



ÉCOLE CENTRALE LYON

NOUVELLES TECHNOLOGIES DE L'INFORMATION ET LA
COMMUNICATION
SYNTHÈSE DE VEILLE

Faut-il s'inquiéter du contrôle de contenu effectué par les géants du numériques ?

Élèves :
Pierre GRIMAL

Tuteur :
René CHALON

15 mars 2021

Table des matières

1	Différents types de réseaux	3
1.1	Population et utilisation	3
1.2	Contrôle et gestion du contenu	4
2	La construction d'un outil de parole	5
2.1	La question politique	5
2.2	La question légale	7

Introduction

L'idée de ce sujet m'est venue après les récents événements : l'invasion de militants pro-Trump au capitole, et la suspension définitive des comptes de certains réseaux sociaux importants du président des Etats-Unis. Même si certains adversaires politiques félicitent l'initiative, d'autres contestent cette décision prise à quelques jours de la fin du mandat du président américain. Aurore Bergé, député LREM déclare : « On peut combattre Trump et le chaos qu'il a entraîné, mais refuser de se réjouir de voir les GAFA décider seuls, sans contrôle du juge, sans recours possible, quel président en exercice a le droit d'avoir un compte ou non sur Twitter » [8].

C'est dans ce contexte que je me suis interrogé sur la question du contrôle de contenu de manière plus globale, en choisissant un sujet large et volontairement à caractère dissertatif. Le but principal de cette étude est d'essayer de comprendre tous les mécanismes et événements qui ont mené à la situation actuelle, et de dresser un panorama de la légitimité des réseaux à contrôler leur contenu.

1 Différents types de réseaux

1.1 Population et utilisation

Il existe aujourd'hui une pluralité de réseaux sociaux ayant tous plus ou moins leur spécificité ou leur spécialisation. Bien que ce rapport se concentre principalement sur Facebook et Twitter, il est intéressant d'analyser les différences des règles d'utilisations et des populations de chacun de ces réseaux.

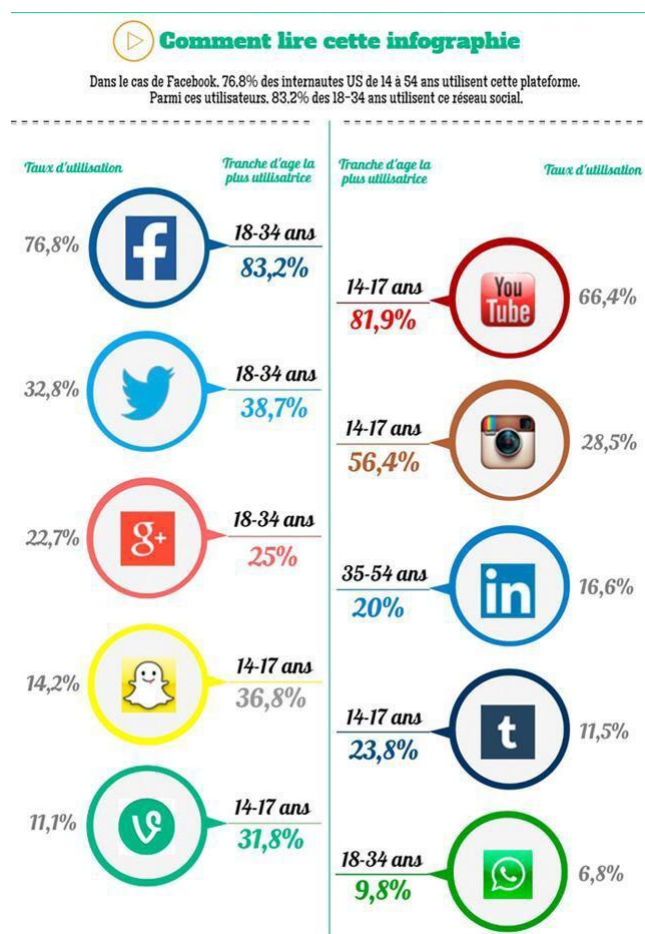


FIGURE 1 – Les tranches d'âge les plus utilisatrices des réseaux sociaux (2015)

Comme le montre ce graphique, les réseaux sociaux possèdent des populations différentes, en effet la plateforme de Microsoft : LinkedIn qui se veut être un réseau social professionnel concentre une population bien plus vieille et présente un besoin de modération bien moindre, tandis qu'Instagram, dont la majorité des utilisateurs ont entre 16 et 24 ans aujourd'hui possède des règles très strictes et une modération très agressive. De plus, les intentions et habitudes selon le réseau varient, certains sont d'avantages utilisés pour du divertissement ou de l'art, d'autres pour l'information. Facebook et Twitter sont notamment devenus au fil des années des outils essentiels de communication et d'information, offrant le droit à tous d'exprimer sa pensée tant qu'elle reste dans les règles d'utilisation.

Chaque réseau a ses propres règles, certaines sont un applicatif direct des lois, d'autres sont contingentes au réseau. Par exemple, Twitter possède un ensemble de

Key Moments in the History of the Facebook Algorithm

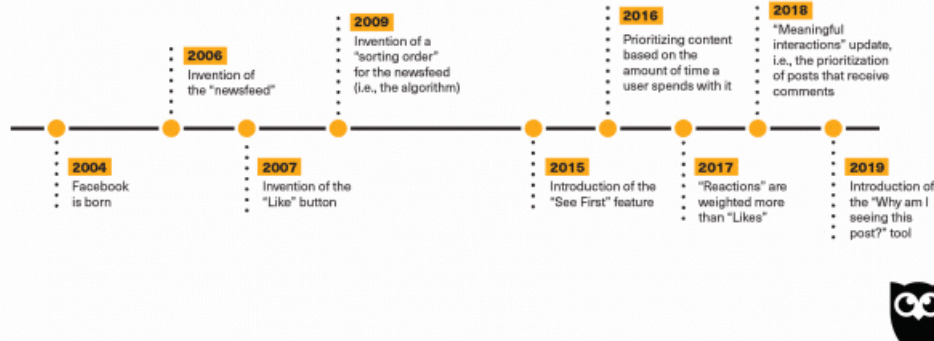


FIGURE 2 – Brève chronologie de l'algorithme Facebook

règle assez souple [9], en effet c'est un rare réseau tolérant des contenus explicites, des points de vue très tranchés et du langage grossier. Finalement aller à l'encontre d'une des règles de Twitter équivaut la plupart du temps à aller à l'encontre de la loi. Facebook et Instagram ont quant à eux des règles beaucoup plus strictes, Instagram interdit par exemple la semi-nudité féminine. En tant que acteur privé, les réseaux ont toute la légitimité à réguler leur contenu selon leurs normes, le problème apparaît dès lors qu'il y a une orientation politique du filtrage effectué.

1.2 Contrôle et gestion du contenu

Chaque réseau possède plus ou moins deux niveaux de gestion, le premier constitue le classement du contenu : pourquoi un utilisateur voit ce type de posts et a-t-il la main sur son fil d'actualités .

Comme le montre cette frise, Facebook a su très tôt faire intervenir l'utilisateur dans la gestion de son fil, d'abord de manière passive : l'utilisateur ajoute un like à un post, Facebook va donc lui proposer plus souvent du contenu similaire, les utilisateurs ayant également aimé le post sont ajoutés au même "cluster" d'individus et des suggestions venant de contenu apprécié par chacun d'entre eux sont proposées à ceux qui n'auraient pas encore labellisé de tel posts [12]. Facebook entame depuis 2016 une démarche plus participative, dans laquelle l'utilisateur est sollicité activement sur la gestion de son fil avec l'apparition des options "Ne plus afficher", "Voir en premier". Pour rappel l'algorithme de classement et les données récoltées sont essentiels pour l'équation économique de Facebook dont 98% du CA est issu des revenus publicitaires.

Le deuxième niveau de contrôle s'effectue au niveau du respect des règles d'utilisation. Il existe depuis quelques années des algorithmes de Deep Learning permettant avec un taux très élevé de précision, de détecter des images ou vidéos de nudité, des paroles vulgaires. Cependant certains contenus échappent au filtrage, volontairement plus souple sur les données textuelles, c'est pourquoi l'option "Signaler" est apparue sur la plupart des réseaux. Cette action permet à l'équipe de modération (puisque'il n'est pas encore possible d'automatiser tout le contrôle) de présélectionner le contenu sensible afin de légiférer sur sa présence.

Depuis ces dernières années, les géants (Amazon, Facebook, Google, etc.) se sont attachés à développer des technologies d'Intelligence artificielle de plus en plus abouti. Avec la démocratisation des fake news, l'avènement des Deepfake (des photos/vidéos/audio ultra réalistes qui utilisent des réseaux de type GAN pour générer le contenu souhaité) il était impératif de développer des algorithmes permettant leur détection [7], Facebook a même lancé en 2019 le *Deepfake Detection Challenge* pour concevoir le meilleur algorithme de reconnaissance de Deepfake. A titre d'exemple, au Gabon, l'utilisation de Deepfakes a même conduit à la tentative d'un coup d'État de leur gouvernement. Il est cependant bien plus difficile d'assurer la détection de fake news avec un haut taux de précision [11], politisant ainsi l'action des algorithmes.

2 La construction d'un outil de parole

2.1 La question politique

C'est en 2011, qu'on a pu constater les possibilités qu'offrait les réseaux sociaux pour contourner l'appareil de communication classique inhérents aux états. En effet, la majorité des révolution du printemps arabe [5] se sont organisées sur les réseaux sociaux, technologies encore trop récentes pour que les moyens étatiques de ces pays soient suffisants pour en assurer la surveillance. Ainsi les réseaux sociaux deviennent progressivement un outil essentiel de la liberté d'expression, une alternative viable offrant le droit à tous d'exprimer son point de vue et voyant la création de nombreux néo-médias.

C'est à partir de 2013/2014 qu'on constate l'autre pendant de son utilisation, et les failles de sécurité qu'elle offre. Effectivement, ce sont devenus les socles du recrutement de réseaux terroristes (Al-Qaida et Daesh notamment), jamais il n'a été aussi aisé de diffuser du contenu et manipuler les gens. Dans la foulée, des vidéos d'exécutions, du contenu incitant à la haine, des diffusions de contenu sexuel non consenti commencent leur prolifération sur les réseaux. Il devient impératif d'effectuer un filtrage afin d'assurer la fiabilité du réseau. C'est au même moment, que les technologies de Machine Learning connaissent d'énormes progrès.

En 2016, éclate le scandale de Cambridge Analytica. C'est l'aube d'une nouvelle ère à l'utilisation des données des électeurs pour les campagnes politiques, avec l'élection notable de Trump. Cambridge Analytica est une filiale du groupe de communication britannique Strategic Communication Laboratories. Le scandale est directement lié à l'utilisation non consentie de données labellisées passivement par le cœur électoral. En effet, Facebook a rendu possible aux statisticiens et expert en psychométrie l'identification des habitudes et préférences culturelles, sociales et religieuses de utilisateurs. Ce sont les données de 50 millions de compte qui ont pu être utilisé pour inciter et favoriser l'élection de Trump, puis le vote du Brexit [10]. Mark Zuckerberg sera notamment auditionné par le sénat américain, son réseau contraint de payer une amende de 5 milliards infligé par l'autorité américaine de régulation des communications. Dans la même lancée, Facebook est directement visé par loi de protection de données Européenne (RGPD), et sera condamné à payé de lourdes amendes. Aujourd'hui Facebook est valorisée à 525 milliards et a réalisé en 2020 un CA de 90 milliards, un système d'amende est-il suffisant pour limiter les potentielles actions de Facebook? Les géants

du numérique dominant à présent la dissémination de l'information et la mobilisation politique. Il ne s'agit pas seulement de concentration économique et de constitution de monopoles, ni d'atteintes à la vie privée des utilisateurs, ce qui est en cause à présent, c'est l'influence politique. "Ils ne se contentent pas de fausser le marché, ils menacent la démocratie", écrivent Francis Fukuyama, Barak Richman et Ashish Goel [4].

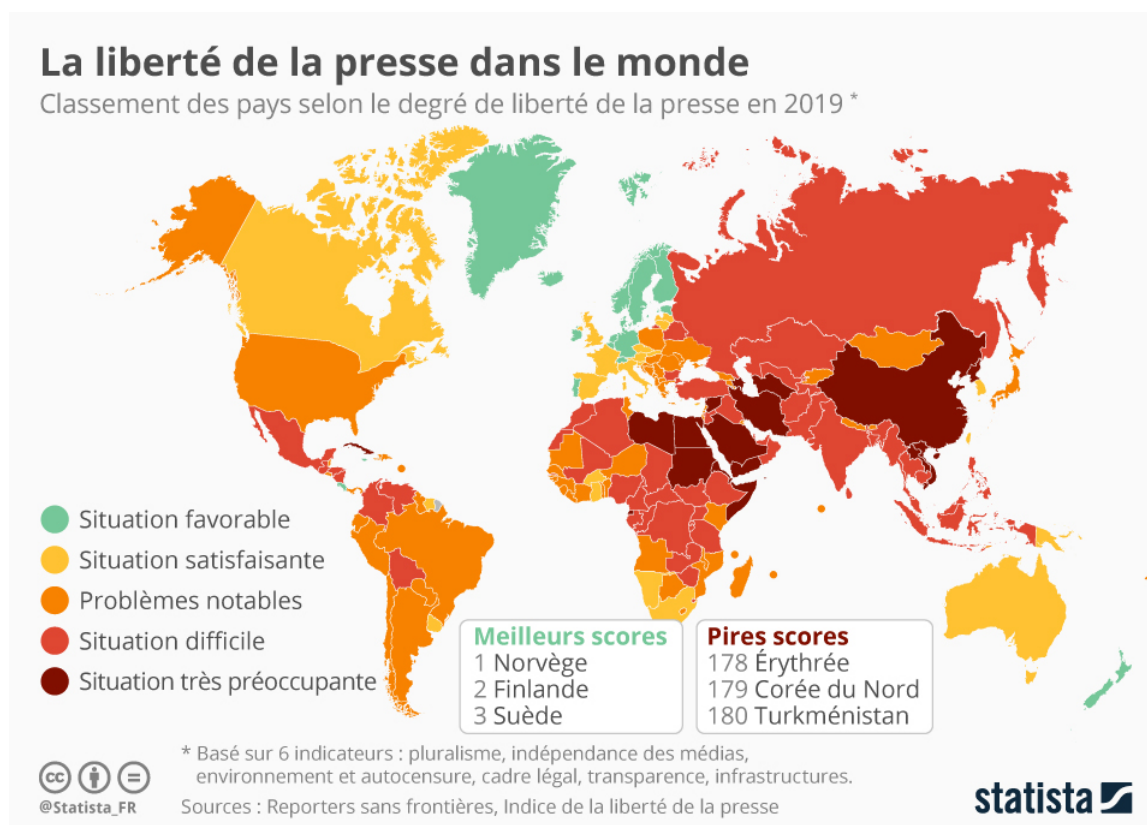


FIGURE 3 – Niveau de liberté de la presse selon le pays

Si au début des années 2010, les pays souffrant d'une dictature de l'opinion pouvaient compter sur les réseaux sociaux pour assurer un contre pouvoir, aujourd'hui ce sont des outils maîtrisés par les états qui assurent même une prolongation des systèmes historiques de surveillance. Par exemple, à Hong Kong [1], publier sur un réseau peut désormais être considéré comme un crime : demander la démocratie ou l'autonomie de Hong Kong, agiter un drapeau britannique ou américain, prôner des idées « étrangères », publier sur un réseau social... peut être considéré comme un crime. En Egypte, alors qu'en 2013 Facebook était utilisé par les militants pour s'organiser face au pouvoir dictatorial en position, depuis 2018 une loi autorise l'état à une surveillance très étroite des utilisateurs. La Russie s'est servit à de nombreuses reprises des réseaux pour polariser l'opinion et favoriser les élections via la création de faux comptes et la diffusion de faux contenus. Les réseaux sont très permissifs dès lors qu'il s'agit de l'état, aujourd'hui ils s'apparentent plus à des outils puissants au service des gouvernements plutôt qu'au service de la liberté d'expression. Finalement il n'y a plus trop de différences entre liberté de la presse et liberté de parole sur Internet.

2.2 La question légale

Le fait de donner le droit à tous de s'exprimer a engendré ces dernières années des déferlements de haine, des vagues de cyber-harcèlement, une violence exacerbée par l'impression d'anonymat. On peut récemment citer deux affaires emblématiques de la question en France.

L'affaire Mila est une affaire médiatique et judiciaire française qui commence en janvier 2020 lorsque Mila, une adolescente de 16 ans, blasphème l'islam de manière virulente sur Instagram. Ses propos sont considérés par certains internautes comme dénigrants et elle est l'objet de dizaines de milliers de messages homophobes, misogynes et d'appels à la violence et au meurtre [2]. L'adolescente est contrainte de changer de déménagement, suivre des cours par correspondance et fuir la sphère publique pour assurer sa sécurité et retrouver un semblant d'équilibre. En France, le blasphème est légal, quoi qu'on pense de ses propos, cette jeune fille avait le droit de les tenir. Certaines des personnes l'ayant menacé de mort ont été conduits en justice, mais ces mesures sont trop marginales pour empêcher qu'une situation similaire se reproduise.

Encore plus récemment, l'enseignant d'histoire Samuel Paty a été assassiné par un terroriste islamiste. Il avait donné quelques jours auparavant un cours portant sur la liberté d'expression dans lequel il présentait deux caricatures de Mahomet issues du journal Charlie Hebdo qui avait lui-même connu un attentat en 2015. Un parent d'élève accuse l'enseignant d'avoir demandé « à toute la classe que les élèves musulmans lèvent la main » alors même que sa fille était absente du cours. Samuel Paty avait indiqué aux élèves pensant que le contenu pourrait les heurter de détourner le regard quelques secondes. S'en est suivi un limogeage de l'enseignant sur les réseaux sociaux, fortement relayé sur WhatsApp est devenue très virale dans les milieux musulmans. Abdoullakh Anzorov l'auteur des crimes, tombe sur la polémique et décide un jour de venir à la sortie du collège de l'enseignant pour l'assassiner. Un élève lui permettrait de l'identifier contre quelques centaines d'euros, l'enseignant sera égorgé peu de temps après la photo de sa tête découpée mise sur le compte Twitter du criminel. Il sera neutralisé mortellement par des policiers quelques heures après.

C'est la prolongation de cette polémique sur les réseaux qui a déclenché le drame, le parent n'ayant même pas soupçonné une seule seconde sa fille de mentir, n'ayant pas désiré s'entretenir avec l'enseignant avant de le limoger sur internet. En ce sens, la liberté de parole sur les réseaux sociaux est aussi une de ses caractéristiques très dangereuse pouvant mener à du harcèlement voire à un assassinat.

Afin de limiter l'ampleur du cyber-harcèlement, le projet de loi Avia est né en 2018 [6], était destinée à retirer des contenus terroristes et pédopornographiques de n'importe quel site et les contenus haineux et pornographiques sous 24 h des principaux réseaux sociaux. Des personnalités politiques, un grand nombre d'organisations et des juristes critiquent la loi, qu'ils présentent comme un danger pour la liberté d'expression, notamment en raison de la possibilité que des décisions de retrait de contenus soient prises par un opérateur privé sans intervention du juge judiciaire. La loi est finalement largement censurée par le conseil constitutionnel [3]. Cette loi rend bien compte de la difficulté à introduire un cadre légal pour le contrôle de la parole sur les réseaux.

Conclusion

En effet, on a montré que sans contrôle, on s'expose à des problèmes de sécurité très importants. Mais trop de contrôle inhiberait la liberté d'expression. La mouvance *#metoo* a montré une vraie libération de la parole concernant le harcèlement sexuel constant et plus récemment divulguant des affaires de viol et d'inceste. On ne peut que encourager cette initiative et désirer qu'elle devienne durable et virale. Il est donc normal et nécessaire que les réseaux contrôlent leurs contenus, cependant il faut impérativement un cadre légal suffisamment précis pour détecter des suppressions/suspensions de contenus s'apparentant à une censure ou une orientation politique.