

Haine en ligne et régulation des Réseaux sociaux



Point de départ : projection de la Une du Midi Libre du 21 octobre 2020

faire parler les élèves : quel rôle des réseaux sociaux, avant, pendant et après l'assassinat de S. Paty ?

Introduction

Marlène Schiappa, la ministre chargée de la citoyenneté auprès du ministère de l'Intérieur, en réunion avec les patrons des grands réseaux sociaux Twitter, Facebook, Instagram, Snapchat, Tik Tok, Youtube pour la France : **lieu de la radicalisation = téléphone**

<https://www.publicsenat.fr/article/parlementaire/haine-en-ligne-ce-que-dit-la-loi-185155>

1/ Les réseaux sociaux sont-ils, par essence des vecteurs de haine ?

<https://www.youtube.com/watch?v=HXP2HOYtS20>

video mise en ligne par L'Express le 26/10/2020 : L'assassinat de Samuel Paty repose la question de la responsabilité des plateformes.

➤ Réseaux sociaux = résonance et amplification

Modèle économique = retenir l'utilisateur, capter l'attention (=> pubs ciblées) en permettant de réagir, commenter, partager

Augmentent l'exposition des contenus : liker, retweeter

Donnent une plus grande **viralité** aux contenus qui font réagir et donc contenus haineux

définition de viralité selon les élèves

La **viralité** dans les médias est la diffusion rapide et imprévisible d'un contenu sur internet. La **viralité** en médecine est un ancien terme de médecine qui **désigne** le fait d'être contagieux.

source : wiktionnaire

➤ Filtrer, bloquer en amont ?

Europe matin du 20/10/2020 : Haine en ligne, terrorisme : les réseaux sociaux en font-ils assez ?

<https://www.europe1.fr/technologies/haine-en-ligne-terrorisme-les-reseaux-sociaux-en-font-ils-assez-3999710>

Sur le papier, les réseaux sociaux font de plus en plus d'efforts pour filtrer au mieux les contenus haineux, violents et choquants. Twitter et Facebook publient régulièrement des **rapports sur leur action de modération**. On sait donc que l'an dernier, par exemple, Twitter a suspendu 270.000 comptes pour incitation à la haine et quelque 180.000 autres pour apologie du terrorisme. Chez Facebook, on fait savoir que 15 millions de contenus liés d'une manière ou d'une autre au terrorisme ont été sanctionnés depuis le début de l'année. Et 99% d'entre eux l'ont été **avant même d'avoir été signalés par la communauté**.

Le problème, c'est le 1% restant, l'angle mort, car il suffit de quelques minutes pour qu'un contenu haineux produise son effet.

définition de modération selon les élèves

La **modération** est l'action de contrôler et filtrer les contenus sur le web (blogs, réseaux sociaux). Elle peut être automatique, effectuée par des machines grâce à des algorithmes, ou humaine.

Algorithmes imparfaits : certains contenus ne sont pas identifiés comme haineux (ex vidéo du père d'élève appelant à la mobilisation contre Samuel Paty)

➤ Et en aval, après publication ?

Outils à mettre au point pour **ralentir le flux**, diminuer la viralité // liberté d'expression

Plus d'un million de messages publiés chaque seconde sur Facebook. Plus de 400 heures de vidéo diffusées sur YouTube chaque minute... un contrôle strict des contenus paraît illusoire. Le contrôle dépend beaucoup de la vigilance des utilisateurs eux-mêmes. Mais les plateformes sont régulièrement mises en cause pour ne pas agir efficacement contre les contenus haineux qui leur sont signalés. Les **derniers chiffres de la commission européenne**, publiés fin juin 2020, indiquaient notamment que sur 1 400 signalements envoyés à Twitter, seuls 36 % ont été retirés. De son côté, YouTube avait retiré 80 % des contenus signalés, Facebook 87 %.

Réponse **politique** = légiférer, Thierry Breton, commissaire européen et **projet Digital services Act** : responsabiliser les plateformes, obliger les auteurs de comptes très suivis à s'identifier mais limite ex : Trump

2/ Comment agir, quelle régulation ?

➤ Statut des réseaux sociaux :

statut **d'hébergeurs**, ce qui veut dire que la loi - en principe - leur demande d'**intervenir sur signalement**. Ils ne sont pas éditeurs, c'est-à-dire qu'ils seraient responsables dès la première seconde de publication, comme peut l'être par exemple franceinfo en tant que média.

➤ Ce que dit la loi

○ Loi sur la liberté de la presse de 1881

La loi de 1881 sur la liberté de la presse et la liberté d'expression est le **premier texte à réprimer notamment la diffamation et l'injure**. Ces délits peuvent faire encourir à leurs auteurs une amende de 12 000 euros pour diffamation publique et d'un an de prison et de 45 000 € d'amende si les caractères racistes, antisémites ou sexistes sont retenus. Les délits d'incitation à la haine, à la violence ou à la discrimination raciale sont eux passibles d'un an d'emprisonnement et de 45 000 euros d'amende.

Ce **texte fondateur** qui se veut le **garant de l'équilibre délicat entre liberté d'expression et répression des abus**, a déjà été remis en cause avec l'essor d'internet.

Internet > web > prolifération de contenus et d'auteurs, tout le monde peut publier de façon plus ou moins anonyme...

Sur le web : médias en ligne = presse = responsabilité éditoriale du contenu

Réseaux sociaux = hébergeurs, responsabilité différente

○ loi « Confiance dans l'économie numérique » en 2004

= **protection de la vie privée dans le secteur des communications électroniques**.

Ce texte permet notamment à toute personne victime de propos haineux de demander **l'identification d'une personne via la requête d'un juge**. L'identification procède généralement d'une enquête de police qui recherche l'auteur du délit via son adresse IP ou en effectuant une demande auprès du fournisseur d'accès à internet (FAI). Cette loi impose également **aux hébergeurs de retirer les contenus « manifestement illicites »** dont ils ont connaissance. responsabilité de l'hébergeur mais manque de réactivité et régulation

○ La loi antiterroriste de 2014

= fait passer le délit d'apologie du terrorisme de la loi sur la presse au code pénal.

○ Loi Avia juin 2020

La mesure phare de ce texte fixait aux plateformes un **délai de 24 heures, après notification par une ou plusieurs personnes, pour supprimer tout contenu manifestement illicite** : les incitations à la haine, les injures à caractère raciste ou anti-religieuses. Un délai jugé bien trop court par le Conseil constitutionnel qui avait été saisi par les sénateurs. Le Conseil constitutionnel avait estimé que le dispositif de signalement par les internautes, pouvait « inciter les opérateurs de plateforme en ligne à retirer les contenus qui leur sont signalés, qu'ils soient ou non manifestement illicites », au risque d'attenter à la liberté d'expression.

Pour Thomas Rohmer, président de [l'association Open](https://www.ouest-france.fr/europe/france/contre-la-haine-en-ligne-quelle-regulation-sur-les-reseaux-sociaux-7021454) qui promeut notamment l'éducation au numérique. « **Retirer un contenu en une heure n'est pas raisonnable. Le suspendre temporairement, avant d'en faire un examen approfondi, serait plus habile** »,

<https://www.ouest-france.fr/europe/france/contre-la-haine-en-ligne-quelle-regulation-sur-les-reseaux-sociaux-7021454>

<https://www.publicsenat.fr/article/parlementaire/haine-en-ligne-ce-que-dit-la-loi-185155>

➤ En pratique : Les instances

○ Plateforme PHAROS : rôle et limites

Les élèves font une recherche sur **Pharos**

Création/ Signification du nom / Rôle / limites

Créée en 2009, [Pharos](#) (Plateforme d'harmonisation, d'analyse, de recoupement et d'orientation des signalements) permet de signaler aux services de police les contenus illicites. La plateforme est intégrée à l'Office central de lutte contre la criminalité liée aux technologies de l'information et de la communication (OCLCTIC). La question des moyens alloués à ce dispositif peut aujourd'hui se poser puisque ce service comptait, en 2018, seulement 27 policiers et gendarmes spécialisés dans la cybercriminalité pour 163 000 signalements, soit plus de 3 000 chaque semaine, comme le souligne [20 minutes](#).

A signaler :

- Pédophilie et pédopornographie
- Expression du racisme, de l'antisémitisme et de la xénophobie
- Incitation à la haine raciale, ethnique et religieuse
- Terrorisme et apologie du terrorisme
- Escroquerie et arnaque financières utilisant internet

Concrètement, les **policiers et gendarmes** qui sont affectés à cette structure dépendant de l'Office central de lutte contre la criminalité liée aux technologies de l'information et de la communication (OCLCTIC), et donc de la police nationale **consultent les contenus, vérifient leur caractère illicite et décident de les transférer ou non à un service compétent. En revanche, ils ne sont pas responsables de l'utilisation qui est faite ensuite des informations transmises, par exemple à la Direction générale de la sécurité intérieure (DGSI) dans le cadre de contenus liés au terrorisme.** Si erreur d'appréciation il y a eu, celle-ci se situerait donc plutôt du côté des policiers de la DGSI qui ont reçu les informations transmises par Pharos.

La suspension des comptes ou le filtrage des réseaux sociaux n'est pas un objectif absolu partagé par les forces de l'ordre, particulièrement en matière de lutte antiterroriste. C'est le plus souvent **contre-productif**, puisqu'il faut alors repartir de zéro pour identifier les nouveaux comptes que la personne ira certainement ouvrir. « Les terroristes font des erreurs, notamment sur les réseaux sociaux, et supprimer leurs comptes revient à se priver d'un moyen de surveillance assez facile à mettre en œuvre et peu coûteux », nous confie une source policière spécialisée dans le cyber. « Cela n'empêche pas les plateformes de réagir très rapidement lorsqu'un événement se produit pour supprimer des photos ou des vidéos d'actes de terrorisme », poursuit-il. Il y a toutefois de gros progrès à faire. Si [Facebook avait coupé la vidéo en direct de l'attentat de Christchurch en seulement 17 minutes](#), Twitter a mis au moins une heure à couper l'accès à la photo de la tête de Samuel Paty gisant sur le bitume.

<https://www.lesechos.fr/tech-medias/medias/haine-en-ligne-la-plateforme-de-signalement-pharos-en-cinq-questions-1257316>

https://www.lepoint.fr/high-tech-internet/pharos-une-plateforme-simple-pour-des-problemes-compliques-21-10-2020-2397406_47.php#

[https://www.francetvinfo.fr/faits-divers/terrorisme/enseignant-decapite-dans-les-yvelines/moderation-sur-les-reseaux-sociaux-il-ne-faudrait-pas-deleguer-l-analyse-des-contenus-aux-plateformes-preconise-nicolas-arpagian_4148047.html#xtor=AL-79-\[article\]-\[connexe\]](https://www.francetvinfo.fr/faits-divers/terrorisme/enseignant-decapite-dans-les-yvelines/moderation-sur-les-reseaux-sociaux-il-ne-faudrait-pas-deleguer-l-analyse-des-contenus-aux-plateformes-preconise-nicolas-arpagian_4148047.html#xtor=AL-79-[article]-[connexe])

<https://www.interieur.gouv.fr/A-votre-service/Ma-securite/Conseils-pratiques/Sur-internet/Signaler-un-contenu-suspect-ou-illicite-avec-PHAROS>

○ Observatoire de la haine en ligne, créé en juillet 2020 : fonctionnement et missions composé d'opérateurs, d'associations, d'administrations et de chercheurs concernés par la lutte et la prévention de la haine en ligne.

Analyser pour mieux comprendre et mieux lutter : scruter le web, quantifier, voir ce qui se fait ailleurs

Prévenir, éduquer

<https://www.csa.fr/Informer/Toutes-les-actualites/Actualites/Observatoire-de-la-haine-en-ligne-analyser-pour-mieux-lutter>

Conclusion :

Pistes juridiques

- instauration d'un délit « de mise en danger par la publication de données personnelles ». À l'Assemblée nationale, ce mardi 20 octobre 2020, le Premier ministre s'est dit favorable à l'instauration d'un délit « de mise en danger par la publication de données personnelles ».
- **Augmentation du nombre de magistrats** totalement dédiés aux traitements des contenus haineux en ligne
- sortir les délits d'incitation à la haine de loi 1881 afin de permettre des **procédures rapides**.
- un texte à l'échelle européenne, le **Digital Services Act**, plus contraignant pour les plateformes des réseaux sociaux avec obligations graduées en matière de signalement et de retrait de jugement jugés nocifs (+ sécurité numérique, concurrence...) sera présenté en décembre 2020. consultation de juin à septembre 2020.

Eloge de la lenteur : le sociologue Dominique Boullier parle de « réchauffement médiatique » et plaide pour des **quotas** sur le nombre de contenus qu'on pourrait partager quotidiennement par exemple*.

Action individuelle : signaler/ne pas relayer

C'est notre responsabilité à chacun, toujours prendre le temps d'analyser, de réfléchir avant de partager du contenu

* Tesquet, Olivier. *Combattre la haine en ligne* in Télérama 3694, 28/10/2020. p. 22-24

Un procès en cours contre Twitter

Hasard de calendrier, lundi 19 octobre, plusieurs associations dont l'Union des étudiants juifs de France et SOS racisme assignaient Twitter, estimant que le réseau social manquait de manière ancienne et persistante à ses obligations en matière de modération des contenus. Les associations avaient lancé un testing : plus de 1 100 signalements avaient été envoyés à Twitter, allant de sale juif à bougnoule et seulement 12 % avaient été retirés au bout de trois à cinq jours. La justice française a ordonné une médiation. Prochaine audience le 1^{er} décembre 2020.