

Chronique

DROIT DES NOUVELLES TECHNOLOGIES

RÉGULATION DES RÉSEAUX SOCIAUX : OÙ EN EST-ON ?

PAR FABRICE PERBOST* ET ALEXANDRE EUVERTE**

À l'ère du numérique, les réseaux sociaux constituent le nouveau vecteur d'une liberté d'expression sans commune mesure. D'abord parce que toute forme d'expression est aujourd'hui affranchie de l'idée de frontière, chacun ayant la possibilité de diffuser des contenus à une audience potentiellement illimitée, et chacun ayant accès aux contenus publiés par les autres utilisateurs, en direct et sans restriction. Ensuite parce que la dimension virtuelle d'internet permet de se délier des limites imposées par le réel (facilité et rapidité d'utilisation, possibilité de rester anonyme, sentiment d'impunité, etc.)

En 2015, un peu plus de 2 milliards d'individus se sont connectés sur les réseaux sociaux¹. Ceux-ci permettent principalement la communication de contenus entre amis (Facebook, 1,55 milliards d'utilisateurs actifs par mois²), les connexions professionnelles (LinkedIn, 100 millions d'utilisateurs actifs par mois³), le *microblogging* (Twitter, 310 millions d'utilisateurs actifs par mois⁴), le *dating* (Tinder⁵) ou encore la diffusion de vidéos en direct (Periscope, plus de 10 millions de comptes⁶).

Il est indéniable que les réseaux sociaux forment aujourd'hui un nouveau tissu social globalisé, qui a vu naître la formation de communautés regroupées en fonction des centres d'intérêts, convictions ou affinités de leurs membres. Il serait donc régressif de nier leur propension à faciliter l'accès à l'information et l'interaction entre les internautes.

Néanmoins, cette immense toile n'est pas exempte de dérives, et y naviguer sans restriction peut s'avérer dangereux, en particulier pour les individus les plus vulnérables tels que les mineurs. Les risques d'addiction ou d'endoctrinement, la propagation de contenus obscènes, violents ou haineux constituent indéniablement la face sombre des réseaux sociaux favorisant, paradoxalement, la généralisation et la banalisation des comportements antisociaux.

Prenons par exemple Periscope, application pour mobiles proposant à ses utilisateurs de diffuser publiquement (ou non) en direct ce qu'ils sont en train de filmer. Sur son site, l'application est décrite comme un concept innovant permettant d'explorer le monde « à travers le regard des autres » et, par exemple, de « voir ce que voit un manifestant en Ukraine », ou encore « regarder le lever du soleil depuis une montagne en Cappadoce »⁷. L'intention est éminemment louable et la majorité des usages qui en sont faits ne sont pas répréhensibles.

Mais en pratique, Periscope a aussi été le théâtre de plusieurs scandales récents, les derniers en date étant :

* Fabrice Perbost est avocat associé au sein du cabinet Harlay Avocats.

** Alexandre Euverte est juriste au sein du cabinet Harlay Avocats.

1 - <http://www.novius.com/medias-sociaux-chiffres-2015.html>.

2 - <http://www.blogdumoderateur.com/50-chiffres-medias-sociaux-2016/>.

3 - <http://www.blogdumoderateur.com/50-chiffres-medias-sociaux-2016/>.

4 - <https://about.twitter.com/fr/company>.

5 - Tinder ne communique pas sur son nombre d'utilisateurs, mais revendique 26 millions de « matches »

par jour, c'est-à-dire de mises en relation entre deux personnes.

6 - <http://www.blogdumoderateur.com/periscope-chiffres>.

7 - <https://www.periscope.tv/about>.

- l'affaire « SFR » (mai 2016) : deux employés de SFR ont tenu des propos misogynes et antisémites à l'égard d'un client⁸ ;
- un suicide en direct (mai 2016) : une jeune femme de 19 ans se donne la mort en se jetant sous les rails du RER sous les yeux de plus d'un millier d'utilisateurs⁹ ;
- une agression retransmise en direct à Bordeaux (avril 2016) : deux mineurs se lancent le défi de frapper gratuitement des passants, afin d'atteindre un maximum de vues¹⁰ ;
- l'affaire « Serge Aurier » (février 2016) : dans une vidéo publique, le footballeur a proféré à l'encontre de son entraîneur, Laurent Blanc, des insultes à caractère homophobe¹¹.

En outre, les vidéos mettant en ligne ce type de contenus – par ailleurs loin d'être des cas isolés – sont généralement accompagnées d'un déferlement de commentaires haineux et provocants, créant un effet d'emballement et conduisant inéluctablement leurs auteurs au passage à l'acte, la surenchère devenant loi. Dès lors que l'on veut appréhender le contenu de ce type de réseau social, il convient donc également de prendre en compte les commentaires des internautes qui pour certains suivent passivement ce qui est diffusé et pour d'autres encouragent les comportements.

Souvent, le réseau lui-même, formé par la communauté des internautes, la presse et tous types de média, se mue en tribunal virtuel géant, se chargeant de mettre en lumière les auteurs de ces actes répréhensibles, qui sont sanctionnés *a posteriori* par les personnes appropriées (mise à pied de Serge Aurier, licenciement des employés SFR, mises en examen des mineurs dans le cas de l'agression de Bordeaux, ouverture d'une enquête dans l'affaire du suicide de la jeune femme).

Indépendamment des suites juridiques qui sont ou seront données à ces agissements, quels sont les outils permettant de réguler les contenus diffusés ?

La régulation de toutes les formes de publications sur les réseaux sociaux passe, dans un

premier temps, par le recours au droit commun (I), mais également par une autorégulation des acteurs du marché dont l'efficacité reste, à ce jour, incertaine (II).

I - L'appréhension des réseaux sociaux par le droit commun

La liberté d'expression est un fondement de nos démocraties, protégée constitutionnellement par l'article 11 de la Déclaration des Droits de l'Homme et du Citoyen de 1789 : « *La libre communication des pensées et des opinions est un des droits les plus précieux de l'Homme : tout Citoyen peut donc parler, écrire, imprimer librement, sauf à répondre de l'abus de cette liberté dans les cas déterminés par la Loi* ». C'est dans l'abus, strictement circonscrit par la loi, que se trouve la source des limites apportées à la liberté d'expression.

S'agissant des réseaux sociaux en particulier, le droit commun adapté à internet semble, pour l'heure, offrir tous les outils nécessaires à la régulation des propos qui y sont tenus par les utilisateurs.

A - La sanction des contenus illicites par les délits de presse (loi du 29 juillet 1881)

Cette liberté fondamentale qu'est la liberté d'expression trouve son prolongement dans la loi du 29 juillet 1881 sur la liberté de la presse, qui s'est adaptée aux technologies de l'information et à la communication de masse incarnée aujourd'hui par les réseaux sociaux.

Cette loi réprime notamment la provocation à commettre des crimes ou des délits, la diffamation ou encore l'injure publique. Le législateur a prévu que la loi s'applique lorsque ces délits sont commis par tout moyen, y compris par voie électronique¹². Les réseaux sociaux ne sauraient donc y échapper. La jurisprudence en la matière est d'ailleurs foisonnante, notamment pour les infractions commises sur les réseaux sociaux.

Ainsi, dans un jugement du 18 mars 2015¹³, le tribunal de grande instance de Paris a condamné l'humoriste Dieudonné à deux ans d'emprisonnement avec sursis et à 1 000 euros d'amende pour apologie du terrorisme commis

8 - Le Parisien, *Deux salariés de SFR mis à pied après des propos antisémites sur Periscope*, 31 mai 2016.

9 - LCI/TF1, *Suicide en direct sur Periscope : « une scène traumatisante pour les utilisateurs »*, 17 mai 2016.

10 - Le Monde, *Bordeaux : une agression diffusée en direct sur Periscope*, 29 avril 2016.

11 - Le Monde, *PSG : Serge Aurier mis à pied après ses propos sur Laurent Blanc*, 14 février 2016.

12 - Article 23 de la loi du 29 juillet 1881 sur la liberté de la presse.

13 - TGI Paris, 16e ch. corr., 18 mars 2015 : JurisData n°2015-005323.

au moyen d'un service de communication au public en ligne. En l'espèce, le prévenu avait publié sur sa page Facebook officielle le message « *Sachez que ce soir, je me sens Charlie Coulibaly* » au lendemain des attentats de Charlie-Hebdo. Le tribunal a pris toute la mesure de la portée d'un tel contenu diffusé sur un réseau social tel que Facebook, estimant qu'« *il convient de rappeler qu'internet est un média utilisé de façon massive, abritant des forums de discussion qui libèrent une parole dénonçant les impérialismes américain et juif pour justifier le « djihad* ». En tenant, d'une part, dans le contexte décrit, des propos volontairement provocateurs faisant l'amalgame entre la liberté d'expression et des actes terroristes qu'il a ainsi contribué à banaliser, et en décidant, d'autre part, de les publier sur internet à destination d'un large public auprès duquel il entretient un sentiment d'hostilité à l'égard de la communauté juive, M. D., qui ne pouvait, ignorer l'impact de ses propos, s'est rendu coupable du délit d'apologie du terrorisme. »

Il est d'ailleurs à noter que le code pénal punit plus sévèrement ce délit lorsqu'il est commis par voie électronique : sept ans d'emprisonnement et 100 000 euros d'amende, au lieu de cinq ans d'emprisonnement et 75 000 euros d'amende¹⁴. Cette sévérité peut se comprendre dans la mesure où le dommage causé peut être bien plus important sur internet, dès lors qu'une audience potentiellement illimitée peut être captée par l'auteur.

Dans une autre affaire rendue en mars 2015¹⁵, un élu du Front National s'est rendu coupable du délit de provocation à la haine ou à la violence à l'égard d'une personne ou d'un groupe de personnes à raison de d'une religion déterminée¹⁶ (en l'occurrence l'islam). Cet élu local avait publié sur son mur Facebook des propos virulents créant une assimilation entre musulmans et délinquance et insécurité dans la ville de Nîmes.

Selon cette jurisprudence établie, l'appréhension des contenus illicites ne pose pas de difficulté particulière concernant les réseaux sociaux.

B - La liberté d'expression sur les réseaux sociaux limitée par le caractère public des contenus

L'encadrement de la liberté d'expression n'est pas le même selon que les contenus litigieux sont dirigés vers un nombre limité de personnes (cercle privé), ou selon qu'ils sont publiquement diffusés¹⁷. Ceci étant, force est de constater que, compte tenu de la nature des réseaux sociaux, il existe, de fait, une certaine porosité entre les sphères privées et publiques sur ces réseaux¹⁸.

Les juridictions prud'homales sont un terrain privilégié de ce débat¹⁹. Les risques liés à l'utilisation des réseaux sociaux par le salarié sont, entre autres, l'atteinte à la réputation (de l'employeur ou de l'entreprise), à l'image ou à l'honneur, le dénigrement, la violation de son obligation de confidentialité, de loyauté vis-à-vis de son employeur, la diffamation, la violation du devoir de neutralité dans le cas d'un agent de la fonction publique, etc.

La liberté d'expression du salarié n'est en principe pas limitée lorsque la diffusion du contenu est limitée à un cercle restreint (par exemple, ses « amis » Facebook). C'est le sens de la jurisprudence sociale depuis que le conseil des prud'hommes de Boulogne-Billancourt a condamné, en 2010, une salariée pour dénigrement et atteinte à la réputation de sa supérieure hiérarchique²⁰. Cette employée s'était en l'espèce livrée, sur la page Facebook de l'un de ses collègues, à une incitation à la rébellion contre la hiérarchie et au dénigrement de son employeur. Or, la page Facebook en question était accessible par les « amis » du titulaire du compte ainsi que leurs « amis ». Selon les juges, cette communication dépassait la sphère privée, la salariée ayant en conséquence abusé de son droit d'expression visé à l'article L. 1121-1 du code du travail.

Dans la même veine et sur le terrain de la loi du 29 juillet 1881, la chambre sociale de la Cour de cassation a précisé que le délit d'injure

14 - Article 421-2-5 du code pénal.

15 - Cass. crim., 17 mars 2015, n°13-87.922 JurisData n°2015-005832.

16 - Article 24, alinéa 6, de la loi du 29 juillet 1881.

17 - Emmanuel Netter, *La liberté d'expression sur les réseaux sociaux en droit français*, 26 octobre 2015.

18 - J. Le Clainche, *Licenciement pour des propos tenus sur Facebook ou les dangers de la porosité des sphères publique et privée* : RLDI janv. 2011, n°2208.

19 - Merav Griguer, *La gestion des risques liés à l'utilisation des nouveaux outils de communication*, Cahiers de droit de l'entreprise n°5, septembre 2010, prat. 25.

20 - Cons. prud'h. Boulogne-Billancourt, 19 nov. 2010, n°09/00343.

publique ne peut être caractérisé que si les propos en cause sont publiés sur un compte ouvert, c'est-à-dire qu'ils sont accessibles par des personnes qui n'ont pas été agréées par l'utilisateur du compte. La haute juridiction considère en conséquence que sur un compte fermé (à un nombre limité de personnes agréées), les individus en relation forment une « communauté d'intérêts », et les contenus qu'ils s'échangent ne peuvent être considérés comme des injures publiques au sens de la loi de 1881²¹.

Cette jurisprudence est désormais constante²². Dans un arrêt récent du 25 juin 2015, la cour d'appel de Paris a commencé par définir le réseau Facebook comme ayant pour objectif « de créer entre ses membres un maillage relationnel destiné à leur permettre de partager toutes sortes d'informations » qui « permet que les échanges s'effectuent librement via « le mur » de chacun des membres auquel toute personne peut accéder si son titulaire n'a pas apporté de restrictions ». Dans la droite ligne des arrêts précédents, elle confirme que l'auteur des propos litigieux aurait dû utiliser les fonctionnalités offertes par Facebook pour préserver la confidentialité de tels propos. En l'espèce, il s'agissait, comme en 2010, d'un cas de dénigrement de la part d'une salariée sur son « mur » Facebook, laissé en libre accès.

Tout dépend donc du paramétrage opéré par l'auteur du contenu, qui doit lui-même définir un espace privé au sein duquel sa liberté d'expression ne sera pas limitée par ses engagements professionnels. S'il ne le fait pas, les propos sont considérés être diffusés dans un espace public²³, ouvrant droit aux sanctions appropriées et permettant à ces propos d'être utilisés comme moyen de preuve valable pris à l'appui d'une sanction disciplinaire, comme un licenciement.

C - L'encadrement des réseaux sociaux par la loi pour la confiance dans l'économie numérique (« LCEN »)

Nous avons cité plus haut le principe fondamental de la liberté d'expression. Celle-ci trouve un prolongement essentiel dans la liberté de communication, notamment par la voie élec-

tronique. C'est ce que rappelle l'article 1^{er} de la LCEN, faisant écho à la loi du 30 septembre 1986 relative à la liberté de communication : « la communication au public par voie électronique est libre ». Et l'alinéa 2 d'en préciser les limites : « L'exercice de cette liberté ne peut être limité que dans la mesure requise, d'une part, par le respect de la dignité de la personne humaine, de la liberté et de la propriété d'autrui, du caractère pluraliste de l'expression des courants de pensée et d'opinion et, d'autre part, par la sauvegarde de l'ordre public, par les besoins de la défense nationale (...) ».

Pour rappel, les réseaux sociaux sont considérés comme des hébergeurs au sens de la LCEN, dès lors qu'ils n'ont aucun contrôle ni pouvoir éditorial sur les contenus²⁴. Si cette loi instaure un mécanisme de responsabilité limitée à leur égard, deux séries d'obligations leur incombent néanmoins :

- **obligations relatives aux données** : les hébergeurs et les fournisseurs d'accès à internet sont tenus de conserver « les données relatives à quiconque a contribué à la création du contenu des services dont ils sont prestataires »²⁵. Ils doivent communiquer ces données si l'autorité judiciaire en fait la demande²⁶.
- **obligations relatives au blocage et à la cessation de l'illicite** : les hébergeurs et les fournisseurs d'accès à internet doivent, « compte tenu de l'intérêt général attaché à la répression de l'apologie des crimes contre l'humanité, de la provocation à la commission d'actes de terrorisme et de leur apologie, de l'incitation à la haine raciale, à la haine à l'égard de personnes à raison de leur sexe, de leur orientation ou identité sexuelle ou de leur

24 - Voir par exemple, s'agissant du site de partage de vidéos en ligne Dailymotion : CA Paris, pôle 5, ch. 1, 2 déc. 2014, n°13/08052, TF1 et a. c/ Dailymotion et a. : JurisData n°2014-029711. La cour d'appel relève que « l'organisation des espaces personnels des utilisateurs du site, le postage, l'accessibilité et le retrait des vidéos s'effectuent par les utilisateurs eux-mêmes sous leur seule responsabilité, sans possibilité d'interférence de la SA Dailymotion ». Par conséquent, cette dernière doit être qualifiée d'hébergeur. Ce raisonnement se tient pour l'ensemble des grands réseaux sociaux, puisque ce sont les utilisateurs seulement qui postent leurs propres contributions.

25 - Article 6-II de la loi n° 2004-575 du 21 juin 2004 pour la confiance dans l'économie numérique.

26 - Les modalités de conservation et de communication des données sont fixées par le décret n° 2011-219 du 25 février 2011 relatif à la conservation et à la communication des données permettant d'identifier toute personne ayant contribué à la création d'un contenu mis en ligne.

21 - Cass. 1^{re} civ., 10 avr. 2013, n°11-19.530.

22 - Voir par exemple : Cons. prud'h. Longwy, 24 sept. 2012, n°12/00077 : JurisData n°2012-021777, ou CA Toulouse, 5 déc. 2013, n°12/00526.

23 - CA Besançon, 15 novembre 2011, n°10/02642 : JurisData n°2011-031655.

handicap ainsi que de la pornographie enfantine, de l'incitation à la violence, notamment l'incitation aux violences faites aux femmes, ainsi que des atteintes à la dignité humaine »²⁷, concourir à la lutte contre la diffusion de ces infractions prévues par la loi du 29 juillet 1881 et par le code pénal.

- Par ailleurs, ils doivent mettre en place un « dispositif facilement accessible et visible permettant à toute personne de porter à leur connaissance ce type de données », informer promptement les autorités dès lors que des contenus illicites leur sont signalés, et rendre publics les moyens employés à ces fins.
- Les hébergeurs et les fournisseurs d'accès à internet engagent par ailleurs leur responsabilité civile s'ils n'agissent pas promptement pour retirer les contenus manifestement illicites ou en rendre l'accès impossible²⁸.
- Une procédure de référé ou de requête est également prévue, afin que l'hébergeur ou à défaut le fournisseur d'accès à internet prenne les mesures nécessaires pour faire cesser un dommage causé par le contenu d'un de leurs services (procédure dite « référé LCEN »)²⁹.

L'ensemble de ces dispositions a pu trouver application devant les tribunaux, notamment en vue de la fermeture d'un site ouvertement négationniste sur le fondement du référé LCEN³⁰, ou encore pour ordonner au réseau social Twitter la communication des données d'identification des auteurs de tweets antisémites (#UnBonJuif ou #siJétaisNazi entre autres...)³¹ ainsi que la mise en place d'un dispositif plus efficace de signalement des contenus illicites (pour qu'il soit « facilement accessible et lisible », conformément aux termes de la LCEN)³².

Il résulte de ces décisions que les contenus diffusés sur les réseaux sociaux ne sauraient, en tant que « communication au public en ligne »,

échapper aux dispositions de la loi, et en particulier de la LCEN.

II - L'autorégulation des réseaux sociaux par la modération

A - Quelques exemples d'autorégulation

1 - Actuellement, Twitter propose à ses utilisateurs des formulaires à remplir en ligne permettant de signaler une violation. Plusieurs cas sont envisagés, parmi lesquels : usurpation d'identité, contrefaçon, harcèlement, ou encore signaler un cas d'automutilation³³. Un autre type de formulaire est destiné aux forces de l'ordre, afin que leur soient communiquées d'urgence les données d'identification d'un utilisateur. Au dernier semestre 2015, la France a adressé 707 demandes d'informations à Twitter concernant 866 comptes. 61% des requêtes ont été satisfaites³⁴. Dans ses conditions générales, le géant américain énumère ce qu'il considère être un « comportement inapproprié », et se réserve le droit de pouvoir temporairement bloquer le compte ou même le suspendre définitivement.

Par ailleurs, sur son application Periscope, Twitter a mis en place une adresse e-mail spécialement dédiée au signalement de vidéos illicites : safety@periscope.tv ainsi qu'un bouton générique qui ne requiert aucune précision sur la nature du signalement³⁵. Il peut s'agir tout aussi bien d'une violation de droit d'auteur que d'un cas plus grave nécessitant une intervention d'urgence des autorités.

2 - De son côté, Facebook cherche aussi à moderniser ses mécanismes de modération.

D'une part, en misant sur l'intelligence artificielle. Le 1^{er} juin 2016, le premier réseau social mondial a dévoilé son projet DeepText, un programme capable de comprendre, avec une précision « quasi-humaine », selon Facebook, le contenu de plusieurs milliers de posts par secondes, dans plus de 20 langues³⁶. L'objectif serait d'assurer une meilleure efficacité de la modération. Il sera à cet égard intéressant de vérifier si ce système d'intelligence artificielle, qui vient en renfort et facilite le travail des

27 - Article 6-I-7 de la loi n° 2004-575 du 21 juin 2004 pour la confiance dans l'économie numérique.

28 - Article 6-I-1 de la loi n° -2004-575 du 21 juin 2004 pour la confiance dans l'économie numérique.

29 - Article 6-I-8 de la loi n° 2004-575 du 21 juin 2004 pour la confiance dans l'économie numérique.

30 - Cass. 1^{re} civ., 19 juin 2008, n°07-12.244 ; AFA c/ UEJF : JurisData 2008-044403.

31 - Fabrice Perbost et Henri Alterman, *Twitter dans le viseur des juges français*, RJCom. Juillet/Août 2013.

32 - Fabrice Perbost, *La lutte contre les propos antisémites sur internet*, 21 juin 2014, RJCom. Juillet/Août 2014.

33 - <https://support.twitter.com/forms>.

34 - <https://transparency.twitter.com/information-requests/2015/jul-dec>.

35 - Numerama, *Le suicide en direct sur Periscope, ou l'urgence d'un bouton d'urgence !*, 11 mai 2016.

36 - <https://code.facebook.com/posts/181565595577955>.

équipes affectées à la gestion des signalements, est compatible avec les exigences de signalement posées par la LCEN. Tout dépendra de l'efficacité dudit dispositif, à savoir sa capacité réelle à comprendre le contexte des contenus.

D'autre part, Facebook a lancé, en juin 2016, un bouton anti-suicide intitulé « signaler un contenu de nature suicidaire »³⁷. Ce dispositif, qui existe déjà aux États-Unis depuis un an, permet de signaler les publications laissant penser que leur auteur pourrait avoir l'intention d'attenter à ses jours. Plusieurs options sont ouvertes, dès que l'on actionne ce bouton : entrer en contact avec la personne concernée, contacter une assistance dédiée, ou consulter un support.

B - Les limites des systèmes de modération mis en place

1 - Le problème de certains réseaux sociaux réside dans l'immédiateté de leurs contenus (Periscope), leur rapidité d'exécution ou leur caractère éphémère (Snapchat). Dans le cas de Periscope, signaler un cas d'automutilation en remplissant un formulaire en ligne n'aurait probablement pas empêché le suicide de la jeune femme. Les *Community Guidelines* de Periscope le reconnaissent dès leurs premières lignes : « Utiliser Periscope, c'est être dans l'instant. (...) Cette immédiateté encourage une participation directe et non filtrée à une histoire qui est en train de se dérouler »³⁸, même lorsque cette histoire est vouée à une fin tragique et profondément choquante pour ceux qui la suivent. Periscope indique dans ces mêmes *guidelines* se réserver la faculté de supprimer tout contenu en infraction sans préavis et à sa seule discrétion, en sus de supprimer le compte en cause si nécessaire. Mais avant d'y procéder, il faudra qu'un utilisateur signale tel contenu, avant d'attendre que l'équipe dédiée de l'entreprise traite la demande (c'est ainsi que cela fonctionne sur la majorité des réseaux sociaux), ce qui pose un réel problème de réactivité de la procédure de signalement.

Le 31 mai 2016, Periscope a annoncé mettre en place une nouvelle forme de modération s'agissant des commentaires³⁹. Selon ce nouveau dispositif, dit « démocratique », lorsqu'un

commentaire est signalé comme abusif, d'autres utilisateurs sont sélectionnés de façon aléatoire et devront juger si le commentaire est réellement abusif. Si tel est le cas, son auteur sera sanctionné et ne pourra plus commenter pendant une certaine durée ; durée qui sera décuplée en cas de récidive. Le concept, novateur, tient compte de l'immédiateté de Periscope. Néanmoins, il est permis de douter de son efficacité : les utilisateurs, souvent mineurs, prendront-ils la peine de signaler chaque commentaire abusif ? Seront-ils à même de remplir efficacement leur rôle de « jury populaire », en déterminant si un contenu est offensant ou non ? Et d'une manière plus générale, quelle est la légitimité de ces jurys d'utilisateurs ? Leurs décisions ne pourraient-elles pas être contestées devant les tribunaux ?

Pour l'heure, ce réseau social en pleine explosion regorge toujours autant de vidéos à caractère sexuel ou violent, mettant très souvent en scène des personnes mineures, sans parler des commentaires qui les accompagnent et les attisent.

2 - La faiblesse de ces mécanismes de modération est régulièrement dénoncée. Par exemple, du 31 mars au 10 mai 2016, dans le cadre d'un test de masse, il a été démontré que sur 586 contenus haineux (explicitement racistes, homophobes et antisémites) signalés sur YouTube, Facebook et Twitter, 4% ont été supprimés sur Twitter, 7% sur YouTube, et 34% sur Facebook. En conséquence, les associations SOS-Racisme, l'Union des étudiants juifs de France et SOS-Homophobie ont assigné ces trois géants du Web en référé⁴⁰.

De par leurs défaillances dans leur capacité à modérer leurs réseaux, ces entreprises se trouvent en violation constante de leurs propres conditions générales ainsi que de la loi (LCEN en particulier). L'enjeu concerne avant tout le dispositif de signalement obligatoire permettant de bloquer les contenus constitutifs d'infractions pénales et d'en sanctionner les auteurs le cas échéant. Nous avons déjà souligné, en son temps et dans ces mêmes colonnes, que bien que ces sociétés ne soient pas à l'origine des propos ou actes antisémites ou puissent être taxées d'antisémitisme, elles ont toutes, involontairement et/ou par complaisance ou neutralité coupable, de par leurs formidables outils, favo-

37 - <https://fr-fr.facebook.com/help/contact/305410456169423>.

38 - <https://www.periscope.tv/content>.

39 - Le Monde, *Facebook, Periscope : la modération évolue*, 8 juin 2016.

40 - Le Point, *Haine sur internet : YouTube, Facebook et Twitter assignés par l'UEJF et SOS-Racisme*, 15 mai 2016.

risé voire encouragé la parole antisémite, engageant par là même leur responsabilité⁴¹.

C - Vers une modération plus efficace ?

1 - Les initiatives mentionnées plus haut semblent toutefois aller dans le bon sens, et la période actuelle est marquée d'une volonté des acteurs du secteur à améliorer la performance de leur modération. En témoigne l'adoption d'un code de conduite relatif aux discours haineux en ligne⁴² par Facebook, Twitter, YouTube et Microsoft le 31 mai 2016. Dans le communiqué de presse, la Commission européenne a déclaré que « *Si l'application effective des dispositions qui criminalisent les discours haineux dépend de l'existence d'un système solide d'application de sanctions pénales aux auteurs de ce type de discours, elle doit être complétée par des actions visant à garantir que, dès réception d'un signalement valide, les intermédiaires en ligne et les plateformes de médias sociaux examinent rapidement, dans un délai approprié, les contenus signalés en tant que discours haineux en ligne* ». Le code de conduite comporte un engagement de la part des quatre entreprises américaines d'examiner les signalements dans un délai de 24 heures, et les retirer le cas échéant.

Ceci étant, cet objectif, louable, est *a priori* loin d'être atteint pour le moment, au vu des 4% de contenus haineux signalés et retirés sur Twitter par exemple.

2 - Au niveau européen, une proposition de révision de la directive Services de médias audiovisuels (Directive « SMA ») a été publiée le 25 mai 2016⁴³. Cette révision vise à faire entrer dans le champ de la réglementation les plateformes de partage de vidéos générées par les utilisateurs (« User Generated Content » ou « UGC »), à l'instar de YouTube ou Periscope⁴⁴. La proposition rappelle qu'en l'état actuel, la directive SMA ne s'applique pas aux UGC publiés sur ces plateformes de partage, puisque celles-ci n'ont, généralement, pas de pouvoir éditorial sur le contenu qu'elles stockent. Elles sont soumises à la directive n°2000/31 sur le

commerce électronique, transposée en droit français par la LCEN. À ce titre, elles n'ont pas d'obligation générale de surveillance à l'égard du contenu qu'elles hébergent⁴⁵, et bénéficient d'une exemption de responsabilité en vertu de leur statut d'hébergeur (à la condition de retirer promptement un contenu illicite dont elles auraient connaissance ou en rendre l'accès impossible)⁴⁶.

La révision de la directive SMA viendrait compléter ces dispositions, sans les remettre en cause, et notamment (article 28) obligerait ces plateformes de partage à prendre les mesures appropriées pour protéger les mineurs des contenus pouvant être néfastes pour leur développement physique, mental ou moral, et protéger tous les citoyens contre l'incitation à la violence ou à la haine directement à l'encontre d'un groupe de personne ou d'un membre d'un tel groupe défini par son sexe, sa race, sa couleur, sa religion, sa descendance ou son origine nationale ou ethnique. Ces mesures peuvent consister dans la définition et l'application des concepts d'incitation à la violence et à la haine dans les conditions générales des plateformes, un suivi des signalements, des limitations d'âge en fonction des contenus pouvant être néfastes pour les mineurs, ainsi que la mise en place de contrôles parentaux.

Ceci dit, il convient de noter que ces nouvelles et futures normes européennes ne brillent pas par leur précision : leur contenu ne semble pas révolutionner les dispositifs existants, notamment en droit français avec la LCEN.

En résumé et en conclusion, la lutte contre les contenus illicites sur les réseaux sociaux doit s'appuyer sur un arsenal législatif complet, combiné à un système de signalement effectif et efficace, qui résulte de l'action des acteurs du numérique. Si le dispositif légal est pour le moment étoffé, il n'en reste pas moins que les mécanismes d'autorégulation doivent être nettement renforcés, afin que les termes de la loi ne restent pas lettre morte. À la lumière de l'actualité et compte tenu de l'agitation qui entoure ces sujets, il est en effet possible que la marge laissée aux entreprises jusqu'à présent s'amenuise, à la faveur d'un durcissement législatif et réglementaire à leur encontre.

41 - Fabrice Perbost, *La lutte contre les propos antisémites sur internet*, 21 juin 2014, RJCom. Juillet/Août 2014.

42 - Commission européenne, Communiqué de presse, 31 mai 2016.

43 - <https://ec.europa.eu/digital-single-market/en/news/proposal-updated-audiovisual-media-services-directive>.

44 - Numerama, *Comment Bruxelles veut réguler YouTube, Twitch ou Periscope*, 25 mai 2016.

45 - Article 15 de la directive n°2000/31, Commerce électronique du 8 juin 2000.

46 - Article 14 de la directive n°2000/31, Commerce électronique du 8 juin 2000.