

L'enjeu de l'ouverture du code des algorithmes, tout comme celui de l'ouverture des jeux de données utilisés dans le cadre des artefacts génératifs (MidJourney, ChatGPT, etc) n'est plus une question de compréhension qui vise à documenter un « comment » et une ingénierie. C'est une question d'appréhension sur les nouveaux et lancinants régimes discursifs qui sont venus se substituer à la matrice initiale du web dans laquelle chaque individu pouvait posséder (et administrer) une page et une adresse et où l'on pouvait entretenir des proximités sans jamais sacrifier à des formes de promiscuités.

Après les mensonges de l'industrie du tabac sur sa responsabilité dans la conduite vers la mort de centaines de millions de personnes, après les mensonges de l'industrie du pétrole sur sa responsabilité dans le dérèglement climatique, nous faisons face aujourd'hui au troisième grand mensonge de notre modernité. Et ce mensonge est celui des industries extractivistes de l'information, sous toutes leurs formes. Nous avons aimé croire que le calcul intensif se ferait sans travail intensif, que le Data Mining ne nécessiterait pas de mineurs de fond, que l'informatique en nuage (Cloud Computing) ne dissimulait pas la réalité d'une industrie lourde.

Nous ne pouvons plus aujourd'hui nous réfugier dans ces mensonges. Sur les industries extractivistes de l'information, nous avons l'avantage d'en connaître déjà les mécanismes et les routines ; la chance d'en observer les infrastructures de marché (du Cloud Computing au High Frequency Trading en passant par la précarisation des différentes formes de Digital Labor) ; la chance d'être capables de documenter la toxicité de ces prismes dans le cadre de certains sujets de société ; la chance d'avoir pu documenter et prouver à de trop nombreuses reprises l'insincérité fondamentale et aujourd'hui fondatrice de toutes ces plateformes et de leurs créateurs et administrateurs. Et même s'ils s'inscrivent, comme je le rappelais plus haut, dans un écosystème médiatique, économique et politique bien plus vaste qu'eux, leur part émergée, c'est à dire les médias sociaux, sont aujourd'hui pour l'essentiel de même nature que la publicité et le lobbying le furent pour l'industrie du tabac et du pétrole : des outils au service d'une diversion elle-même au service d'une perversion qui n'est alimentée que par la recherche permanente du profit.



Ouvrir le code des algorithmes ne suffit plus

Olivier Ertzscheid
Paru sur AOC – Septembre 2023

propriété qu'il n'est pas sensitif (sic) à l'échelle. Plus vous mettez de serveurs et mieux ça va. »

Aujourd'hui, ce principe fondamental est entièrement nié. Non seulement « l'hypertexte » au sein des médias sociaux est réduit à peau de chagrin et ne fonctionne le plus souvent que comme un sous-texte normatif contraint et souvent empêché (le « like »¹⁰ a tué le lien), mais plus on dispose de serveurs et plus ils sont centralisés au service des quelques grandes entreprises qui sont la matrice dominante du web contemporain. Ainsi ce passage à l'échelle, cette « scalabilité » technique, qui formait, dans l'esprit de Cailliau et des pionniers de l'internet et du web, une garantie de diversité, s'est aujourd'hui presque entièrement dissoute et renversée. Plus on met de serveurs (et de contenus et de gens les produisant et de gens les faisant circuler) dans des architectures techniques centralisées et « moins bien ça va ».

Et s'il demeure un point particulièrement saillant dans la compréhension de la nature profonde de ces fonctionnements algorithmiques, c'est aujourd'hui davantage la manière dont ils rendent possible le fait que quelques très grands ou gros comptes (que l'on appelle des « power users » ou « super utilisateurs ») suffisent à entièrement dicter un agenda discursif et médiatique qui force les effets de récursivité dans le champ social. Pour le dire simplement, c'est donc un (gros) problème de démocratie « interne ».

Et c'est enfin, bien sûr, un problème de gouvernance. « Il n'y a pas d'algorithme, il n'y a que la décision de quelqu'un d'autre » écrivait en 2017 Antonio Casilli dans la préface qu'il me faisait l'honneur d'écrire pour mon premier livre chez C&F Éditions, *L'appétit des géants*¹¹. Dans le paysage actuel des médias sociaux, ces décisions sont, pour l'essentiel, devenues lisibles et explicites. Et nous affrontons leurs effets à l'échelle de collectifs que nous n'imaginions pas pouvoir devenir aussi puissants. Il nous faut donc, tout en continuant de documenter leurs causes et leurs modes opératoires, enquêter sans relâche sur leurs effets dans le champ social, économique, politique et médiatique.

¹⁰ <https://affordance.framasoft.org/2010/05/le-like-tuera-le-lien/>

¹¹ <https://cfeditions.com/geants/>

Parvenir à ouvrir le code algorithmique était, il y a quelques années et à l'échelle des moteurs de recherche, une urgence démocratique vitale et une fin en soi. Et une option suffisante pour en comprendre les principales dynamiques informationnelles et y détecter les matrices économiques et idéologiques à l'œuvre. Ce n'est plus le cas aujourd'hui car ces matrices se sont déplacées et sont, la plupart du temps, désormais explicites et assumées. Ainsi Google ne tente même plus de cacher qu'il met en avant ses propres produits et son propre écosystème de services dans le cadre des requêtes de son moteur de recherche et assume être allé au bout de la logique consistant à mettre les liens publicitaires devant les résultats « organiques ». Musk assume sa conception « maximaliste » de la liberté d'expression. Il assume de rétablir des comptes masculinistes violents, des comptes d'extrême-droite, et ainsi de suite. Point n'est besoin d'ouvrir le code de l'algorithme pour l'observer et le documenter. C'est également le cas de toute une série de médias sociaux qui portent et assument une ligne idéologique souvent radicale d'un bout à l'autre de l'échiquier politique (mais surtout à l'extrême-droite quand même...).

Dès lors qu'un média social est lisible dans ses intentions, dès qu'il se dote d'une ligne éditoriale, la question de l'importance d'avoir accès au code source de son algorithme change d'objet : il ne s'agit plus d'aller y enquêter pour dresser la liste des éléments qui permettent de porter cette ligne éditoriale au premier plan, mais de détecter ce qui rend encore possible l'existence de poches d'altérité qui permettent de la combattre, de s'en éloigner ou de s'en protéger. C'est d'ailleurs l'une des ambivalences fondamentales de ces médias sociaux que de proposer, lorsqu'on les observe à l'échelle macro, une récurrence délétère de discours toxiques, mais lorsqu'on les enquête à l'échelle individuelle, de permettre de maintenir une matrice de liens qui fonctionne comme un abri idéologique. Le problème étant que les grands enjeux politiques et sociétaux ne sont mis en débat et ne se négocient que très rarement ... à l'échelle individuelle.

L'un des pionniers de l'internet et du web, Robert Cailliau livrait ce témoignage dans une série d'entretiens à France Culture⁹ autour de l'histoire du web : « le web est l'hypertexte le plus cru, le plus stupide, le plus con que l'on puisse imaginer mais il a comme grande

Depuis son arrivée à la tête de Twitter rebaptisé X, Elon Musk a bien tenu sa promesse d'ouvrir certaines parties du code source de la plateforme. Cette réponse à de vieilles revendications militantes pour plus de transparence des plateformes ne manque, en vérité, pas de cynisme. Il est devenu transparent que le code de Twitter favorisait certains contenus et en défavorisait d'autres, au bon vouloir de Musk. L'ouverture du code ne suffit plus, aujourd'hui, pour donner des outils de contrôle à ce que les grandes plateformes font de nos espaces d'expression et de nos démocraties.



C'est un débat qui existe depuis que le web est devenu un média de masse. Celui autour de l'ouverture des algorithmes qui le rendent, pour chacun d'entre nous, traversable, navigable et praticable. Un débat pour comprendre comment fonctionnent les moteurs de recherche et désormais ces « jardins fermés » que sont les médias sociaux. Longtemps les grands acteurs du web s'abritèrent derrière deux arguments, le « secret industriel » d'une part, et l'absence de nécessité d'autre part, au motif que « nous » serions les premiers et les seuls responsables de l'organisation de l'information en ligne, les classements des moteurs de recherche ou les « fils » et autres « murs » des médias sociaux n'en étant que le reflet instrumental. Les Big Tech ne se voulaient qu'hébergeurs et en aucun cas éditeurs.

Et puis deux choses se produisirent. D'abord le fait que cette ouverture du code devienne un argument sociétal puis politique. Car plus les usages du web se massifiaient, plus les moteurs de recherche et plus les réseaux et médias sociaux devenaient l'essentiel de cette masse, et plus il devenait évident qu'ils avaient une lourde responsabilité dans le choix de ce qu'ils rendaient visible ou invisible. Et qu'il fallait faire quelque chose : « Il est vain de réclamer la dissolution de Google ou d'un autre acteur majeur comme il est vain d'espérer un jour voir ces acteurs "ouvrir" complètement leurs algorithmes. Mais il devient essentiel d'inscrire enfin clairement, dans l'agenda politique, la question du rendu public de fonctionnements algorithmiques directement assimilables à des formes classiques d'éditorialisation. » (*Libération*, tribune de votre serviteur, 2015)

⁹ <https://www.radiofrance.fr/franceculture/podcasts/serie-une-histoire-de-l-internet>

Personne ne conteste aujourd'hui le rôle que l'éditorialisation algorithmique¹ joue dans l'organisation des informations en ligne et donc dans les débats publics de notre société. Mais ce constat fut long à établir (j'avais avec d'autres, commencé à la souligner dès 2004). Et le temps d'y parvenir, le web et les médias sociaux avaient, en partie, changé de nature. La clé de leur modèle économique restait leur capacité de capter et de monétiser notre attention, mais le cœur du réacteur n'était plus uniquement et prioritairement les algorithmes eux-mêmes mais aussi et surtout les très larges jeux de données (personnelles² ou non) qu'ils étaient en capacité de capter, de conserver et d'adresser à des fins marchandes. Voilà pourquoi on vit ce que l'on pensait alors impensable avec des acteurs comme Amazon ouvrant le code de l'algorithme de recommandation qui avait forgé son succès et son hégémonie.

Lors de son arrivée à la tête de Twitter rebaptisé X, Elon Musk avait fait tout un tas de promesses. Notamment celle « d'ouvrir » le code de l'algorithme. Ce qu'il a fait en grande partie. Et nous avons alors pu constater ce que nous étions jusqu'ici soit contraints de conjecturer en pratiquant des méthodes « d'ingénierie inversée » (en gros par essai et erreur), soit d'observer grâce aux témoignages d'ingénieurs ou de cadres repentis devenus autant de lanceurs et lanceuses d'alerte : Tristan Harris, Frances Haugen, etc.

Avec la promesse tenue de Musk, nous avons constaté, dès le mois d'avril³ 2023 avec le dépôt dans Github d'une partie importante du code de la plateforme, quelles étaient les métriques et le poids qui leur était attribué en termes de visibilité, découvert qu'un Like valait plus qu'un Retweet, qu'il valait mieux mettre des images et des gifs animés, qu'il fallait éviter les liens externes vers des concurrents, et ainsi de suite. Tout une série de métriques⁴ comme autant de coups de trique.

Mais nous avons découvert tout cela à la lumière de ce que nous avons déjà appris de ces plateformes technologiques : d'abord qu'un algorithme est toujours en mouvement et que ses paramètres et le poids qui leur est accordé changent (très) régulièrement, et ensuite que ces plateformes mentent délibérément y compris sur les

la fabrique d'une opinion alignée sur la ligne idéologique de leurs propriétaires (et/ou jouant sur les plus viles pulsions de son audience) n'a rien de nouveau. Mais la nature si particulière de notre rapport aux médias numériques vient d'une réalité simple : si nous voulons lutter contre Bolloré nous pouvons regarder ou ne pas regarder, acheter ou ne pas acheter, écouter ou ne pas écouter. Notre choix reste binaire parce qu'à aucun moment nous n'envisageons la possibilité de négocier avec lui. A l'échelle des médias sociaux, nos expressions publiques comme privées sont en permanence situées dans des espaces troubles de négociation qui sont eux-mêmes traversés de saillances le plus souvent imprévisibles. Nous sommes par exemple aussi surpris de la viralité soudaine d'un de nos posts que nous n'imaginions même pas, que nous sommes déçus de l'absence de réaction conforme à nos attentes sur un autre post que nous pensions devoir mériter d'être mieux exposé.

La radio, la télévision, la presse sont des médias qui jamais ne déçoivent car ils sont strictement alignés sur des représentations qui sont aussi attendues que connues de leurs publics. D'ailleurs « la » radio, « la » télévision et « la » presse n'existent pas. France Culture et Radio Courtoisie existent. CNews et Arte existent. *L'Humanité* et *Valeurs Actuelles* existent. Ils ne négocient pas avec nous autre chose que notre approbation. Les médias sociaux ont, eux, une part déceptive bien plus constante qui est le comburant premier de nos assignations attentionnelles. Cette part déceptive est entretenue, déterminée, orchestrée par les logiques instrumentales de la FOMO (la peur de manquer quelque chose – *Fear Of Missing Out*). Le plus souvent nous constatons à la fin que nous n'avons rien raté d'important et nous sommes déçus.

Mais à d'autres occasions nous avons confirmé l'hypothèse selon laquelle nous aurions en effet pu rater quelque chose d'important. Cette possibilité de voir nos déceptions comblées et réparées est une assignation à négocier en permanence l'articulation des discours auxquels nous sommes exposés. Ne reste aux algorithmes qu'à entretenir l'illusion que nous sommes à l'initiative de la négociation, et que nos déceptions s'apparentent à un désordre amoureux alors qu'elles ne sont au fond, qu'un ordonnancement idéologique le plus souvent soluble dans l'ordre du marché.

¹ <https://affordance.framasoft.org/2016/05/editorialisation-algorithmique/>

² <https://affordance.framasoft.org/2016/05/nationaliser-code-republique-algorithmique/>

³ <https://www.radiofrance.fr/franceinter/ce-que-nous-apprend-le-code-de-l-algorithme-de-twitter-rendu-public-par-elon-musk-7968092>

⁴ <https://affordance.framasoft.org/2014/07/maudites-metriques-merdiques-de-mots/>

Le problème de Twitter / X aujourd'hui, comme celui de Facebook, d'Instagram, de TikTok, de Snapchat et d'autres plateformes numériques n'est plus réductible à la seule question de leurs déterminismes algorithmiques. Il doit être réfléchi à la hauteur de l'ensemble de ce que j'appelle leurs architectures techniques toxiques, c'est à dire en partie, mais en partie seulement, la question algorithmique ; mais également le maillage de ces espaces d'expression avec d'autres espaces sociaux et médiatiques et la manière dont ils s'entremêlent ; mais aussi avec l'ensemble des secteurs dans lesquels ces plateformes opèrent (qui pour la majorité d'entre elles sont loin d'être réductibles au seul secteur de la tech) avec tout l'impact et les influences économiques et politiques que cela mobilise ; et enfin en regardant ces plateformes et médias comme la partie visible d'infrastructures extractivistes bien plus vastes et plus massives dans lesquelles le « média social » constitue de plus en plus une forme de leurre stratégique qui tend à gommer ou à dissimuler les matrices de pouvoir dont elles se nourrissent et qu'elles orchestrent.

Je m'explique rapidement sur ce dernier point : lorsque Google s'est mis en 2005 à investir massivement et à perte dans la numérisation d'ouvrages du domaine public, personne n'a compris assez vite qu'il poursuivait en fait un autre but qui était d'optimiser la sémantique de son moteur de recherche et de ce que l'on appelle aujourd'hui un grand modèle de langage (Large Language Model, celui de Google s'appelle PaLM 2). De la même manière, j'avais souligné à quel point le rachat de Twitter par Musk, qui apparaissait à certains comme la pochade d'un multi-milliardaire excentrique, poursuivait un autre but que de « simplement » s'offrir un outil pour étayer sa conception maximaliste de la liberté d'expression, et qu'il fallait y détecter la mainmise sur la plus grande base de donnée conversationnelle⁸ dynamique de la planète qu'il aurait tôt fait de pouvoir utiliser dans le cadre de ses plans de développement sur le secteur de l'intelligence artificielle et de faire percoler dans plusieurs des autres secteurs où il est implanté.

Derrière les algorithmes, ces plateformes posent aujourd'hui la question de leur place dans un espace public médiatique démocratique. Ô bien sûr pas de manière totalement inédite. De William Randolph Hearst à Vincent Bolloré, la mainmise sur des médias au service de

métriques qu'elles mettent en place et font semblant de laisser voir ou entrevoir (comme Facebook⁵ l'avait fait par exemple sur le nombre de vues des vidéos postées ou sur leur taux d'engagement).

Disposer du code source est donc une chose, savoir si la version dont on dispose est sincère en est une seconde, et détecter les modifications effectuées sur un temps donné par rapport au code source mis en accès public ... en est une troisième. D'autant qu'il n'est pas, à l'échelle de chacune de ces plateformes, « un » algorithme unique et omnipotent, mais plusieurs algorithmes, co-occurents et concurrents, qui sont autant de déclinaisons d'une matrice d'intention et d'objectifs, et qui remplissent chacun des tâches définies. Et qu'en plus de cela, la matrice algorithmique, si elle comprend des éléments figés (les instructions et les règles), s'inscrit aussi dans une dynamique contextuelle capable de faire varier certaines de ces instructions et de ces règles.

Depuis son arrivée à la tête de Twitter désormais X, Elon Musk a effectué un certain nombre de changements dans cette matrice algorithmique. Les derniers en date ont été relevés et chroniqués par différents comptes sur la plateforme X, notamment « NFT_GOD » et « Tibo_Maker »⁶.

On y apprend cette fois que dans l'algorithme voulu par Elon Musk, outre ses travers égocentriques consistant à proposer systématiquement son compte (et ses tweets) aux nouveaux venus, il vaut mieux « répondre » que « reposer » et qu'il faut éviter de faire les deux sur un même contenu ; qu'il faut utiliser des images et de la vidéo pour être « boosté » par l'algorithme ; qu'il faut parler des tendances les plus populaires (trending topics) pour augmenter ses chances de voir son compte proposé dans l'onglet « pour vous » ; qu'il ne faut surtout pas mentionner des concurrents ou mettre des liens vers des sites concurrents (vers des Reels Instagram par exemple) sous peine de se voir invisibiliser ; qu'il ne faut pas non plus publier du contenu « qui ne fasse pas écho auprès de vos followers habituels » (sic) ; que la durée de vie d'un post est de 24h s'il rencontre une forte audience ou un nombre important d'interactions et qu'elle peut être portée à 48h si vraiment il cartonne ; que vos réponses à l'un de vos posts

⁵ <https://www.lefigaro.fr/secteur/high-tech/2018/11/04/32001-20181104ARTFIG00181-video-sur-facebook-petit-mensonge-grosses-consequences.php>

⁶ https://nitter.net/NFT_GOD/status/1688584001466875904?s=20 et https://nitter.net/tibo_maker/status/1691428242924740608

⁸ <https://aoc.media/opinion/2022/05/01/loiseau-le-milliardaire-et-le-precipice/>

valent le double de celles des autres et que vous avez donc intérêt (pour la plateforme ...) à répondre aux gens qui commentent vos posts ; qu'il faut ajouter à vos favoris des posts à forte visibilité ; et que bien sûr il vaut mieux choisir ... la formule payante de Twitter / X.

On apprend aussi, mais cette fois en dehors du code disponible, qu'Elon Musk aurait fait le choix de ralentir le trafic entrant vers un certain nombre de sites dont, excusez du peu, Facebook, Instagram, Bluesky, Substack mais aussi Reuters et ... le *New-York Times*, qu'Elon Musk a souvent critiqué et accusé de diffuser des Fake News (sic). Si vous publiez, sur X, un lien qui pointe vers ces sites, non seulement la visibilité de votre compte sera diminuée, mais celles et ceux qui cliqueront sur votre lien auront un délai de 5 secondes avant de pouvoir l'afficher et y accéder. Quand on sait que nous n'avons en général pas la patience d'attendre plus de 2 ou 3 secondes avant d'accéder à un contenu en ligne et que nous préférons souvent passer à autre chose et y renoncer s'il ne répond pas dans ce délai, et quand on voit la nature de sites ciblés par Elon Musk, on mesure toute la toxicité démocratique de ces paramètres algorithmiques et ce qu'ils disent de l'éthique ou de l'idéologie de celles et ceux qui les déterminent et les mettent en œuvre. Depuis l'article sur cette possible malversation par le *Washington Post*⁷, il semble que ce délai n'ait plus cours.

Voilà donc quelques-uns des éléments que l'on apprend grâce à « l'ouverture de l'algorithme » de Twitter / X.

Mais qu'apprend-on réellement ? Les points évoqués ci-dessus sont, pour l'essentiel, des points clés de tout écosystème numérique et l'héritage des règles de référencement ou de positionnement (SEO et SMO pour Search Engine Optimization et Social Media Optimization). Et y compris dans le cadre de plateformes qui n'ont pas ouvert leur algorithme – Facebook par exemple – chacun sait que le même type de règles sont appliquées.

On n'apprend donc plus grand-chose de l'ouverture des algorithmes mais l'on a des confirmations, des preuves, et cette ouverture est et demeure le seul moyen de vérifier l'authenticité de ce que nous en savons par intuition, par comparaison, par réflexion et par

observation. Voilà pourquoi il est tout à fait impératif de continuer à œuvrer pour une ouverture maximale de la partie du code algorithmique relevant de l'éditorialisation.

Mais pour le reste, pour l'essentiel des logiques de visibilité ou d'invisibilité, pour ce qui est de comprendre quels contenus sont les plus mis en avant et pourquoi, nous savons déjà à peu près tout ce qu'il est nécessaire de savoir pour comprendre que cela n'est plus suffisant pour réguler l'usage de ces plateformes.

Par ailleurs ces plateformes et leurs algorithmes n'échappent pas aux grandes lois de la cybernétique et de la théorie des systèmes : l'information est une manière de lutter contre l'entropie, contre le chaos et le désordre. Plus un système dispose d'information, et plus il « s'ordonne », plus il se met « en forme » (in-forma). Les médias sociaux, traversés qu'ils sont par des volumes considérables d'information et guidés par ces autres flux informationnels que sont les algorithmes sont tout sauf le chaos désordonné et aléatoire dans lequel on les range parfois : ils sont au contraire la quintessence d'un ordre, parfaitement déterministe.

Il faut donc, bien sûr continuer de travailler à ouvrir les algorithmes pour que chacun soit à même de mieux comprendre la vision et l'ordre du monde qu'ils véhiculent. Il n'est d'ailleurs pour cela, contrairement à ce que l'on a longtemps cru et expliqué, pas nécessairement besoin d'être un expert du code informatique pour en approcher la logique profonde. Les algorithmes sont comme des ritournelles auxquelles on s'accoutume à force de les fréquenter, que l'on retient – et que l'on maîtrise parfois – dans une forme d'intelligence situationnelle altérée par l'expérience sans cesse renouvelée de cette fréquentation. Comme la ritournelle chez Deleuze et Guattari dans leur ouvrage *Mille Plateaux*, les algorithmes sont trois choses à la fois. D'abord ce qui nous rassure par une forme de régularité attendue, que l'on devine et anticipe. Ensuite ce qui installe l'organisation nous semblant familière d'un espace que l'on sait public mais que l'on perçoit et que l'on investit en partie comme intime : ils « enchantent » l'éventail de nos affects et sont l'état de nature de nos artifices sociaux. Enfin ils sont ce qui, parfois, nous accompagne et nous équipe aussi dans la découverte d'un ailleurs, parce qu'y compris au sein de représentations cloisonnées, ils sont des chants traversants.

⁷ <https://www.washingtonpost.com/technology/2023/08/15/twitter-x-links-delayed/>