

Auditions du 15 novembre 2021

Twitter

Audrey Herblin-Stoop, Directrice des affaires publiques de Twitter
M. Francois Lesage, Directeur de la communication

« On ne résout pas les problèmes sociaux en accusant la technologie » : Internet est un outil, qui peut être aussi merveilleux, dans des mains bien intentionnées, que nuisible quand il est asservi à de mauvaises intentions. Mais ce n'est pas aux plateformes de réseaux sociaux de porter toute la faute des manques de la société.

Fort de la conviction qu'Internet peut-être une source de progrès, et que la plateforme Twitter facilite l'information des citoyens lors des moments de crises, l'entreprise veut développer la plus grande transparence vis-à-vis de ses utilisateurs.

Ainsi, en parallèle d'un **fil d'actualité à régulation algorithmique classique** (pondéré par la popularité de l'historique des tweets et des likes de l'utilisateur, avec mise en avant des contenus aimés par mes abonnements), Twitter rend possible d'opter pour un **fil d'actualité antéchronologique** pur.

Ce système était au départ le seul existant, et beaucoup d'utilisateurs y sont encore très attachés.

Le CSA a montré qu'il n'y a pas d'effet bulle sur Twitter.

Pourquoi certains comptes tenant des propos faux sont-ils labellisés ?

La **labellisation des comptes** n'est pas une validation des informations que ces comptes diffusent, ni un moyen de donner un surcroît de visibilité, mais un simple élément contextuel qui garantit que la personne qui tient le compte est bien celle qu'elle prétend être.

Aujourd'hui, sur les réseaux sociaux, la popularité donne plus de droit à la visibilité que la compétence. Pourriez-vous valider l'augmentation de visibilité de certains comptes à certains moments de crise ?

Ce système a été mis en place pendant le Covid, à travers un *feed covid* rassemblant des comptes sources faisant autorité. Mais il serait difficile de le généraliser : sur la plateforme, tous les utilisateurs doivent avoir les mêmes droits et devoirs. On sait par expérience que la renommée, l'expertise ne garantissent pas le respect des conditions générales d'utilisation (CGU).

Quel traitement réservez-vous aux *fake news* ?

Twitter se met à distance du terme de *fake news*, vague et sans véritable base juridique.

Les contenus trompeurs avec l'intention de provoquer un tort sont mis hors ligne. En revanche, Twitter n'a pas vocation à éliminer les contenus simplement erronés.

Les boutons de signalement sur les contenus trompeurs sont réservés aux périodes électorales. En dehors d'un danger électoral, même si des contenus sont faux (par exemple, si L'EU Disinfo Lab repère une fausse nouvelle sur le Covid), il n'est pas possible de le signaler ; Twitter en revanche peut le repérer et le supprimer.

La ligne de l'entreprise est claire : bien que certains contenus dérogent aux lois de Twitter, il peut être important de les conserver sur la plateforme à titre exceptionnel, parce que leur auteur occupent une place dans la société. Cela peut donc être utile que l'utilisateur puisse en prendre connaissance s'il le souhaite : ces contenus sont dits *newsworthy*.

Il existe 3 sous-exceptions cependant au sein de cette exception qui ne sont pas conservés : les contenus de pédopornographie, le terrorisme, et les appels à la violence.

Les membres de la Commission font remarquer que même parmi ces sous-exceptions, il y a parfois des manquements comme le tweet d'appel à la violence de l'ancien premier ministre polynésien, Mahathir Mohamad qui n'a été retiré que sur demande de Cédric O.)

Tout viol de ces règles entraîne suspension du compte. La liberté d'expression n'en sera pas bafouée, puisque Twitter n'est pas la seule plateforme où l'on puisse s'exprimer.

Twitter estime que sa responsabilité est dans le contenu qu'il met en avant, et non pas dans le contenu qu'il laisse voir à l'utilisateur s'il l'a cherché et donc choisi.

Comment limiter la viralité de certains contenus ?

Certains tweets sont empêchés de retweet, et l'utilisateur ne peut y accéder qu'en cliquant sur une bannière l'avertissant que le contenu auquel il s'apprête à accéder est susceptible d'enfreindre les règles de la communauté.

Twitter a également travaillé à éduquer les comportements de ses utilisateurs, en rendant les règles parfaitement intelligibles. Quand un compte est gelé ou un tweet refusé, un message pop-up explique quelle règle le contenu/compte a violé.

Transparence et explicabilité algorithmique

La transparence algorithmique est fondamentale ; il est nécessaire par exemple d'expliquer sur quel jeu de données a été entraîné l'algorithme.

Mais rendre public ne signifie pas que le grand public pourra s'approprié ces informations, souvent complexes, réservées à un public averti. Aussi faut-il aller plus loin, et expliquer les enjeux fondamentaux de ces algorithmes et la logique qui y préside.

L'exception journalistique du DSA

Le DSA prévoit une exemption pour les contenus journalistiques. Twitter y est opposé, car cela crée une inégalité entre les utilisateurs.

Quelle politique d'Éducation aux médias (EMI) chez Twitter ?

Il y a une collaboration avec le CLEMI et l'UNESCO:

- CLEMI: création de contenus et de ressources, actions menées pour la semaine de la presse (ateliers, etc.)
- UNESCO: guide à destination des professeurs et éducateurs actualisé très fréquemment (éléments pratiques en matière de sécurité en ligne et de harcèlement en ligne).

Quelle politique de lutte contre les ingérences étrangères ?

Twitter travaille à l'identification des comptes fautifs, à stopper les bots, etc. À partir de 2018, Twitter a commencé à rendre public des corpus de données pour la recherche.

La menace de l'ingérence étrangère est présente, mais la menace domestique est également très forte.

Y a-t-il un "Election Integrity Project" en France en 2022 en cours ?

Aucun projet n'est encore monté, tout est ouvert à ce stade.

On peut envisager une collaboration avec les organes de presse et les fact-checkers, sur le modèle américain.