Désinformation, le rapport – 3

La traduction suivante est la suite et la continuation du travail entamé la semaine dernière sur le long rapport final élaboré par le comité « Digital, Culture, Media and Sport » du Parlement britannique, publié le 14 février dernier, sur la désinformation et la mésinformation.

Maintenant que le décor est posé, on aborde les questions réglementaires. Après avoir clairement défini ce qu'est une fake news*, que nous avons traduit par « infox » et que les auteurs regroupent sous le terme plus précis de « désinformation », il est question de définir une nouvelle catégorie de fournisseurs de service pour caractériser leur responsabilité dans les préjudices faits à la société ainsi que des solutions pour protéger le public et financer l'action des structures de contrôle.*

Le groupe Framalang a en effet entrepris de vous communiquer l'intégralité du rapport en feuilleton suivant l'avancement de la traduction.

Vous trouverez le texte intégral en suivant ce lien vers le PDF original (3,8 Mo).

La traduction est effectuée par le groupe Framalang, avec l'aide de toutes celles et ceux qui veulent bien participer et pour cet opus :

Traducteurs : Khrys, Lumibd, Maestox, simon, Fabrice, serici, Barbara, Angie, Fabrice, simon

La réglementation, le rôle, la

définition et la responsabilité juridique des entreprises de technologie

Définitions

11. Dans notre rapport intermédiaire, nous avons désavoué le terme d'« infox » puisqu'il a « pris de nombreux sens, notamment une description de toute affirmation qui n'est pas appréciée ou en accord avec l'opinion du lecteur » et nous avons recommandé à la place les termes de « mésinformation » ou de « désinformation ». Avec ces termes viennent « des directives claires à suivre pour les compagnies, organisations et le Gouvernement » liées à « une cohérence partagée de la définition sur les plateformes, qui peuvent être utilisées comme la base de la régulation et de l'application de la loi »[1].

12. Nous avons eu le plaisir de voir que le Gouvernement a accepté notre point de vue sur le fait que le terme « infox » soit trompeur, et ait essayé à la place d'employer les termes de « désinformation » et de « mésinformation ». Dans sa réponse, le gouvernement a affirmé :

> Dans notre travail, nous avons défini le mot « désinformation » comme la création et le partage délibérés d'informations fausses et/ou manipulées dans le but de tromper et d'induire en erreur le public, peu importe que ce soit pour porter préjudice, ou pour des raisons politiques, personnelles ou financières. La « mésinformation » se réfère au partage par inadvertance de fausses informations[2].

13. Nous avons aussi recommandé une nouvelle catégorie d'entreprises de réseaux sociaux, qui resserrent les responsabilités des entreprises de technologie et qui ne sont

pas forcément « une plateforme » ou un « éditeur ». le gouvernement n'a pas du tout répondu à cette recommandation, mais Sharon White, Pdg de Of.com a qualifié cette catégorie de « très soignée » car les « plateformes ont vraiment des responsabilités, même si elles ne génèrent pas les contenus, concernant ce qu'elles hébergent et promeuvent sur leur site ».[3].

14. Les entreprises de réseaux sociaux ne peuvent se cacher derrière le fait qu'elles seraient simplement une plateforme, et maintenir qu'elles n'ont elles-mêmes aucune responsabilité sur la régulation du contenu de leurs sites. Nous répétons la recommandation de notre rapport provisoire, qui stipule qu'une nouvelle catégorie d'entreprises technologiques doit être définie qui renforcera les responsabilités des entreprises technologiques et qui ne sont pas forcément « une plateforme » ou un « éditeur ». Cette approche voudrait que les entreprises de technologie prennent leur responsabilité en cas de contenu identifié comme étant abusif après qu'il a été posté par des utilisateurs. Nous demandons au gouvernement de prendre en compte cette nouvelle catégorie de compagnies technologiques dans son livre blanc qui va paraître prochainement.

Préjudices et réglementation en ligne

15. Plus tôt dans le cadre de notre enquête, nous avons écouté le témoignage de Sandy Parakilas et Tristan Harris, qui étaient tous deux à l'époque impliqués dans le Center for Human Technology, situé aux États-Unis. Le centre a compilé un « Recueil de Préjudices » qui résume les « impacts négatifs de la technologie qui n'apparaissent pas dans les bilans des entreprises, mais dans le bilan de la société ».[4] Le Recueil de Préjudices contient les impacts négatifs de la technologie, notamment la perte d'attention, les problèmes de santé mentale, les confusions sur les relations personnelles, les risques qui pèsent sur nos démocraties et les problèmes qui

touchent les enfants.[5].

16. La prolifération des préjudices en ligne est rendu plus dangereuse si on axe des messages spécifiques sur des individus suite à des « messages micro-ciblés », qui jouent souvent sur les opinions négatives qu'ont les gens d'eux-mêmes et des autres et en les déformant. Cette déformation est rendue encore plus extrême par l'utilisation de « deepfakes » [6] audio et vidéos qui sonnent et ressemblent à une personne existante tenant des propos qui ne lui appartiennent pas.[7] Comme nous l'avons dit dans notre rapport intermédiaire, la détection de ces exemples ne deviendra que plus complexe et plus difficile à démasquer au fur et à mesure de la sophistication des logiciels [8].

17. Le ministre de la santé, le député Hon Matthew Hancock, a récemment mis en garde les sociétés informatiques, notamment Facebook, Google et Twitter, qu'elles étaient en charge de la suppression des contenus inappropriés, blessants suite à la mort de Molly Russel, qui à 14 ans s'est suicidée en novembre 2017. Son compte Instagram contenait du contenu en lien avec la dépression, l'auto-mutilation et le suicide. Facebook, propriétaire d'Instagram, s'est déclaré profondément désolé de l'affaire.[9] Le directeur d'Instagram, Adam Mosseri, a rencontré le secrétaire de la Santé début février 2019 et déclaré qu'Instagram n'était pas « dans une situation où il était nécessaire de traiter le problème de l'auto-mutilation et du suicide » et que cela revenait à arbitrer entre « agir maintenant et agir de manière responsable » [10]

18. Nous relevons également que dans son discours du 5 février 2019, la députée Margot James, ministre du numérique dans le département du numérique, de la culture, des médias et du sport a exprimé ses craintes :

> La réponse des principales plateformes est depuis trop

> longtemps inefficace. Il y a eu moins de 15 chartes de bonne conduite mises en place volontairement depuis 2008. Il faut maintenant remettre absolument en cause un système qui n'a jamais été suffisamment encadré par la loi. Le livre blanc, que le DCMS produit en collaboration avec le ministère de l'intérieur sera suivi d'une consultation durant l'été et débouchera sur des mesures législatives permettant de s'assurer que les plateformes supprimeront les contenus illégaux et privilégieront la protection des utilisateurs, particulièrement des enfants, adolescents et adultes vulnérables. [11]

Le nouveau Centre pour des algorithmes et des données éthiques

19. Comme nous l'avons écrit dans notre rapport intermédiaire, les sociétés fournissant des réseaux sociaux tout comme celles fournissant des moteurs de recherche utilisent des algorithmes ou des séquences d'instructions pour personnaliser les informations et autres contenus aux utilisateurs. Ces algorithmes sélectionnent le contenu sur la base de facteurs tels que l'activité numérique passée de l'utilisateur, ses connexions sociales et leur localisation. Le modèle de revenus des compagnies d'Internet repose sur les revenus provenant de la vente d'espaces publicitaires et parce qu'il faut faire du profit, toute forme de contenu augmentant celui-ci sera priorisé. C'est pourquoi les histoires négatives seront toujours mises en avant par les algorithmes parce qu'elles sont plus fréquemment partagées que les histoires positives.[12]

20. Tout autant que les informations sur les compagnies de l'internet, les informations sur leurs algorithmes doivent être plus transparentes. Ils comportent intrinsèquement des travers, inhérents à la façon dont ils ont été développés par les ingénieurs ; ces travers sont ensuite reproduits diffusés et renforcés. Monica Bickert, de Facebook, a admis « que sa

compagnie était attentive à toute forme de déviance, sur le genre, la race ou autre qui pourrait affecter les produits de l'entreprise et que cela inclut les algorithmes ». Facebook devrait mettre plus d'ardeur à lutter contre ces défauts dans les algorithmes de ses ingénieurs pour éviter leur propagation.

[13]

21. Dans le budget de 2017, le Centre des données Ethiques et de l'innovation a été créé par le gouvernement pour conseiller sur « l'usage éthique, respectueux et innovant des données, incluant l'IA ». Le secrétaire d'état a décrit son rôle ainsi:

> Le Centre est un composant central de la charte numérique du gouvernement, qui définit des normes et des règles communes pour le monde numérique. Le centre permettra au Royaume-Uni de mener le débat concernant l'usage correct des données et de l'intelligence artificielle.[14]

22. Le centre agira comme un organisme de recommandation pour le gouvernement et parmi ses fonctions essentielles figurent : l'analyse et l'anticipation des manques en termes de régulation et de gestion; définition et orchestration des bonnes pratiques, codes de conduites et standards d'utilisations de l'Intelligence Artificielle; recommandation au gouvernement sur les règles et actions réglementaires à mettre en place en relation avec l'usage responsable et innovant des données. [15]

23. La réponse du gouvernement à notre rapport intermédiaire a mis en lumière certaines réponses à la consultation telle que la priorité de l'action immédiate du centre, telle que « le monopole sur la donnée, l'utilisation d'algorithme prédictifs dans la police, l'utilisation de l'analyse des données dans les campagnes politiques ainsi que l'éventualité de discrimination automatisée dans les décisions de recrutement ». Nous nous félicitons de la création du Centre

et nous nous réjouissons à la perspective d'en recueillir les fruits de ses prochaines travaux.

La loi en Allemagne et en France

24. D'autres pays ont légiféré contre le contenu malveillant sur les plateformes numériques. Comme nous l'avons relevé dans notre rapport intermédiaire, les compagnies d'internet en Allemagne ont été contraintes initialement de supprimer les propos haineux en moins de 24 heures. Quand cette auto-régulation s'est montrée inefficace, le gouvernement allemand a voté le Network Enforcement Act, aussi connu sous le nom de NetzDG, qui a été adopté en janvier 2018. Cette loi force les compagnies technologiques à retirer les propos haineux de leurs sites en moins de 24 heures et les condamne à une amende de 20 millions d'euros si ces contenus ne sont pas retirés[16]. Par conséquent, un modérateur sur six de Facebook travaille désormais en Allemagne, ce qui prouve bien que la loi peut être efficace.[17].

25. Une nouvelle loi en France, adoptée en novembre 2018 permet aux juges d'ordonner le retrait immédiat d'articles en ligne s'ils estiment qu'ils diffusent de la désinformation pendant les campagnes d'élection. La loi stipule que les utilisateurs doivent recevoir « d'informations qui sont justes, claires et transparentes » sur l'utilisation de leurs données personnelles, que les sites doivent divulguer les sommes qu'elles reçoivent pour promouvoir des informations, et la loi autorise le CSA français à pouvoir suspendre des chaînes de télévision contrôlées ou sous influence d'un état étranger, s'il estime que cette chaîne dissémine de manière délibérée des fausses informations qui pourraient affecter l'authenticité du vote. Les sanctions imposées en violation de la loi comprennent un an de prison et une amende de 75000 euros[18].

Le Royaume-Uni

26. Comme la Commissaire de l'Information du Royaume-Uni, Elisabeth Denham, nous l'a expliqué en novembre 2018, il y a une tension entre le modèle économique des médias sociaux, centré sur la publicité, et les droits humains tels que la protection de la vie privée. « C'est notre situation actuelle et il s'agit d'une tâche importante à la fois pour les régulateurs et le législateur de s'assurer que les bonnes exigences, la surveillance et sanctions sont en place » [19]. Elle nous a dit que Facebook, par exemple, devrait en faire plus et devrait faire « l'objet d'une régulation et d'une surveillance plus stricte »[20]. Les activités de Facebook dans la scène politique sont en augmentation; l'entreprise a récemment lancé un fil d'actualités intitulé « Community Actions » avec une fonctionnalité de pétition pour, par exemple, permettre aux utilisateurs de résoudre des problèmes politiques locaux en créant ou soutenant des pétitions. Il est difficile de comprendre comment Facebook sera capable d'auto-réguler une telle fonctionnalité; plus le problème local va être sujet à controverse et litiges, plus il entrainera de l'engagement sur Facebook et donc de revenus associés grâce aux publicités [21].

Facebook et la loi

27. En dépit de toutes les excuses formulées par Facebook pour ses erreurs passées, il semble encore réticent à être correctement surveillé. Lors de la session de témoignage verbal au « Grand Comité International », Richard Alland, vice-président des solutions politiques de Facebook, a été interrogé à plusieurs reprises sur les opinions de Facebook sur la régulation, et à chaque fois il a déclaré que Facebook était très ouvert au débat sur la régulation, et que travailler ensemble avec les gouvernements seraient la meilleure option possible :

> « Je suis ravi, personnellement, et l'entreprise est vraiment engagé, de la base jusqu'à notre PDG – il en a parlé en public – à l'idée d'obtenir le bon type de régulation afin que l'on puisse arrêter d'être dans ce mode de confrontation. Cela ne sert ni notre société ni nos utilisateurs. Essayons de trouver le juste milieu, où vous êtes d'accord pour dire que nous faisons un travail suffisamment bon et où vous avez le pouvoir de nous tenir responsable si nous ne le faisons pas, et nous comprenons quel le travail que nous avons à faire. C'est la partie régulation[22]. »

28. Ashkan Soltani, un chercheur et consultant indépendant, et ancien Responsable Technologique de la Commission Fédérale du Commerce des USA [23], a questionné la volonté de Facebook à être régulé. À propos de la culture interne de Facebook, il a dit : « Il y a ce mépris – cette capacité à penser que l'entreprise sait mieux que tout le monde et tous les législateurs » [24]. Il a discuté de la loi californienne pour la vie privée des consommateurs [25] que Facebook a supporté en public, mais a combattu en coulisses [26].

29. Facebook ne semble pas vouloir être régulé ou surveillé. C'est considéré comme normal pour les ressortissants étrangers de témoigner devant les comités. En effet, en juin 2011, le Comité pour la Culture, les Médias et le Sport [27] a entendu le témoignage de Rupert Murdoch lors de l'enquête sur le hacking téléphonique [28] et le Comité au Trésor [29] a récemment entendu le témoignage de trois ressortissants étrangers [30]. **En choisissant de ne pas se présenter devant le Comité et en choisissant de ne pas répondre personnellement à aucune de nos invitations, Mark Zuckerberg a fait preuve de mépris envers à la fois le parlement du Royaume-Uni et le « Grand Comité International », qui compte des représentants de neufs législatures dans le monde.**

30. La structure managériale de Facebook est opaque pour les personnes extérieures, et semble conçue pour dissimuler la connaissance et la responsabilité de certaines décisions. Facebook a pour stratégie d'envoyer des témoins dont ils disent qu'ils sont les plus adéquats, mais qui n'ont pas été suffisamment informés sur les points cruciaux, et ne peuvent répondre ou choisissent de ne pas répondre à nombre de nos questions. Ils promettent ensuite d'y répondre par lettre, qui —sans surprise— échouent à répondre à toutes nos questions. Il ne fait pas de doute que cette stratégie est délibérée.

Régulateurs britanniques existants

31. Au Royaume-Uni, les principales autorités compétentes — Ofcom, l'autorité pour les standards publicitaires [31], le bureau du commissaire à l'information [32], la commission électorale [33] et l'autorité pour la compétition et le marché [34] — ont des responsabilités spécifiques sur l'utilisation de contenus, données et comportements. Quand Sharon White, responsable de Ofcom, est passé devant le comité en octobre 2018, après la publication de notre rapport intermédiaire, nous lui avons posé la question si leur expérience comme régulateur de diffusion audiovisuelle pourrait être utile pour réguler les contenus en ligne. Elle a répondu :

> « On a essayé d'identifier quelles synergies seraient possibles. […] On a été frappé de voir qu'il y a deux ou trois domaines qui pourraient être applicable en ligne. […]
>
> Le fait que le Parlement [35] ait mis en place des standards, ainsi que des objectifs plutôt ambitieux, nous a semblé très important, mais aussi durable avec des objectifs clés, que ce soit la protection de l'enfance ou les préoccupations autour des agressions et injures. Vous pouvez le voir comme un processus démocratique sur quels sont les maux que l'on croit en tant que société être fréquent en ligne. L'autre chose qui est très importante

dans le code de diffusion audiovisuelle est qu'il explicite clairement le fait que ces choses peuvent varier au cours du temps comme la notion d'agression se modifie et les inquiétudes des consommateurs changent. La mise en œuvre est ensuite déléguée à un régulateur indépendant qui traduit en pratique ces objectifs de standards. Il y a aussi la transparence, le fait que l'on publie nos décisions dans le cas d'infractions, et que tout soit accessible au public. Il y a la surveillance de nos décisions et l'indépendance du jugement [36] ».

32. Elle a également ajouté que la fonction du régulateur de contenu en ligne devrait évaluer l'efficacité des compagnies technologiques sur leurs mesures prises contre les contenus qui ont été signalés comme abusifs. « Une approche serait de se dire si les compagnies ont les systèmes, les processus, et la gouvernance en place avec la transparence qui amène la responsabilité publique et la responsabilité devant le Parlement, que le pays serait satisfait du devoir de vigilance ou que les abus seront traités de manière constante et efficace ».[37].

33. Cependant, si on demandait à Ofcom de prendre en charge la régulation des capacités des compagnies des réseaux sociaux, il faudrait qu'il soit doté de nouveaux pouvoirs d'enquête. Sharon White a déclaré au comité « qu'il serait absolument fondamental d'avoir des informations statutaires, réunissant des pouvoirs sur un domaine large ».[38].

34. UK Council for Internet Safety(UKCIS) est un nouvel organisme, sponsorisé par le Ministère du Numérique, de la Culture, des Médias et du Sport, le Ministère de l'Éducation et le Ministère de l'Intérieur, il réunit plus de 200 organisations qui ont pour but de garantir la sécurité des enfants en ligne. Son site web affirme « si c'est inacceptable hors ligne, c'est inacceptable en ligne ». Son attention tiendra compte des abus en lignes comme le cyberharcèlement et

l’exploitation sexuelle, la radicalisation et l’extrémisme, la violence contre les femmes et les jeunes filles, les crimes motivés par la haine et les discours haineux, et les formes de discrimination vis à vis de groupes protégés par l’Equality Act.[39]. Guy Parker, Pdg d’Advertising Standards Authority nous a informé que le Gouvernement pourrait se décider à intégrer les abus dans la publicité dans leur définition d’abus en ligne[40].

35. Nous pensons que UK Council for Internet Safety devrait inclure dans le périmètre de ses attributions « le risque envers la démocratie » tel qu’identifié dans le « Registre des Préjudices » du Center for Human Technology, en particulier par rapport aux reportages profondément faux. Nous notons que Facebook est inclus en tant que membre d’UKCIS, compte tenu de son influence éventuelle, et nous comprenons pourquoi. Cependant, étant donné l’attitude de Facebook dans cette enquête, nous avons des réserves quant à sa bonne foi des affaires et sa capacité à participer au travail d’UKCIS dans l’intérêt du public, par opposition à ses intérêts personnels.

36. Lorsqu’il a été demandé au Secrétaire du Numérique, de la Culture, des Médias et des Sports, le Très Honorable député Jeremy Wright, de formuler un spectre des abus en ligne, sa réponse était limitée. « Ce que nous devons comprendre est à quel point les gens sont induits en erreur ou à quel point les élections ont été entravées de manière délibérée ou influencée, et si elle le sont […] nous devons trouver des réponses appropriées et des moyens de défense. Cela fait partie d’un paysage bien plus global et je ne crois par que c’est juste de le segmenter[41]. Cependant, une fois que nous avons défini les difficultés autour de la définition, l’étendue et la responsabilité des abus en ligne, le Secrétaire d’État était plus coopératif lorsqu’on lui a posé la question sur la régulation des compagnies de réseaux sociaux, et a déclaré que le Royaume-Uni devrait prendre

l’initia

37. **Notre rapport intermédiaire recommandait que des responsabilités juridiques claires soient définies pour les compagnies technologiques, afin qu’elles puissent prendre des mesures allant contre des contenus abusifs ou illégaux sur leurs sites. À l’heure actuelle, il est urgent de mettre en œuvre des règlements indépendants. Nous croyons qu’un Code d’Éthique obligatoire doit être implémenté, supervisé par un régulateur indépendant, définissant ce que constitue un contenu abusif. Le régulateur indépendant aurait des pouvoirs conférés par la loi pour surveiller les différentes compagnies technologiques, cela pourrait créer un système réglementaire pour les contenus en ligne qui est aussi effectif que pour ceux des industries de contenu hors ligne.**

38. **Comme nous l’avons énoncé dans notre rapport intermédiaire, un tel Code d’Éthique devrait ressembler à celui du Broadcasting Code publiée par Ofcom, qui se base sur des lignes directrices définies dans la section 319 du Communications Acts de 2003. Le Code d’Éthique devrait être mis au point par des experts techniques et supervisés par un régulateur indépendant, pour pouvoir mettre noir sur blanc ce qui est acceptable et ce qui ne l’est pas sur les réseaux sociaux, notamment les contenus abusifs et illégaux qui ont été signalés par leurs utilisateurs pour être retirés, ou qu’il aurait été facile d’identifier pour les compagnies technologiques elles-mêmes.**

39. **Le processus devrait définir une responsabilité juridique claire pour les compagnies technologiques de prendre des mesures contre les contenus abusifs et illégaux sur leur plateforme et ces compagnies devraient mettre en place des systèmes adaptés pour marquer et retirer des « types d’abus » et s’assurer que les structures de cybersécurité soient implémentées. Si les compagnies techniques (y compris les ingénieurs informaticiens en charge de la création des logiciels pour ces compagnies) sont reconnues fautifs de ne**

pas avoir respecté leurs obligations en vertu d'un tel code, et n'ont pas pris de mesure allant contre la diffusion de contenus abusifs et illégaux, le régulateur indépendant devrait pouvoir engager des poursuites judiciaires à leur encontre, dans l'objectif de les condamner à payer des amendes élevées en cas de non-respect du Code.

40. C'est le même organisme public qui devrait avoir des droits statutaires pour obtenir toute information de la part des compagnies de réseaux sociaux qui sont en lien avec son enquête. Cela pourrait concerner la capacité de vérifier les données qui sont conservées sur un utilisateur, s'il demandait ces informations. Cet organisme devrait avoir accès aux mécanismes de sécurité des compagnies technologiques et aux algorithmes, pour s'assurer qu'ils travaillent de manière responsable. Cet organisme public devrait être accessible au public et recevoir les plaintes sur les compagnies des réseaux sociaux. Nous demandons au gouvernement de soumettre ces propositions dans son prochain livre blanc.

Utilisation des données personnelles et inférence

41. Lorsque Mark Zuckerberg a fourni les preuves au congrès en avril 2018, dans la suite du scandal Cambridge Analytica, il a fait la déclaration suivante : « Vous devriez avoir un contrôle complet sur vos données […] Si nous ne communiquons pas cela clairement, c'est un point important sur lequel nous devons travailler ». Lorsqu'il lui a été demandé à qui était cet « alterego virtuel », Zuckerberg a répondu que les gens eux-mêmes possèdent tout le « contenu » qu'ils hébergent sur la plateforme, et qu'ils peuvent l'effacer à leur gré[42]. Cependant, le profil publicitaire que Facebook construit sur les utilisateurs ne peut être accédé, contrôlé ni effacé par ces utilisateurs. Il est difficile de concilier ce fait avec l'affirmation que les utilisateurs possèdent tout « le contenu » qu'ils uploadent.

42. Au Royaume-Uni, la protection des données utilisateur est couverte par le RGPD (Règlement Général de Protection des Données)[43]. Cependant, les données « inférées » ne sont pas protégées ; cela inclut les caractéristiques qui peuvent être inférées sur les utilisateurs et qui ne sont pas basées sur des informations qu'ils ont partagées, mais sur l'analyse des données de leur profil. Ceci, par exemple, permet aux partis politiques d'identifier des sympathisants sur des sites comme Facebook, grâce aux profils correspondants et aux outils de ciblage publicitaire sur les « publics similaires ». Selon la propre description de Facebook des « publics similaires », les publicitaires ont l'avantage d'atteindre de nouvelles personnes sur Facebook « qui ont des chances d'être intéressées par leurs produits car ils sont semblables à leurs clients existants » [44].

43. Le rapport de l'ICO, publié en juillet 2018, interroge sur la présomption des partis politques à ne pas considérer les données inférées comme des données personnelles:

> « Nos investigations montrent que les partis politiques n'ont pas considéré les données inférées comme des informations personnelles car ce ne sont pas des informations factuelles. Cependant, le point de vue de l'ICO est que ces informations sont basées sur des hypothèses sur les intérets des personnes et leurs préférences, et peuvent être attribuées à des individus spécifiques, donc ce sont des informations personnelles et elles sont soumises aux contraintes de la protection des données [45]. »

44. Les données inférées sont donc considérées par l'ICO comme des données personnelles, ce qui devient un problème lorsque les utilisateurs sont informés qu'ils disposent de leurs propres données, et qu'ils ont un pouvoir sur où les données vont, et ce pour quoi elles sont utilisées. Protéger nos données nous aide à sécuriser le passé, mais protéger les

inférences et l'utilisation de l'Intelligence Artificielle (IA) est ce dont nous avons besoin pour protéger notre futur.

45. La commissaire à l'information, Elizabeth Denham, a souligné son intérêt sur l'utilisation des données inférées dans les campagnes politiques lorsqu'elle a fourni des preuves au comité en novembre 2018, déclarant qu'il y a eu :

> « Un nombre dérangeant de manque de respect des données personnelles des votants et des votants potentiels. Ce qui s'est passé ici est que le modèle familier aux gens du secteur commercial sur le ciblage des comportements a été transféré – je pense transformé – dans l'arène politique. C'est pour cela que j'appelle à une pause éthique, afin que nous puissions y remédier. Nous ne voulons pas utiliser le même modèle qui nous vend des vacances, des chaussures et des voitures pour collaborer avec des personnes et des votants. Les gens veulent plus que ça. C'est le moment pour faire un pause pour regarder les codes, regarder les pratiques des entreprises de réseaux sociaux, de prendre des mesures là où ils ont enfreint la loi. Pour nous, le principal but de ceci est de lever le rideau et montrer au public ce qu'il advient de leurs données personnelles [46]. »

46. Avec des références explicites sur l'utilisation des « publics similaires » de Facebook, Elizabeth Denham a expliqué au comité qu'ils « devaient être transparents envers les [utilisateurs] privés. Ils ont besoin de savoir qu'un parti politique, ou un membre du parlement, fait usage des publics similaires. Le manque de transparence est problématique[47]. Lorsque nous avons demandé à la commissaire à l'information si elle pensait que l'utilisation des « publics similaires » était légal selon le RGPD, elle a répondu : « Nous avons besoin de l'analyser en détail sous la loupe du RGPD, mais je pense que le public est mal à l'aise avec les publics similaires, et il a besoin que ce soit transparent » [48]. Les gens ont besoin de savoir que l'information qu'ils

donnent pour un besoin spécifique va être utilisé pour inférer des informations sur eux dans d'autres buts.

47. Le secrétaire d'état, le très honorable membre du parlement Jeremy Wright, nous a également informé que le framework éthique et législatif entourant l'IA devait se développer parallèlement à la technologie, ne pas « courir pour [la] rattraper », comme cela s'est produit avec d'autres technologies dans le passé [49]. Nous devons explorer les problèmes entourant l'IA en détail, dans nos enquêtes sur les technologies immersives et d'addictives, qui a été lancée en décembre 2018 [50].

48. **Nous soutenons la recommandation de l'ICO comme quoi les données inférées devraient être protégées par la loi comme les informations personnelles. Les lois sur la protection de la vie privée devraient être étendues au-delà des informations personnelles pour inclure les modèles utilisés pour les inférences sur les individus. Nous recommandons que le gouvernement étudie les manières dont les protections de la vie privée peuvent être étendues pour inclure les modèles qui sont utilisés pour les inférences sur les individus, en particulier lors des campagnes politiques. Cela nous assurerait que les inférences sur les individus sont traitées de manière aussi importante que les informations personnelles des individus.**

Rôle accru de l'OIC et taxe sur les entreprises de technologie

49. Dans notre rapport intérimaire, nous avons demandé que l'OIC soit mieux à même d'être à la fois un « shérif efficace dans le Far West de l'Internet » et d'anticiper les technologies futures. L'OIC doit avoir les mêmes connaissances techniques, sinon plus, que les organisations examinées[51]. Nous avons recommandé qu'une redevance soit prélevée sur les

sociétés de technologie opérant au Royaume-Uni, pour aider à payer ces travaux, dans le même esprit que la façon dont le secteur bancaire paie les frais de fonctionnement de l'autorité de régulation Financière [52] [53].

50. Lorsque l'on a demandé au secrétaire d'État ce qu'il pensait d'une redevance, il a répondu, en ce qui concerne Facebook en particulier: « Le Comité est rassuré que ce n'est pas parce que Facebook dit qu'il ne veut pas payer une redevance, qu'il ne sera pas question de savoir si nous devrions ou non avoir une redevance »[54]. Il nous a également dit que « ni moi, ni, je pense franchement, l'OIC, ne pensons qu'elle soit sous-financée pour le travail qu'elle a à faire actuellement. […] Si nous devons mener d'autres activités, que ce soit en raison d'une réglementation ou d'une formation supplémentaires, par exemple, il faudra bien qu'elles soient financées d'une façon ou d'une autre. Par conséquent, je pense que la redevance vaut la peine d'être envisagée »[55].

51. **Dans notre rapport intermédiaire, nous avons recommandé qu'une redevance soit prélevée sur les sociétés de technologie opérant au Royaume-Uni pour soutenir le travail renforcé de l'OIC. Nous réitérons cette recommandation. La décision du chancelier, dans son budget de 2018, d'imposer une nouvelle taxe de 2% sur les services numériques sur les revenus des grandes entreprises technologiques du Royaume-Uni à partir d'avril 2020, montre que le gouvernement est ouvert à l'idée d'une taxe sur les entreprises technologiques. Dans sa réponse à notre rapport intermédiaire, le gouvernement a laissé entendre qu'il n'appuierait plus financièrement l'OIC, contrairement à notre recommandation. Nous exhortons le gouvernement à réévaluer cette position.**

52. Le nouveau système indépendant et la nouvelle réglementation que nous recommandons d'établir doivent être financés adéquatement. Nous recommandons qu'une taxe soit prélevée sur les sociétés de technologie opérant au Royaume-

Uni pour financer leur travail.

La nouvelle dystopie, c’est maintenant

L’article qui suit n’est pas une traduction intégrale mais un survol aussi fidèle que possible de la conférence TED effectuée par la sociologue des technologies Zeynep Tufecki. Cette conférence intitulée : « Nous créons une dystopie simplement pour obliger les gens à cliquer sur des publicités »
(We’re building a dystopia just to make people click on ads) ~~*est en cours de traduction sur la plateforme Amara préconisée par TED, mais la révision n’étant pas effectuée, il faudra patienter pour en découvrir l’intégralité sous-titrée en français.*~~ *est maintenant traduite en français \o/*

En attendant, voici 4 minutes de lecture qui s’achèvent hélas sur des perspectives assez vagues ou plutôt un peu vastes : il faut tout changer. Du côté de Framasoft, nous proposons de commencer par outiller la société de contribution avec la campagne Contributopia… car dégoogliser ne suffira pas !

Mettez un peu à jour vos contre-modèles, demande Zeynep : oubliez les références aux menaces de Terminator et du 1984 d’Orwell, ces dystopies ne sont pas adaptées à notre débutant XXI^{e} siècle.

Cliquez sur l'image pour afficher la vidéo sur le site de TED (vous pourrez afficher les sous-titres via un bouton en bas de la vidéo)

Ce qui est à craindre aujourd'hui, car c'est *déjà là*, c'est plutôt comment ceux qui détiennent le pouvoir utilisent et vont utiliser l'intelligence artificielle pour exercer sur nous des formes de contrôle nouvelles et malheureusement peu détectables. Les technologies qui menacent notre liberté et notre jardin secret (celui de notre bulle d'intimité absolue) sont développées par des entreprises-léviathans qui le font d'abord pour vendre nos données et notre attention aux GAFAM (Tristan Nitot, dans sa veille attentive, signale qu'on les appelle les *frightful five,* les 5 qui font peur, aux États-Unis). Zeynep ajoute d'ailleurs Alibaba et Tencent. D'autres à venir sont sur les rangs, peut-on facilement concevoir.

Ne pas se figurer que c'est seulement l'étape suivante qui prolonge la publicité en ligne, c'est au contraire un véritable saut vers une autre catégorie « un monde différent » à la fois exaltant par son potentiel extraordinaire mais aussi terriblement dangereux.

Voyons un peu la mécanique de la publicité. Dans le monde physique, les friandises à portée des enfants au passage en

caisse de supermarché sont un procédé d'incitation efficace, mais dont la portée est limitée. Dans le monde numérique, ce que Zeynep appelle **l'architecture de la persuasion** est à l'échelle de plusieurs milliards de consommateurs potentiels. Qui plus est, l'intelligence artificielle peut cibler chacun distinctement et envoyer sur l'écran de son smartphone (on devrait dire *spyphone*, non ?) un message incitatif qui ne sera vu que par chacun et le ciblera selon ses points faibles identifiés par algorithmes.

Prenons un exemple : quand hier l'on voulait vendre des billets d'avion pour Las Vegas, on cherchait la tranche d'âge idéale et la carte de crédit bien garnie. Aujourd'hui, les mégadonnées et l'apprentissage machine (*machine learning*) s'appuient sur tout ce que Facebook peut avoir collecté sur vous à travers messages, photos, « *likes* », même sur les textes qu'on a commencés à saisir au clavier et qu'on a ensuite effacés, etc. Tout est analysé en permanence, complété avec ce que fournissent des courtiers en données.

Les algos d'apprentissage, comme leur nom l'indique, apprennent ainsi non seulement votre profil personnel mais également, face à un nouveau compte, à quel type déjà existant on peut le rapprocher. Pour reprendre l'exemple, ils peuvent deviner très vite si telle ou telle personne est susceptible d'acheter un billet pour un séjour à Las Vegas.

Vous pensez que ce n'est pas très grave si on nous propose un billet pour Vegas.

Le problème n'est pas là.
Le problème c'est que les algorithmes complexes à l'œuvre deviennent opaques pour tout le monde, y compris les programmeurs, même s'ils ont accès aux données qui sont généralement propriétaires donc inaccessibles.

> *« Comme si nous cessions de programmer pour laisser se développer une forme d'intelligence que nous ne comprenons*

pas véritablement. Et tout cela marche seulement s'il existe une énorme quantité de données, donc ils encouragent une surveillance étendue : pour que les algos de machine learning puissent opérer. Voilà pourquoi Facebook veut absolument collecter le plus de données possible sur vous. Les algos fonctionneront bien mieux »

Que se passerait-il, continue Zeynep avec l'exemple de Las Vegas, si les algos pouvaient repérer les gens bipolaires, soumis à des phases de dépenses compulsives et donc bons clients pour Vegas, capitale du jeu d'argent ? Eh bien un chercheur qui a contacté Zeynep a démontré que les algos pouvaient détecter les profils à risques psychologiques avec les médias sociaux avant que des symptômes cliniques ne se manifestent…
Les outils de détection existent et sont accessibles, les entreprises s'en servent et les développent.

L'exemple de YouTube est également très intéressant : nous savons bien, continue Zeynep, que nous sommes incités par un algo à écouter/regarder d'autres vidéos sur la page où se trouve celle que nous avons choisie.

Eh bien en 2016, témoigne Zeynep, j'ai reçu de suggestions par YouTube : comme j'étudiais la campagne électorale en sociologue, je regardais des vidéos des meetings de Trump et YouTube m'a suggéré des vidéos de suprématistes (extrême-droite fascisante aux USA) !
Ce n'est pas seulement un problème de politique. L'algorithme construit une idée du comportement humain, en supposant que nous allons pousser toujours notre curiosité vers davantage d'extrêmes, de manière à nous faire demeurer plus longtemps sur un site pendant que Google vous sert davantage de publicités.
Pire encore, comme l'ont prouvé des expériences faites par ProPublica et BuzzFeed, que ce soit sur Facebook ou avec Google, avec un investissement minime, on peut présenter des

messages et profils violemment antisémites à des personnes qui ne sont pas mais *pourraient* (toujours suivant les algorithmes) devenir antisémites.

L'année dernière, le responsable médias de l'équipe de Trump a révélé qu'ils avaient utilisé de messages « non-publics » de Facebook pour démobiliser les électeurs, les inciter à ne pas voter, en particulier dans des villes à forte population d'Afro-américains. Qu'y avait-il dans ces messages « non-publics » ? On ne le saura pas, Twitter ne le dira pas.

Les algorithmes peuvent donc aussi influencer le comportement des électeurs.

Facebook a fait une expérience en 2010 qui a été divulguée après coup.
Certains ont vu ce message les incitant à voter. Voici la version basique :

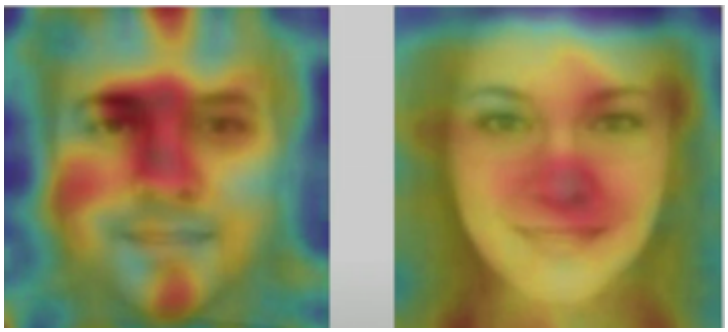
et d'autres ont vu cette version (avec les imagettes des contacts qui ont cliqué sur « j'ai voté »)

Ce message n'a été présenté qu'une fois mais **340 000 électeurs**

de plus ont voté lors de cette élection, selon cette recherche, confirmée par les listes électorales.
En 2012, même expérience, résultats comparables : 270 000 électeurs de plus.
De quoi laisser songeur quand on se souvient que l'élection présidentielle américaine de 2016 s'est décidée à environ 100 000 voix près…

> *« Si une plate-forme dotée d'un tel pouvoir décide de faire passer les partisans d'un candidat avant les autres, comment le saurions-nous ? »*

Les algorithmes peuvent facilement déduire notre appartenance à une communauté ethnique, nos opinions religieuses et politiques, nos traits de personnalité, l'intelligence, la consommation de substances addictives, la séparation parentale, l'âge et le sexe, en se fondant sur les « j'aime » de Facebook. Ces algorithmes peuvent identifier les manifestants même si leurs visages sont partiellement dissimulés, et même l'orientation sexuelle des gens à partir de leurs photos de leur profil de rencontres.
Faut-il rappeler que la Chine utilise déjà la technologie de détection des visages pour identifier et arrêter les personnes ?



Le pire, souligne Zeynep est que

> *« Nous construisons cette infrastructure de surveillance autoritaire uniquement pour inciter les gens à cliquer sur les publicités. »*

Si nous étions dans l'univers terrifiant de 1984 nous aurions peur mais nous saurions de quoi, nous détesterions et pourrions résister. Mais dans ce nouveau monde, si un état nous observe et nous juge, empêche par anticipation les potentiels fauteurs de trouble de s'opposer, manipule individus et masses avec la même facilité, nous n'en saurons rien ou très peu…

> *« Les mêmes algorithmes que ceux qui nous ont été lancés pour nous rendre plus flexibles en matière de publicité organisent également nos flux d'informations politiques, personnelles et sociales… »*

Les dirigeants de Facebook ou Google multiplient les déclarations bien intentionnées pour nous convaincre qu'ils ne nous veulent aucun mal. Mais le problème c'est le *business model* qu'ils élaborent. Ils se défendent en prétendant que leur pouvoir d'influence est limité, mais de deux choses l'une : ou bien Facebook est un énorme escroquerie et les publicités ne fonctionnent pas sur leur site (et dans ce cas pourquoi des entreprises paieraient-elles pour leur publicité sur Facebook ?), ou bien leur pouvoir d'influence est terriblement préoccupant. C'est soit l'un, soit l'autre. Même chose pour Google évidemment.

Que faire ?
C'est toute la structure et le fonctionnement de notre technologie numérique qu'il faudrait modifier…

> *« Nous devons faire face au manque de transparence créé par les algorithmes propriétaires, au défi structurel de l'opacité de l'apprentissage machine, à toutes ces données qui sont recueillies à notre sujet. Nous avons une lourde tâche devant nous. Nous devons mobiliser notre technologie, notre créativité et aussi notre pouvoir politique pour construire une intelligence artificielle qui nous soutienne dans nos objectifs humains, mais qui soit aussi limitée par*

nos valeurs humaines. »

« Nous avons besoin d'une économie numérique où nos données et notre attention ne sont pas destinées à la vente aux plus offrants autoritaires ou démagogues. »

- voir la vidéo : *We're building a dystopia just to make people click on ads*
- une autre conférence de Zeynep avec des sous-titres en français : l'intelligence artificielle rend la morale plus importante.

Quand les recommandations YouTube nous font tourner en bourrique…

Vous avez déjà perdu une soirée à errer de vidéo en vidéo suivante ? À cliquer *play* en se disant « OK c'est la dernière… » puis relever les yeux de votre écran 3 heures plus tard… ?

C'est grâce à (ou la faute de, au choix !) l'algorithme des recommandations, une petite recette qui prend plein d'éléments en compte pour vous signaler les vidéos qui peuvent vous intéresser.

Guillaume Chaslot a travaillé sur cet algorithme. Il a même créé un petit outil open-source pour le tester, afin de

valider sa théorie : ces recommandations nous pousseraient de plus en plus vers les « faits alternatifs » (ça s'appelle aussi une légende urbaine, un complot, une fiction, du *bullshit*… vous voyez l'idée.)

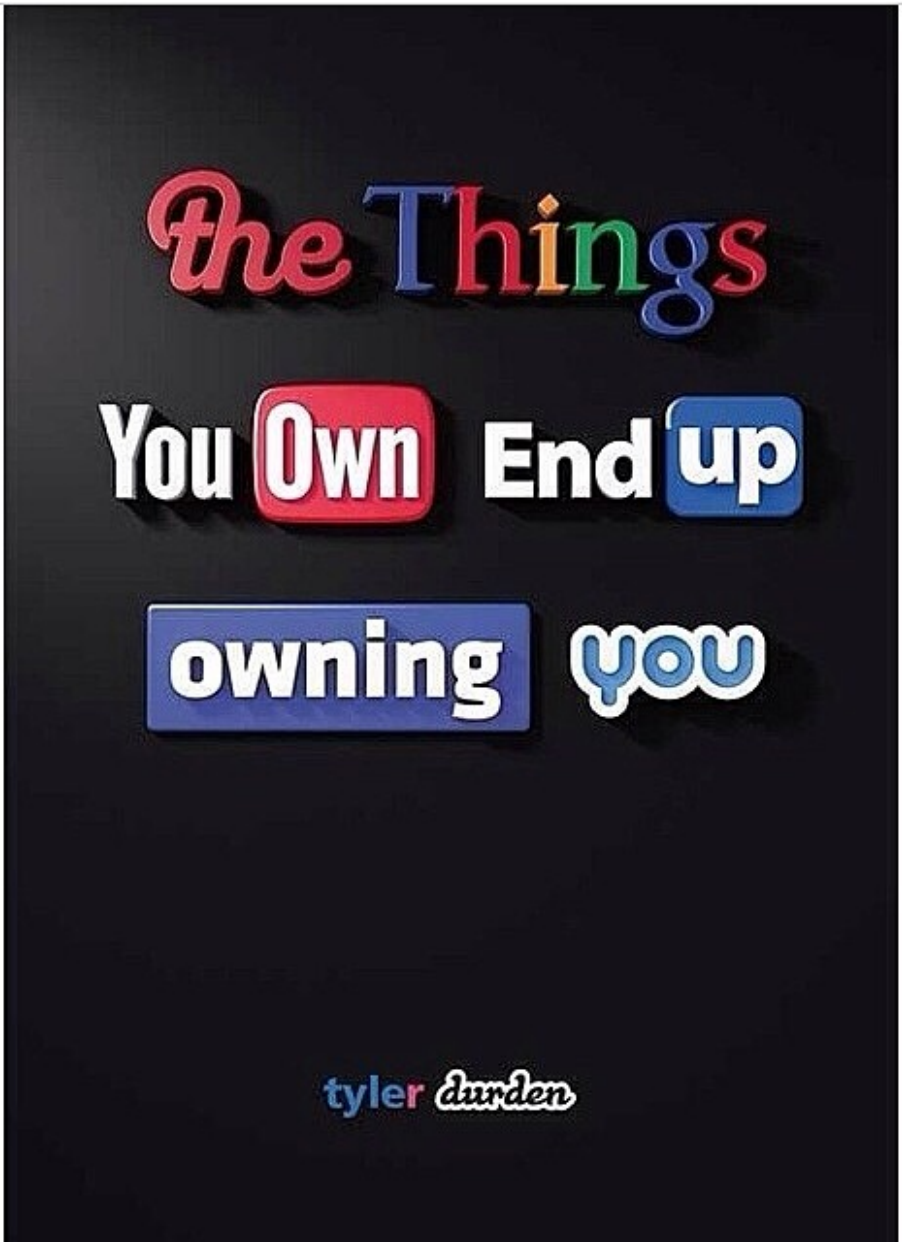
Le groupe Framalang a décidé de traduire cet article passionnant.

Ne soyons pas complotistes à notre tour. Cet article ne dit pas que Google veut nous remplir la tête de mensonges et autres légendes numériques. Il s'agirait là, plutôt, d'un effet de bord de son algorithme.

Nous ne doutons pas, en revanche, qu'un des buts premiers de Google avec ses recommandations YouTube est de captiver notre attention, afin de vendre à ses clients notre temps de cerveau disponible (et d'analyser nos comportements au passage pour remplir ses banques de données avec nos vies numériques).

Sauf qu'avec ce genre de vision (et de buts) à court/moyen terme, on ne réfléchit pas aux conséquences sur le long terme. Lorsque l'on représente l'endroit où une grande portion de notre civilisation passe la majeure partie de son temps… C'est problématique, non ?

Tout comme les révélations de Tristan Harris, ce témoignage nous rappelle que, même chez les géants du web, notre monde numérique est tout jeune, immature, et qu'il est grand temps de prendre du recul sur les constructions que nous y avons dressées : car chacun de ces systèmes implique ses propres conséquences.



“The things you own end up owning you » de *koka_sexton* sous licence CC BY 2.0
(« Ce que vous possédez finit par vous posséder », une citation de Fight Club)

Comment l’I.A. de YouTube favorise les « faits alternatifs »

de Guillaume Chaslot, source : Medium.

Traduction : Jerochat, jaaf, dominix, mo, goofy, Asta, Opsylac, Nimanneau, audionuma, Lyn. + les anonymes

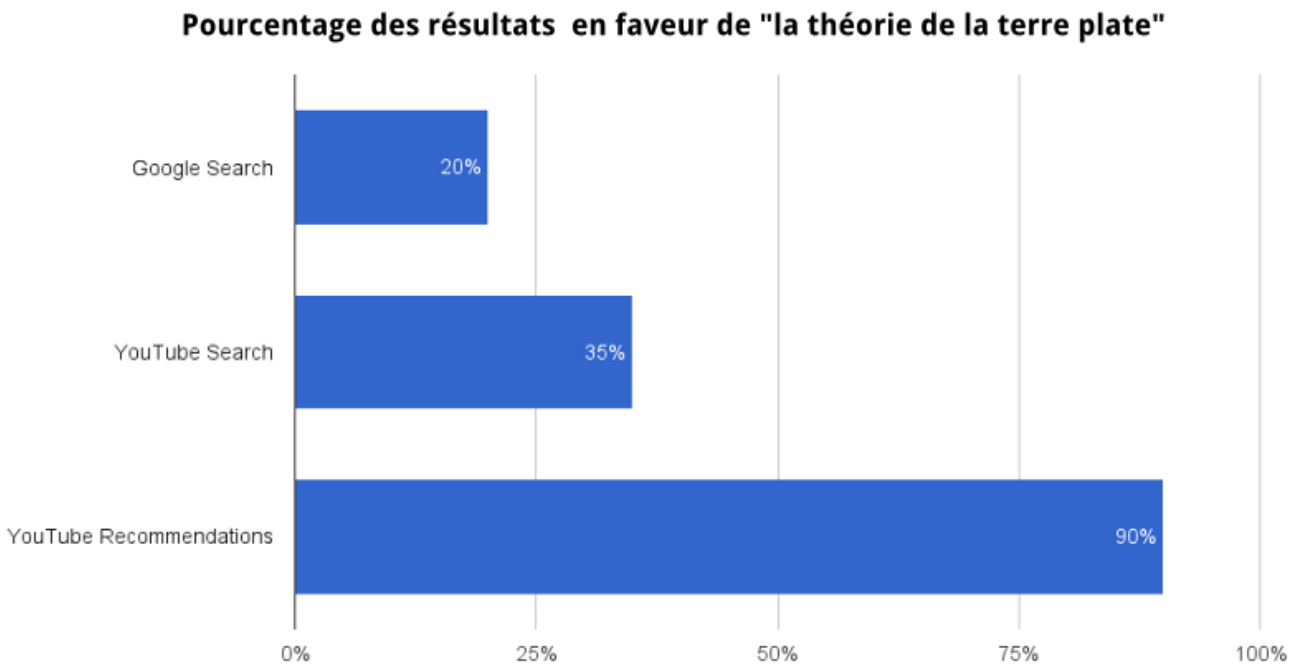
Les I.A. sont conçues pour maximiser le temps que les utilisateurs passent en ligne… Et pour ce faire, la fiction, souvent, dépasse la réalité.

Tout le monde a déjà entendu parler des théories du complot, des faits alternatifs ou des *fake news* qui circulent sur Internet. Comment sont-ils devenus si répandus ? Quel est l'impact des algorithmes de pointe sur leur succès ?

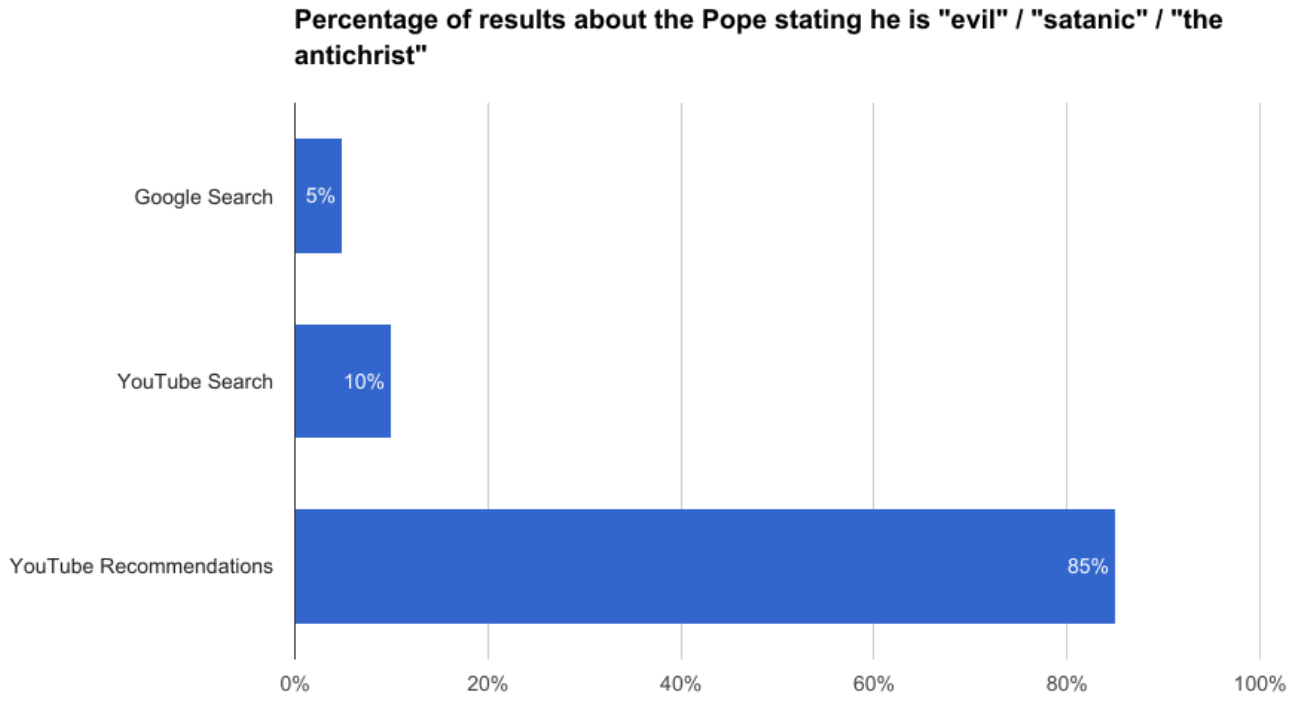
Ayant moi-même travaillé sur l'algorithme de recommandation de YouTube, j'ai commencé à enquêter, et je suis arrivé à la conclusion que le puissant algorithme que j'avais contribué à concevoir joue un rôle important dans la propagation de fausses informations.

Pour voir ce que YouTube promeut actuellement le plus, j'ai développé un explorateur de recommandations *open source* qui extrait les vidéos les plus recommandées sur une requête donnée. Je les ai comparées aux 20 premiers résultats venant de requêtes identiques sur Google et Youtube Search.
Les résultats sur les 5 requêtes suivantes parlent d'eux-mêmes :

1 – Question élémentaire : « La Terre est-elle plate ou ronde ? »

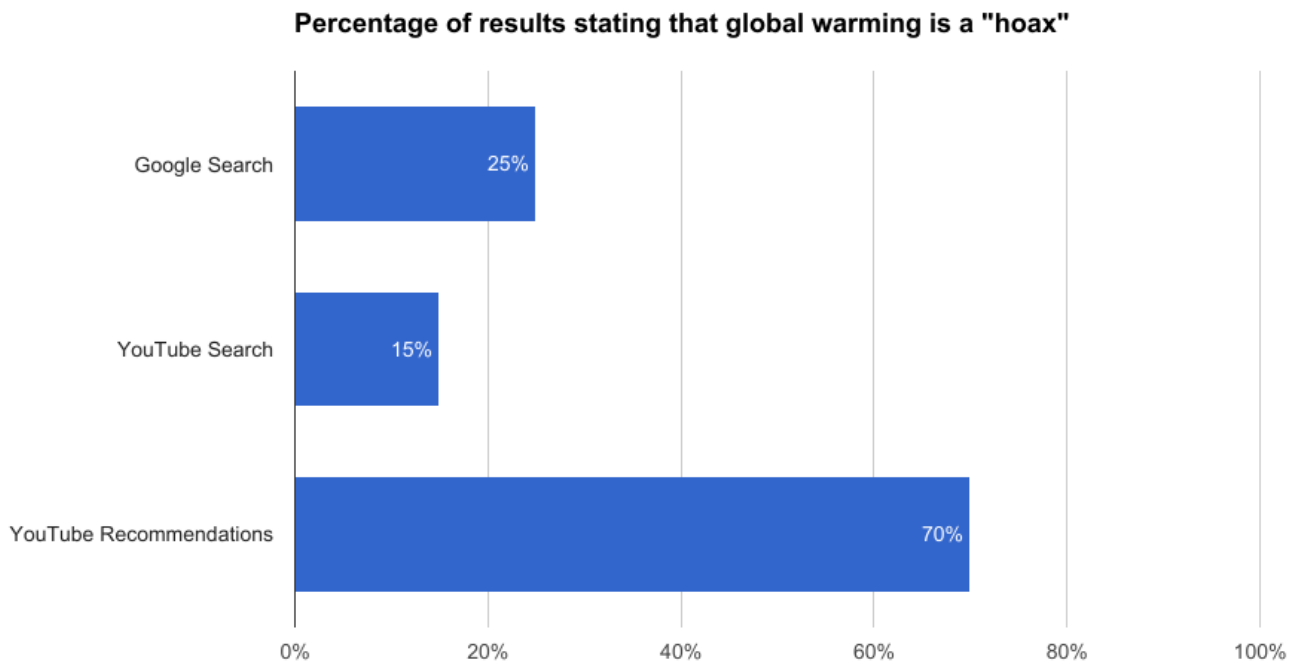


2 – Religion : « Qui est le Pape ? »



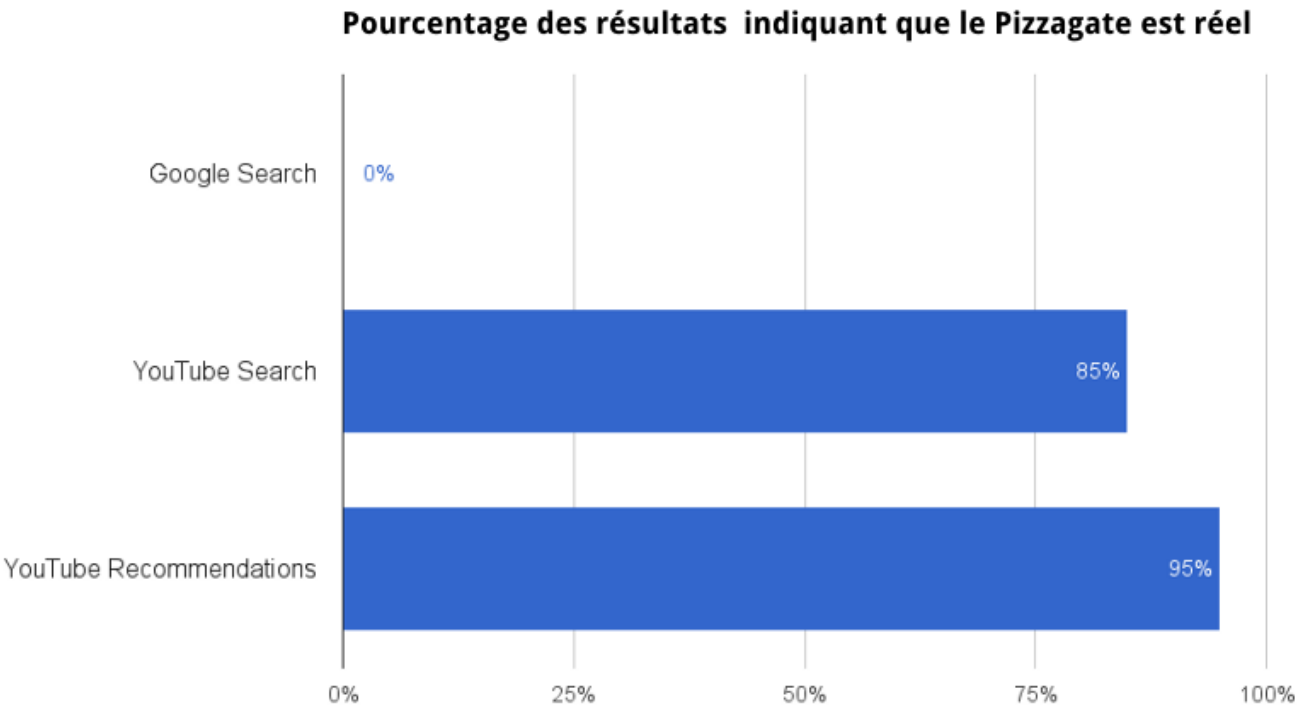
Pourcentage des résultats à propos du pape affirmant qu'il est «le mal», «satanique», «l'antéchrist»

3 –Science : « Le réchauffement climatique est-il une réalité ? »



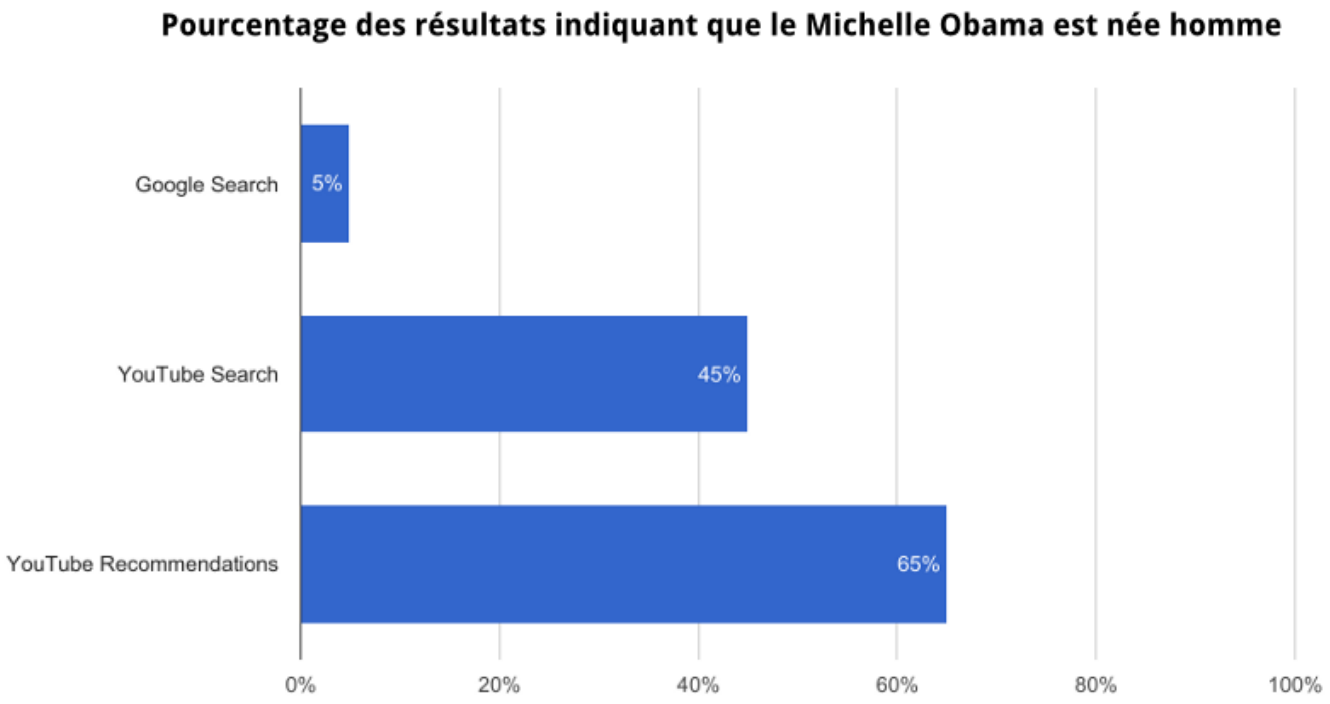
Pourcentage des résultats affirmant que le réchauffement climatique est un canular

4 –Conspirations : « Est-ce que le Pizzagate est vrai ? »



Le Pizzagate est une théorie du complot selon laquelle les Clinton auraient été à la tête d'un réseau pédophile en lien avec une pizzeria de Washington. Des vidéos faisant la promotion de cette théorie ont été recommandées des millions de fois sur YouTube pendant les mois précédant l'élection présidentielle américaine de 2016.

5 – Célébrités: « Qui est Michelle Obama ? »



Pourquoi les recommandations sont-elles différentes des résultats de recherche ?

Dans ces exemples, une recherche YouTube et une recommandation YouTube produisent des résultats étonnamment différents, alors que les deux algorithmes utilisent les mêmes données. Cela montre que de petites différences dans les algorithmes peuvent produire de grosses différences dans les résultats. La recherche est probablement optimisée dans un objectif de pertinence, alors que les recommandations prennent sûrement davantage en compte le temps de visionnage.

YouTube ne recommande pas ce que les gens « aiment »

Étonnamment, on remarque que les « j'aime » ou « je n'aime pas » (pouce bleu ou rouge) ont peu d'impact sur les recommandations. Par exemple, beaucoup de vidéos qui prétendent que Michelle Obama est « née homme » ont plus de pouces rouges que de bleus, et pourtant elles sont toujours fortement recommandées sur YouTube. Il semble que YouTube

accorde davantage d’importance au temps de visionnage qu’aux « j’aime ».

Ainsi, si « la Terre est plate » maintient les utilisateurs connectés plus longtemps que « la Terre est ronde », cette théorie sera favorisée par l’algorithme de recommandation.

L’effet boule de neige favorise les théories du complot.

Une fois qu’une vidéo issue d’une théorie du complot est favorisée par l’I.A., cela incite les créateurs de contenus à charger des vidéos supplémentaires qui confirment le complot. En réponse, ces vidéos supplémentaires font augmenter les statistiques en faveur du complot. Et ainsi, le complot est d’autant plus recommandé.

Finalement, le nombre important de vidéos qui soutiennent une théorie du complot rend cette dernière plus crédible. Par exemple, dans l’une des vidéos sur le thème de « la terre plate », l’auteur a commenté

> *Il y a 2 millions de vidéos sur la « terre plate » sur YouTube, ça ne peut pas être des c***!*

Ce que nous pouvons faire

L’idée ici n’est pas de juger YouTube. Ils ne le font pas intentionnellement, c’est une conséquence involontaire de l’algorithme. Mais chaque jour, les gens regardent plus d’un milliard d’heures de contenu YouTube.

Parce que YouTube a une grande influence sur ce que les gens regardent, il pourrait également jouer un rôle important en empêchant la propagation d’informations alternatives, et le premier pas vers une solution serait de mesurer cela.

Faites des expériences avec l’explorateur de recommandations

si vous souhaitez découvrir ce que YouTube recommande le plus au sujet des thèmes qui vous tiennent à cœur.

Les algos peuvent vous pourrir la vie

Les algorithmes[1] ne sont guère qu'une série d'instructions pas-à-pas généralement exécutées par un programme sur une machine. Cependant leur complexité et leur opacité pour le commun des mortels sont redoutables, et bien plus encore leur omniprésence dans tous les compartiments de notre vie, y compris la plus intime. Si le code fait la loi, c'est justement parce que les algorithmes sont à la fois puissants, invasifs et sont devenus aujourd'hui indispensables.
L'article ci-dessous ne met pas l'accent sur les nombreux domaines où nous utilisons des algorithmes sans en avoir conscience, il pointe davantage les risques et menaces qu'ils représentent lorsque ce sont les algorithmes qui déterminent notre existence, à travers quelques exemples parmi bien d'autres. Il pose également l'intéressante question de la responsabilité de ceux qui élaborent les algorithmes. Suffira-t-il de réclamer des concepteurs d'algorithmes un sympathique engagement solennel à la manière de celui des acteurs du Web ?

Les codeurs dont les algos contrôlent nos vies, qui les contrôle ? Pouvons-nous avoir un droit de regard sur les algorithmes qui désormais menacent de régir nos vies ?

Les algorithmes sont formidables mais peuvent aussi ruiner des vies

Extrait de l'essai (en anglais) The Formula: How Algorithms Solve All Our Problems—and Create More par **Luke Dormehl.**

Source : article du magazine **Wired** Algorithms are great and all, but can also ruin our lives

Traduction Framalang : Wan, r0u, goofy, Sphinx, sinma, Omegax, ylluss, audionuma

Le 5 avril 2011, John Gass, 41 ans, a reçu un courrier du service d'enregistrement des véhicules motorisés (Registry of Motor Vehicles ou RMV) de l'État du Massachusetts. La lettre informait M. Gass que son permis de conduire avait été annulé, qu'il lui était désormais interdit de conduire et que cela prenait effet immédiatement. Le seul problème, c'est qu'en bon conducteur n'ayant pas commis d'infraction grave au code de la route depuis des années, M. Gass n'avait aucune idée du motif de ce courrier.

Après plusieurs appels téléphoniques frénétiques, suivis par une entrevue avec les fonctionnaires du service, il en a appris la raison : son image avait été automatiquement signalée par un algorithme de reconnaissance faciale conçu

pour parcourir une base de données de millions de permis de conduire de l'État, à la recherche de possibles fausses identités criminelles. L'algorithme avait déterminé que Gass ressemblait suffisamment à un autre conducteur du Massachusetts pour présumer d'une usurpation d'identité, d'où le courrier automatisé du RMV.

Les employés du RMV se sont montrés peu compréhensifs, affirmant qu'il revenait à l'individu accusé de prouver son identité en cas d'erreur quelconque et faisant valoir que les avantages de la protection du public l'emportaient largement sur les désagréments subis par les quelques victimes d'une accusation infondée.

John Gass est loin d'être la seule victime de ces erreurs d'algorithmes. En 2007, un bogue dans le nouveau système informatique du Département des services de santé de Californie a automatiquement mis fin aux allocations de milliers de personnes handicapées et de personnes âgées à bas revenus. Leurs frais d'assurance maladie n'étant plus payés, ces citoyens se sont alors retrouvés sans couverture médicale.

Là où le système précédent aurait notifié les personnes concernées qu'elles n'étaient plus considérées comme éligibles aux allocations en leur envoyant un courrier, le logiciel maintenant opérationnel, CalWIN, a été conçu pour les interrompre sans avertissement, à moins de se connecter soi-même et d'empêcher que cela n'arrive. Résultat : un grand nombre de ceux dont les frais n'étaient plus pris en charge ne s'en sont pas rendu compte avant de recevoir des factures médicales salées. Encore beaucoup n'avaient-ils pas les compétences nécessaires en anglais pour naviguer dans le système de santé en ligne et trouver ce qui allait de travers.

Des failles similaires sont à l'origine de la radiation de votants des listes électorales sans notification, de petites entreprises considérées à tort comme inéligibles aux contrats gouvernementaux, et d'individus identifiés par erreur comme

« parents mauvais payeurs ». Comme exemple notable de ce dernier cas, Walter Vollmer, mécanicien de 56 ans, a été ciblé à tort par le Service fédéral de localisation des parents, et s'est vu envoyer une facture de pension alimentaire à hauteur de 206 000 $. L'épouse de M. Vollmer, 32 ans, a par la suite montré des tendances suicidaires, persuadée que son mari avait eu une vie cachée pendant la majeure partie de leur mariage.

Une possibilité tout aussi alarmante : qu'un algorithme puisse ficher par erreur un individu comme terroriste. Un sort qui attend chaque semaine environ 1500 voyageurs malchanceux qui prennent l'avion. Parmi les victimes passées de ces erreurs de corrélation de données, on retrouve d'anciens généraux de l'armée, un garçon de quatre ans, ainsi qu'un pilote d'*American Airlines*, qui a été détenu 80 fois au cours d'une même année.

Beaucoup de ces problèmes sont dus aux nouveaux rôles joués par les algorithmes dans l'application de la loi. Les budgets réduits menant à des réductions de personnel, les systèmes automatisés, auparavant de simples instruments administratifs, sont maintenant des décideurs à part entière.

Dans nombre de cas, le problème est plus vaste que la simple recherche d'un bon algorithme pour une tâche donnée. Il touche à la croyance problématique selon laquelle toutes les tâches possibles et imaginables peuvent être automatisées. Prenez par exemple l'extraction de données, utilisée pour découvrir les complots terroristes : de telles attaques sont statistiquement rares et ne se conforment pas à un profil bien défini comme, par exemple, les achats sur Amazon. Les voyageurs finissent par abandonner une grande partie de leur vie privée au profit des algorithmes d'extraction de données, avec peu de résultats, si ce n'est des faux-positifs. Comme le note Bruce Schneier, le célèbre expert en sécurité informatique :

> *Chercher des complots terroristes… c'est comme chercher une aiguille dans une botte de foin, ce n'est pas en accumulant*

> *davantage de foin sur le tas qu'on va rendre le problème plus facile à résoudre. Nous ferions bien mieux de laisser les personnes chargées d'enquêtes sur de possibles complots prendre la main sur les ordinateurs, plutôt que de laisser les ordinateurs faire le travail et les laisser décider sur qui l'on doit enquêter.*

Bien qu'il soit clair qu'un sujet aussi brûlant que le terrorisme est un candidat parfait pour ce type de solutions, le problème central se résume encore une fois à cette promesse fantomatique de *l'objectivité* des algorithmes. « Nous sommes tous absolument effrayés par la subjectivité et l'inconstance du comportement humain », explique Danielle Citron, professeur de droit à l'Université du Maryland. « Et à l'inverse, nous manifestons une confiance excessive pour tout ce que peuvent accomplir les ordinateurs ».

Le professeur Citron suggère que l'erreur vient de ce que nous « faisons confiance aux algorithmes, parce que nous les percevons comme objectifs, alors qu'en réalité ce sont des humains qui les conçoivent, et peuvent ainsi leur inculquer toutes sortes de préjugés et d'opinions ». Autrement dit, un algorithme informatique a beau être impartial dans son exécution, cela ne veut pas dire qu'il n'a pas de préjugés codés à l'intérieur.

Ces erreurs de jugement, implicites ou explicites, peuvent être causées par un ou deux programmeurs, mais aussi par des difficultés d'ordre technique. Par exemple, les algorithmes utilisés dans la reconnaissance faciale avaient par le passé de meilleurs taux de réussite pour les hommes que pour les femmes, et meilleurs pour les personnes de couleur que pour les Blancs.

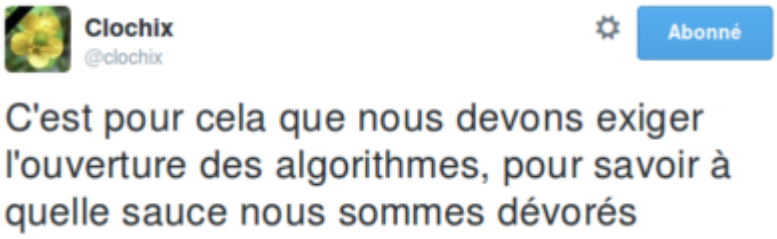
Ce n'est pas par préjugé délibéré qu'un algorithme ciblera plus d'hommes afro-américains que de femmes blanches, mais cela ne change rien au résultat. De tels biais peuvent aussi

venir de combinaisons plus abstraites, enfouies dans le chaos des corrélations de jeux de données.

Prenez par exemple l'histoire de l'afro-américaine Latanya Sweeney, docteure de l'Université d'Harvard. En effectuant des recherches sur Google, elle fut choquée de découvrir que les résultats de ses recherches étaient accompagnés de publicités demandant : « Avez-vous déjà été arrêté(e) ? ». Ces annonces n'apparaissaient pas pour ses collègues blancs. Sweeney se lança alors dans une étude, démontrant que les outils d'apprentissage automatique utilisés par Google étaient incidemment racistes, en associant plus souvent des noms donnés à des personnes noires avec des publicités ayant trait aux rapports d'arrestation.

Le système de recommandation de Google Play révèle un problème similaire : il suggère aux utilisateurs qui téléchargent *Grindr*, un outil de réseautage social basé sur la localisation pour les gays, de télécharger également une application qui assure le suivi géolocalisé des délinquants sexuels. Au vu de ces deux cas, devons-nous conclure que les algorithmes ont fait une erreur, ou plutôt qu'ils sont révélateurs des préjugés inhérents à leurs concepteurs ? Ou, ce qui semble plus probable, ne seraient-ils pas révélateurs d'associations inappropriées et à grande échelle entre – dans le premier cas – les personnes noires et le comportement criminel, et – dans le deuxième cas – l'homosexualité et les agressions sexuelles ?

Peu importe la raison, peu importe la façon répréhensible dont ces corrélations codifiées peuvent exister, elles révèlent une autre face de la culture algorithmique. Quand un seul individu fait explicitement une erreur de jugement, il ne peut jamais affecter qu'un nombre fini de personnes. Un algorithme, quant à lui, a le potentiel d'influer sur un nombre de vies exponentiellement plus grand.



Pour aller plus loin, 4 articles en français sur le même sujet :

- Surveiller les algorithmes
- Ces algorithmes qui vous nous gouvernent
- Ouvrir les modèles, pas seulement les données
- Le jaguar et le bus scolaire

Note

[1] Pour une définition plus élaborée voir Qu’est-ce qu’un algorithme