extrêmement réalistes », explique Justin Hendrix, directeur exécutif du NYC Media Lab, à BuzzFeed News. « Et si ces tentatives fonctionnent, et que les gens en viennent à soupçonner qu'il n'y a aucune réalité sous-jacente aux objets médiatiques quels qu'ils soient, alors nous allons vraiment avoir des problèmes. Il suffira d'un ou deux gros hoax pour **finir de convaincre le public que rien n'est vrai.** » (…)

Ceci dit, Aviv Ovadya concède qu'il reste une petite place à l'optimisme. L'intérêt prêté à l'espace de propagande informatique est plus grand qu'il ne l'a jamais été, et ceux qui tardaient autrefois à prendre les menaces au sérieux sont aujourd'hui plus réceptifs aux avertissements. « Au début c'était vraiment sinistre — quasiment personne n'écoutait », se souvient-il. « Mais les derniers mois ont été réellement prometteurs. Certains contre-pouvoirs sont en train de se mettre en place. » De même, il y a des solutions envisageables — comme la vérification cryptographique des images et des fichiers audio, qui pourrait aider à distinguer ce qui est réel de ce qui a été manipulé. (…)"