

SciencesPo

CHAIR DIGITAL, GOVERNANCE AND
SOVEREIGNTY

LES RESEAUX SOCIAUX, ENTRE ENCADREMENT ET AUTO-REGULATION

Florence G'sell

Professeur de droit privé à l'Université de Lorraine
Titulaire de la Chaire Digital, Gouvernance et Souveraineté

Avril 2021

(version provisoire)

Table des matières

I.	L'adoption de nouvelles règles	13
a/	La réécriture des règles de concurrence : le spectre du démantèlement	13
1/	Le débat américain sur le démantèlement	14
2/	L'imposition d'obligations strictes par les autorités de l'Union Européenne.....	17
b/	Vers une responsabilité accrue des plateformes	20
1/	Vers la fin de l'immunité des plateformes aux Etats-Unis.....	21
	L'immunité des plateformes aux Etats Unis	21
	Vers une réforme de la section 230 du <i>Communications Act of 1934</i> ?	28
2/	Une responsabilité relative des plateformes en Europe	34
	Le régime issu de la Directive E-Commerce de 2000	34
	Les textes en cours : le futur <i>Digital Services Act</i> et la loi française confortant le respect des principes de la République	38
c/	La création d'une autorité de régulation	44
II.	Une nouvelle gouvernance des plateformes.....	48
a/	Le recours à une instance tierce : l'expérience de l' <i>Oversight Board</i>	48
b/	L'implication des utilisateurs	55

Les activités de la Chaire sont soutenues par ses partenaires:



Introduction

Le 7 janvier 2021, au lendemain de l'attaque du Capitole des États-Unis par des militants souhaitant empêcher la certification de l'élection présidentielle de 2020 par le Congrès, Mark Zuckerberg, président de Facebook, [annonçait](#) sa décision de suspendre, pour une durée indéterminée, le compte du Président Donald Trump sur Facebook et Instagram, aux motifs que le Président cherchait à empêcher la transition pacifique et légale du pouvoir à son successeur élu, Joe Biden : dans ce contexte, les risques générés par sa présence sur le réseau étaient trop grands. Le 8 janvier, le réseau social Twitter supprimait, à titre permanent, le compte de Donald Trump en invoquant deux tweets du Président, l'un faisant référence aux "patriotes américains" et l'autre indiquant que M. Trump n'assisterait pas à l'investiture du président élu Joseph R. Biden. Le réseau social estimait que ces tweets étaient susceptibles d'inciter d'autres personnes à reproduire les actes violents commis au Capitole, dans la mesure où existaient de nombreux indices indiquant qu'ils étaient lus et compris comme un encouragement à le faire. Le 9 janvier, Google et Apple retiraient de leur magasin d'applications le réseau social Parler, dominé par l'extrême droite, en raison du refus de la plateforme de modérer les contenus violents¹. Le même jour, Amazon refusait de continuer à héberger Parler en invoquant la violation de ses conditions générales d'hébergeur. A l'origine de toutes ces décisions, un constat : les émeutiers auteurs de l'attaque du Capitole s'étaient concertés et organisés sur des sites comme Facebook, Twitter et Parler, où les thèses complotistes et les fausses allégations de fraude électorale avaient proliféré pendant des semaines. Deux mois plus tard, seul Jack Dorsey, le Président de Twitter, reconnaissait clairement, devant le Congrès, le rôle joué par les réseaux sociaux dans les événements du Capitole².

Si ces décisions à l'encontre de Trump ont, sur le moment, soulagé nombre d'américains alarmés par l'escalade des violences, elles ont aussi montré la puissance des grandes entreprises technologiques, qui ont réussi à réduire au silence le

¹ Parler est réapparu sur Apple en avril 2021, après d'être engagé à modérer davantage les contenus.

² K. Conger, "Jack Dorsey says Twitter played a role in U.S. Capitol riot", The New York Times, March 25, 2021

Président des Etats-Unis sortant, qui venait de remporter 74 millions de suffrages, sur le seul fondement de leurs conditions d'utilisation. D'autres dirigeants ont fait l'objet de décisions de modération, certes moins spectaculaires. Fin mars 2021, Facebook a décidé d'empêcher, pendant 30 jours, le Président vénézuélien, Nicolas Maduro, de publier ou commenter quoi que ce soit sur le réseau au motif que l'intéressé avait enfreint de manière récurrente les conditions générales de Facebook en faisant l'éloge d'un médicament contre le covid-19 dont l'efficacité n'est pas démontrée. En octobre 2020, Twitter a supprimé un message de l'ancien Premier ministre malaisien, Mahatir Mohamad, qui, juste après l'attentat commis dans une église de Nice, avait affirmé que les musulmans avaient le droit de « tuer des millions de français pour les massacres du passé ». Quelques mois plus tôt, en mars 2020, deux tweets du Président brésilien Jair Bolsonaro qui relayaient des vidéos faisant la promotion de la fameuse hydroxychloroquine et appelant à mettre fin aux stratégies de distanciation physique avaient été supprimés. Si toutes ces décisions apparaissent justifiées, elles n'en conduisent pas moins à censurer des dirigeants en exercice et suscitent, à cet égard, le malaise tant elles illustrent le pouvoir écrasant des plateformes. Ce sont, en effet, des sociétés commerciales qui deviennent, ce faisant, les arbitres du débat démocratique et contrôlent ce qui peut être dit, publiquement, par des responsables politiques élus. « On se souviendra de cet événement comme d'un tournant dans la bataille pour le contrôle de la parole numérique », twittait Edward Snowden au lendemain de la « déplateformisation » de Trump, pendant qu'Alexey Navalny dénonçait, toujours sur Twitter, « un acte de censure inacceptable » et que la chancelière allemande Angela Merkel y voyait un acte « problématique ». Quelques jours plus tard, Jack Dorsey avait laconiquement commenté : "Si les gens ne sont pas d'accord avec nos règles et leur application, ils peuvent simplement aller sur un autre service internet".

Ces bras de fer avec des responsables publics, auxquels les plateformes opposent leurs conditions générales comme s'ils étaient de simples utilisateurs, ont le mérite d'illustrer la place désormais prise par les grands réseaux sociaux, leur influence considérable sur le débat public et l'ambiguïté de leur position. La force de frappe d'immenses plateformes transnationales réunissant des millions, voire des milliards d'utilisateurs paraît sans limite. Au premier trimestre 2020, Facebook comportait

2,6 milliards d'utilisateurs actifs par mois et 1,73 milliard d'utilisateurs actifs par jour, à comparer avec les 800 millions d'utilisateurs actifs mensuel de la plateforme TikTok, et aux 192 millions d'utilisateurs quotidiens de Twitter, qui, malgré une surface moindre, comporte des comptes « poids lourds » comme celui de Barack Obama (130 millions d'abonnés) ou, par le passé, celui de Donald Trump (86 millions d'abonnés). Partout dans le monde, les populations sont devenues dépendantes des plateformes pour s'informer, communiquer, débattre, s'instruire, se divertir, acheter ou vendre. Les réseaux sociaux sont désormais des lieux fréquentés quotidiennement, voire en permanence, où se déroule cette vie « on life » décrite par Luciano Floridi³, dans laquelle monde physique et monde réel ne sont jamais entièrement séparés. Il est d'ailleurs reconnu que le rôle majeur des réseaux sociaux dans la prise de parole en ligne, l'exercice de la liberté d'expression et la vie démocratique font qu'il est crucial, pour les citoyens comme pour les élus, de pouvoir y accéder. En 2017, la Cour Suprême des Etats-Unis s'est fondée sur le Premier Amendement à la Constitution des Etats-Unis -dont elle a rappelé qu'il garantit que « *toutes les personnes ont accès aux endroits où elles peuvent parler et écouter puis, après réflexion, parler et écouter encore* »⁴- pour censurer une loi adoptée en Caroline du Nord interdisant de manière générale aux délinquants sexuels d'accéder aux réseaux sociaux où ils pouvaient entrer en contact avec des mineurs. Le juge Kennedy en a profité pour souligner que « *si, par le passé, il a pu être difficile d'identifier les lieux (au sens spatial) les plus importants pour l'échange de vues, aujourd'hui la réponse est claire. Il s'agit du cyberspace -"des vastes forums démocratiques de l'Internet"- en général, et des médias sociaux en particulier* ». En France, le Conseil Constitutionnel a jugé dès 2009, qu'« *en l'état actuel des moyens de communication et eu égard au développement généralisé des services de communication au public en ligne ainsi qu'à l'importance prise par ces services pour la participation à la vie démocratique et l'expression des idées et des opinions, [la libre communication des pensées et des opinions] implique la liberté d'accéder à ces services* »⁵. Autrement dit, la liberté d'accéder aux réseaux sociaux constitue l'une des modalités d'exercice de l'article 11 de la Déclaration des droits de l'homme et du citoyen de 1789, même s'il n'en découle pas pour autant que

³ L. Floridi, *The Onlife Manifesto, Being Human in a Hyperconnected Era*, Springer, 2015

⁴ *Packingham v. North Carolina*, 137 S. Ct. 1730 (2017)

⁵ Décision n° 2009-580 DC du 10 juin 2009 : Loi favorisant la diffusion et la protection de la création sur Internet, Considérant n°12.

cet accès puisse, en tant que tel, s'analyser en un droit fondamental. Le Conseil Constitutionnel a eu l'occasion de rappeler récemment, lors de sa décision relative à la loi dite « Avia », que la liberté d'accéder aux réseaux sociaux et de s'y exprimer est constitutionnellement garantie, si bien que les atteintes portées, par le législateur, à l'exercice de cette liberté doivent être nécessaires, adaptées et proportionnées à l'objectif poursuivi⁶.

Dans le même temps, l'importance démesurée prise par les réseaux sociaux s'accompagne de risques aigus pour la société, la vie démocratique et la sécurité des personnes. Le premier des dangers découle des contenus haineux, violents ou incitant à la violence, qui deviennent d'autant plus toxiques que les utilisateurs favorisent de plus en plus les contenus audiovisuels. L'on peut aujourd'hui visionner en direct des fusillades, des suicides, des viols, des actes de maltraitance, voire des meurtres⁷. En 2017, une mère de famille thaïlandaise a vu son compagnon assassiner son enfant en direct sur Facebook Live⁸ : les deux vidéos montrant cet acte épouvantable sont restées disponibles pendant 24h sur Facebook et ont été téléchargées sur YouTube. La même année, une vidéo montrant un adolescent handicapé mental ligoté et bâillonné se faire torturer par quatre personnes est restée en ligne au moins 23h⁹, et une autre diffusant en direct le viol collectif d'une jeune fille n'a entraîné aucune réaction de ceux qui la visionnaient¹⁰. Les actions violentes s'organisent en ligne : c'est sur un groupe Facebook que des extrémistes ont tenté de préparer l'enlèvement de la gouverneure du Michigan, Gretchen Whitmer et l'attaque du Capitole du Michigan¹¹. En France, les réseaux sociaux ont été accusés d'avoir rendu possible le

⁶ Décision n°2020-801 DC du 18 juin 2020 : Loi visant à lutter contre les contenus haineux sur internet, Considérant n°5.

⁷ Par exemple, le meurtre atroce, par Steve Stephens, de Robert Godwin, dont le petit-fils a supplié les utilisateurs des réseaux sociaux de ne plus diffuser la vidéo du meurtre de son grand-père, qui a au final été visionnée 1,6 millions de fois.

⁸ "Thai Mother Saw Daughter Being Killed on Facebook Live", The Guardian, 27 avril 2017

⁹ "Four Ordered Held Without Bail in Alleged Hate-Crime Attack Streamed Live on Facebook", Chicago Tribune, 6 janvier 2017

¹⁰ "Teen Who Was Gang Raped on Facebook Live Is Receiving Threats, Mom Says", NBC News, 22 mars 2017

¹¹ Ch. Warzel, "Facebook and the Group That Planned to Kidnap Gretchen Whitmer," New York Times, Oct. 8, 2020. Après les événements du Capitole, le 6 janvier 2021, Facebook a annoncé vouloir cesser de recommander des groupes politiques, mais beaucoup estiment que cette mesure vient trop tard et ne suffit pas.

déchaînement d'hostilité ayant conduit à l'assassinat odieux du professeur Samuel Paty¹². Ils ont aussi permis l'apparition du mouvement des « gilets jaunes » et des dérives violentes dont ce mouvement a parfois été le théâtre¹³. Parallèlement, les réseaux sociaux sont accusés de permettre, de manière massive et inédite, la discrimination, le harcèlement et l'intimidation des plus vulnérables. Protégés par des pseudonymes, certains utilisateurs en profitent pour s'attaquer aux minorités en tenant, en toute impunité, des propos sexistes, racistes ou injurieux¹⁴. Une étude récente par des chercheurs de l'Université de New York a d'ailleurs démontré un lien entre le nombre de crimes haineux commis dans une ville et le nombre de tweets comportant des contenus discriminatoires relatifs à la race et l'origine ethnique¹⁵. La lutte contre la violence en ligne est toutefois difficile. Il faut, en effet, pouvoir faire intervenir des modérateurs pouvant supprimer, le plus rapidement possible, les contenus les plus traumatisants, ce qui peut paraître impossible face à des contenus diffusés en direct et pouvant devenir viraux en quelques secondes. Par ailleurs, bien qu'il soit en théorie possible de démasquer les auteurs de propos haineux et de les poursuivre, les plateformes rechignent parfois à communiquer les noms des utilisateurs : récemment, Twitter s'est vu reprocher de n'avoir pas répondu à des réquisitions judiciaires concernant l'identité des auteurs d'actes de cyber-harcèlement¹⁶.

Outre la violence physique ou verbale, l'impact des réseaux sociaux sur la vie politique des démocraties inquiète¹⁷. Les publicités politiques sont lucratives pour les

¹² N. Devillier, « Lynchage de Samuel Paty sur les réseaux sociaux : comment réguler les algorithmes de la haine ? », *The Conversation*, 19 octobre 2020

¹³ N. Souillard, B. Sebbah, L. Loubère, L. Thiong-Kay et N. Smyrniotis, « Les Gilets jaunes, étude d'un mouvement social au prisme de ses arènes médiatiques », *Terminal*, 127, 2020.

¹⁴ Danielle K. Citron, "Cyber Civil Rights", *Boston University Law Review*, Vol. 89, pp. 61-125, 2009.

¹⁵ K. Relia, Z. Li, S.H. Cook, R. Chunara, "Race, ethnicity and national origin-based discrimination in social media and hate crimes across 100 U.S. cities", Paper presented at 13th International Conference on Web and Social Media, ICWSM 2019, Munich, Germany.

¹⁶ « Cyber harcèlement : trois victimes attaquent Twitter en justice faute de coopération », *Le Figaro*, 3 février 2021

¹⁷ I. Fassassi, « Les effets des réseaux sociaux dans les campagnes électorales américaines », *Les Nouveaux Cahiers du Conseil Constitutionnel*, 2017/4, n°57, p. 69-86.

plateformes, qui refusent généralement, tel Facebook, d'en contrôler la véracité¹⁸. Les *fake news* et les contenus à caractère complotiste prolifèrent en ligne et constituent autant de moyens de fausser le jeu d'élections démocratiques et de manipuler les opinions, parfois sous l'influence de puissances étrangères. En 2018, l'affaire Cambridge Analytica a montré que la désinformation et la manipulation de masse sont étroitement associées au modèle d'affaire des plateformes, qui collectent les données personnelles des utilisateurs pour ensuite proposer aux annonceurs des publicités ciblées. En l'occurrence, les données d'environ 87 millions d'utilisateurs de Facebook avaient été récoltées sans leur consentement grâce à un questionnaire de personnalité auquel ils répondaient innocemment, permettant ainsi à Cambridge Analytica de développer un algorithme lui permettant de dresser leur profil psychologique. L'entreprise a pu, sur cette base, proposer ses services dans le cadre des campagnes de Donald Trump et des partisans du Brexit, même si l'on ignore toujours l'impact réel de ces pratiques dans le résultat des élections américaines de 2016¹⁹ ou celui du référendum sur le Brexit²⁰. L'épisode montre que le modèle même des plateformes est en soi générateur de risques. Tirant leurs revenus de la publicité, elles ont intérêt à augmenter leur fréquentation et la durée de connexion des utilisateurs, donc à privilégier les contenus qui génèrent de l'engagement, qui sont souvent les plus extrêmes et les plus toxiques. Les données récoltées en masse sur chaque utilisateur permettent aux algorithmes de déterminer le fil d'actualité, les résultats des recherches faites sur la plateforme et les recommandations de rejoindre tel ou tel groupe ou d'ajouter tel ou tel contact. Or, ces algorithmes, conçus pour maximiser le temps passé en ligne, sont fréquemment accusés d'exacerber la propagation de la désinformation, d'accentuer la polarisation politique, voire de privilégier le racisme, l'extrémisme et la violence. Tout récemment, Nick Clegg, le vice-président des « Affaires Globales » de Facebook, a publié un article soutenant que la forte audience des contenus extrêmes

¹⁸ Facebook a simplement supprimé les publicités politiques pendant quelques mois après l'élection de novembre 2020. Google a fait de même. Twitter a entièrement renoncé aux publicités politiques en 2019.

¹⁹ P. Lewis and P. Hilder, "Leaked: Cambridge Analytica's blueprint for Trump victory", The Guardian, 23 March 2018.

²⁰ En octobre 2020, l'Information Commissioner Office (ICO), l'autorité de protection des données, britannique a publié les résultats d'une enquête de trois ans qui conclut que Cambridge Analytica n'a pas directement utilisé des données en vue d'influencer le vote dans le cadre du référendum sur le Brexit, ni apporté son aide à la Russie sur ce sujet : « L'implication de Cambridge Analytica dans la campagne du Brexit était limitée... tout comme l'efficacité de ses outils », Le Monde, 8 octobre 2020.

ou choquants n'est pas le fait des algorithmes eux-mêmes mais des utilisateurs et de leurs choix en ligne²¹. Il reste que les algorithmes et leurs modes de fonctionnement ne sont pas dévoilés au public : les politiques des plateformes en la matière restent désespérément opaques.

Cette maîtrise de l'algorithme est le principal pouvoir de la plateforme. Comme l'écrivait Lessig, le code constitue, dans le cyberspace, le type de contrainte dominant en ce qu'il produit un effet régulateur plus tranchant que les autres modes de contraintes classiques que sont le droit, le marché et les normes sociales²². Dans le monde virtuel, ce sont ceux qui conçoivent l'architecture et déterminent les modalités d'action et d'échange sur le réseau qui font les règles : *Code is Law*²³. Une autre manière de l'exprimer revient à employer le néologisme d' « algocratie », forgé par le sociologue Aneesh Aneesh²⁴, qui désigne par ce terme des organisations virtuelles dont la gouvernance est organisée par des algorithmes et dans lesquelles l'éventail des actions possibles est programmé à l'avance. Les réseaux sociaux constituent des systèmes algocratiques au sein desquels ceux qui ont la charge de développer et configurer l'architecture sont en situation de domination. Ce pouvoir ne peut qu'inquiéter. Malgré les risques qu'elles génèrent, les plateformes n'ont qu'une responsabilité limitée à l'égard de leurs utilisateurs et ne sont pas réellement tenues à la transparence quant à leurs processus de décision et la configuration de leurs outils. Alors que le gigantisme des plateformes, leur omniprésence dans la vie quotidienne, le pouvoir qu'elles exercent désormais dans l'information des citoyens et le débat public conduisent insensiblement à leur assigner un rôle quasi-étatique²⁵, c'est sans les garanties démocratiques propres aux Etats de droit. Bien au contraire, ces réseaux sont créés et administrés par des entreprises privées à but lucratif, qui agissent dans l'intérêt de leurs actionnaires et ne sont pas soumises aux contraintes pesant sur les personnes publiques. Alors que les autorités étatiques sont responsables devant leurs citoyens, les plateformes sont pilotées de manière verticale par des dirigeants

²¹ Nick Clegg, *You and the Algorithm: It Takes Two to Tango*, Medium, 31 mars 2021.

²² Lawrence Lessig, *The Law of the Horse : What Cyberlaw Might Teach*

²³ L. Lessig, "Code is Law: on Liberty in Cyberspace", Harvard Magazine, January 1, 2000

²⁴ A. Aneesh, « Global Labor. Algocratic Mods of Organization », *Sociological Theory* 2009.

²⁵ Henry Farrell, Margaret Levi, and Tim O'Reilly, "Mark Zuckerberg runs a nation-state, and he's the king", Vox, 10 avril 2018

fondateurs souvent très médiatisés, comme Mark Zuckerberg (Facebook) ou Jack Dorsey (Twitter), qui détiennent la majorité du pouvoir de décision de l'entreprise. En outre, le caractère dématérialisé et transnational de leurs activités leur permet justement de se soustraire à l'emprise des Etats souverains et à l'application des droits nationaux. En matière de politique des contenus, il arrive que les plateformes entendent faire prévaloir leurs propres principes sur les droits nationaux. Twitter a récemment refusé de donner suite aux demandes de suppression émanant du gouvernement indien aux prises avec des manifestations²⁶. De son côté, Facebook refuse de tenir compte de certaines législations nationales et de sanctionner, en dépit d'une incrimination légale, des comportements compatibles avec ses valeurs : une liste d'« infractions non reconnues par Facebook » sert de références aux modérateurs, en dépit des lois nationales²⁷. Certaines plateformes n'hésitent d'ailleurs pas à entrer dans des rapports de force ouverts avec les Etats, tel Facebook décidant brutalement, en février 2021, de supprimer temporairement de son réseau les contenus de la presse australienne afin de faire pression sur le gouvernement australien à propos d'un projet de loi l'obligeant à rémunérer les éditeurs de presse²⁸.

La toute-puissance des plateformes, le nombre vertigineux des personnes qu'elles rassemblent, le fait qu'elles échappent au contrôle des Etats comme à celui de leurs utilisateurs, le sentiment qu'elles sont en mesure d'agir sans contrepouvoir conduisent logiquement à mobiliser à leur sujet, les réflexions et les outils traditionnellement pensés pour analyser et limiter le pouvoir de l'Etat. L'idée est de déterminer comment réduire le pouvoir excessif des grandes entreprises technologiques et restaurer un certain degré de contrôle démocratique sur les infrastructures numériques. Tel est l'objet des travaux qui s'intègrent dans le champ du « constitutionnalisme numérique » (*digital constitutionalism*). Si la pensée constitutionnaliste renvoie à la nécessité de contrôler, de limiter et de restreindre les pouvoirs de l'Etat, les constitutions ne font pas que délimiter les frontières du pouvoir d'un Etat sur ses citoyens : elles créent des institutions permettant la coordination et l'action collective. Dans le monde numérique,

²⁶ « Twitter resists Indian government's orders to block protest accounts », The Verge, 10 février 2021

²⁷ « Facebook leak underscores strategy to operate in repressive regimes », The Guardian, 23 March 2021

²⁸ « Australie : Facebook bloque le partage d'articles, y compris pour les services de secours et de santé », Le Monde, 17 février 2021.

le terme de « constitutionnalisme numérique » désigne ainsi un champ de recherche qui rassemble les initiatives visant à articuler un ensemble de droits politiques, de normes de gouvernance et de limites à l'exercice du pouvoir sur l'Internet²⁹. Il peut s'agir, par exemple, de l'élaboration d'une "charte des droits de l'internet", à l'image du *Marco Civil da Internet*, cette loi brésilienne adoptée en 2014 qui proclame un ensemble de droits au bénéfice des utilisateurs. Il peut s'agir, également, d'une réflexion sur la gouvernance des plateformes, de manière à y instaurer un système de *checks and balances* par la création, par exemple, de nouvelles instances, à l'image de l'Oversight Board de Facebook, ou par la mise en place de référendums permettant d'accorder davantage de place aux utilisateurs. Il peut s'agir, enfin, d'une réflexion approfondie sur la police du discours devant prévaloir sur des réseaux sociaux pratiquant désormais massivement la modération, à l'image de Facebook qui, après avoir longtemps promu l'inaction au nom de la liberté d'expression et de la neutralité, a dû prendre acte de la prolifération des discours de haine et de la désinformation.

Pour l'heure, les plateformes sont sans réel contrepouvoir et s'auto-régulent, en ne se voyant généralement appliquer que des réglementations limitées. Leur approche de la régulation a toutefois évolué du fait de la pression subie depuis l'affaire Cambridge Analytica et des accusations de laxisme face aux messages haineux et les fausses nouvelles. En 2019, Mark Zuckerberg a appelé à une réglementation spécifique aux plateformes³⁰ -qui ne serait ni celle que l'on applique ordinairement aux entreprises de télécommunication, ni celle qui régit les médias- de manière à notamment à les aider à clarifier la distinction entre ce qui peut être dit en ligne et ce qui ne peut pas l'être. Nombreux sont, aujourd'hui, ceux qui croient en la nécessité d'adopter de nouvelles règles, propres aux réseaux sociaux et, plus largement, aux plateformes, même si les voies préconisées varient. Aux Etats-Unis, par exemple, les démocrates souhaitent que les réseaux sociaux modèrent davantage les contenus les plus toxiques alors que les républicains se plaignent de pratiques de modération qu'ils jugent biaisées à l'encontre des plus conservateurs. Quelles que soient les oppositions, il apparaît bien,

²⁹ D. Redeker, L. Gill L, U. Gasser "Towards digital constitutionalism? Mapping attempts to craft an Internet Bill of Rights" *International Communication Gazette*. 2018, 80 (4), pp. 302-319; Edoardo Celeste, "Digital Constitutionalism : Mapping the constitutional response to digital technology's challenges", High Discussion Paper Series, 2018/02.

³⁰M. Zuckerberg, "The Internet needs new rules. Let's start in these four areas", Washington Post, March 30, 2019.

toutefois, que le souhait de limiter le pouvoir écrasant des plateformes soit largement partagé. Plusieurs options peuvent être envisagées pour parvenir à cette fin. L'on pourrait, classiquement, opter en faveur de l'intervention étatique et de l'adoption, par les Etats, de règles plus contraignantes (I). Mais la limitation du pouvoir des plateformes peut également passer une modification en profondeur de leurs modes de gouvernance (II).

I. L'adoption de nouvelles règles

L'intervention étatique pour encadrer davantage les pratiques des plateformes peut passer par la réécriture de règles traditionnelles, qu'il s'agisse des règles de concurrence (A) ou de celles qui prévoient la responsabilité civile des acteurs technologiques (B). Mais elle pourrait également consister à placer les réseaux sous le contrôle d'une autorité de régulation de manière à limiter leur pouvoir excessif (C).

a/ La réécriture des règles de concurrence : le spectre du démantèlement

Les plateformes sont souvent décrites comme des monopoles naturels créés par le double jeu des économies d'échelle et des effets de réseau, qui sont notables dans le cas des réseaux sociaux tant les utilisateurs souhaitent se retrouver en un même lieu. De fait, les principaux réseaux sociaux se comptent sur les doigts d'une main tout en attirant les foules. Au premier trimestre 2020, 95% des jeunes adultes américains utilisaient Facebook, qui capte, avec Google, 63% des publicités sur internet³¹. Dans un tel environnement, les opérations de concentration visant à éliminer tel ou tel concurrent, réel ou potentiel (« acquisitions tueuses »)³² sont fréquentes. De même, l'acquisition par Facebook d'Instagram puis de Whatsapp est venue conforter une situation de domination déjà problématiques. Les positions dominantes et les barrières à l'entrée sont d'autant plus préjudiciables que les plateformes sont en mesure à la fois de contrôler l'architecture, de contraindre leurs utilisateurs à utiliser les applications qui leurs sont propres, de collecter les données générées par le trafic sur leur réseau en les exploitant à leur profit et de faire concurrence à leurs utilisateurs professionnels en proposant leurs propres biens et services sur la plateforme.

Dans un tel contexte, les outils de concurrence traditionnels paraissent échouer à assainir les pratiques et à rectifier la structure du marché. Certes, les autorités de concurrence européennes ou américaines ouvrent régulièrement des enquêtes et n'ont pas hésité à prononcer des condamnations lourdes, à l'image de l'amende de 4,34 milliards d'euros infligée à Google en 2018 pour avoir abusé de sa position dominante en imposant ses applications pré-installées. Mais le droit de la concurrence ne permet pas de sanctionner tous les comportements préjudiciables ni de prévenir

³¹ Elsa Conesa, «Le spectre de la Standard Oil plane sur la Silicon Valley », Les Echos, 26 février 2018

³² C. Cunningham, F. Edered, S. Ma, «Killer Acquisitions», *Working Paper*, 2018.

l'ensemble des opérations de concentrations qui permettent aux plateformes les plus puissantes de consolider leur position de domination et d'en user pour imposer leurs conditions aux utilisateurs. Les remèdes traditionnels du droit de la concurrence arrivent trop tard, à la suite d'enquêtes longues, alors même que les marchés numériques évoluent très vite. Les sanctions prononcées n'ébranlent pas toujours des entreprises aux moyens colossaux. Plus fondamentalement, l'on peut se demander si des marchés plus compétitifs, des entreprises plus petites, une propriété plus largement dispersée- soient les objectifs classiques de l'antitrust- ne seraient pas de nature à régler les difficultés posées par les plateformes. Cette situation explique pourquoi nombreux sont ceux qui, aux Etats-Unis ou en Europe, appellent à l'adoption de réglementations énergiques devant permettre de mettre un terme non seulement au quasi-monopole de certaines plateformes mais aussi aux pratiques les plus condamnables. A l'heure actuelle, la possibilité d'un démantèlement des plateformes les plus puissantes est fréquemment évoquée aux Etats-Unis (1). L'Union Européenne a, en revanche, opté en faveur de l'adoption de réglementations destinées à peser sur le comportement des plateformes par l'imposition d'obligations strictes (2).

1/ Le débat américain sur le démantèlement

En 2019, Chris Hughes, co-fondateur de Facebook, réclamait, dans un éditorial remarqué, le démantèlement de Facebook³³. Cette demande était justifiée par l'insatiable désir de puissance et de croissance imputé à l'entreprise et à son fondateur, illustré par l'acquisition d'Instagram en 2012 et de WhatsApp en 2014. La possibilité d'un démantèlement des grandes plateformes s'est ensuite trouvée au cœur de la campagne présidentielle américaine en raison de la proposition de la candidate démocrate, Elizabeth Warren, d'interdire aux entreprises réalisant plus de 25 millions de dollars de chiffre d'affaires de faire concurrence à leurs utilisateurs en proposant leurs propres services sur la plateforme qu'elles détiennent, ce que font, par exemple, Google ou Amazon.

Le démantèlement est également promu par une nouvelle génération d'universitaires américains souhaitant revenir sur l'approche dominante de l'antitrust depuis les années 1970. Ce dernier courant, désormais traditionnel, largement inspiré par l'Ecole

³³ Chris Hughes, "It's Time to Break Up Facebook", New York Times, May 9. 2019

de Chicago, s'accommode des monopoles dès lors qu'ils paraissent profiter aux consommateurs, notamment par des prix attractifs. Or les prix actuellement pratiqués par les grandes plateformes sont particulièrement bas, et la gratuité fréquente dans le monde numérique. Dans cette perspective, le gigantisme des plateformes serait tolérable dès lors que le consommateur est gagnant. Cette analyse est cependant vigoureusement contestée par de nouveaux auteurs qui invitent à s'affranchir du critère des prix pour revenir à une analyse traditionnelle consistant à prohiber les monopoles en tant que tels et retrouver une stratégie que les autorités américaines semblent avoir abandonnée depuis le démantèlement d'AT&T en 1984. Ce nouveau mouvement est parfois appelé le "*New Brandeis movement*", du nom du juge Louis Brandeis, conseiller du Président Woodrow Wilson avant sa nomination à la Cour Suprême, qui avait incité le Président américain à adopter le *Clayton Antitrust Act* de 1914 et à créer la *Federal Trade Commission*. Brandeis préconisait de combattre les monopoles en tant que tels car « *we can have democracy in this country, or we can have great wealth concentrated in the hands of a few, but we can't have both* ». C'est cette nouvelle approche qui conduit le professeur de droit Tim Wu, récemment nommé par Joe Biden conseiller spécial du Président sur la technologie et la politique de concurrence au sein du *National Economic Council*, à voir dans les grandes plateformes des super monopoles peu vertueux et comparables à la Standard Oil, démantelée en 1911 sur le fondement du Sherman Antitrust Act de 1890³⁴ et à se déclarer partisan du démantèlement de Facebook au motif que certaines plateformes sont devenues plus puissantes que les Etats eux-mêmes. Dans la même optique, Lina Kahn, désignée par Joe Biden pour siéger parmi les 5 commissaires de la *Federal Trade Commission*, a mis en évidence le pouvoir structurel d'Amazon, qui va bien au delà de sa part de marché compte-tenu des données dont la plateforme dispose, de son infrastructure d'expédition et de manutention, et de sa capacité de contrôler l'accès au marché des vendeurs tiers. Surtout, Lina Kahn a montré que le contrôle, par la plateforme, de l'infrastructure sur laquelle se déroulent les transactions lui confère un avantage que le droit antitrust classique ne permet pas de saisir et réguler³⁵, en permettant notamment à la plateforme de faire concurrence à ses

³⁴ Tim Wu, *The Curse of Bigness: Antitrust in the New Gilded Age*, Columbia Global Reports, 2018

³⁵ Lina M. Kahn, "*Amazon's Antitrust Paradox*", *Yale Law Journal*, 2017, vol. 126 n°3, pp. 564-907.

vendeurs tiers en pratiquant des prix prédateurs. Elle a également souligné que les prix faibles ou la gratuité pratiqués par les plateformes est compensé par une utilisation massive des données des utilisateurs et que leur gigantisme pose un problème social et politique. Malgré toute la portée des écrits de Tim Wu ou de Lina Kahn, les arguments mobilisés par les partisans du démantèlement ne suscitent toutefois pas l'unanimité chez les experts. Plusieurs spécialistes, comme Fiona Scott Morton³⁶ ou Herb Hovenkamp ont fait part de leurs doutes et souligné qu'une solution aussi radicale pourrait avoir des effets inattendus et problématique, voire se révéler au final préjudiciable aux consommateurs³⁷. Il est vrai que le démantèlement des plus grands réseaux sociaux pourrait conduire à une compétition plus intense mais de courte durée tant les utilisateurs tendent à vouloir se retrouver en un même lieu : sur le long terme, une plateforme finirait toujours par l'emporter.

De leur côté, les autorités américaines sont déterminées à agir. En octobre 2020, le *Department of Justice* et les *Attorney Generals* de 11 Etats ont décidé de poursuivre Google pour pratiques anticoncurrentielles. En décembre 2020, les autorités fédérales et étatiques ont déclenché des poursuites contre Facebook en raison de sa stratégie monopolistique : la *Federal Trade Commission* et 48 Etats américains reprochent en effet à Facebook un comportement visant à éliminer toute concurrence. En demandant, pour cette raison, que Facebook se voit ordonner de se séparer d'Instagram et Whatsapp, ces autorités ne souhaitent pas obtenir un démantèlement au sens fort mais défaire certaines acquisitions sans désintégrer l'entreprise. La FTC souhaite également obtenir une injonction obligeant Facebook à demander une autorisation pour toute future acquisition et lui interdisant d'imposer des conditions anticoncurrentielles aux développeurs. Quant au législateur fédéral, il devrait prochainement intervenir si l'on en croit le rapport publié par le *House Judiciary Committee's antitrust panel* du Congrès en octobre 2020³⁸, qui a pris acte du caractère très concentré des marchés numériques, souligné qu'à l'horizon d'une décennie (soit 2030), 30% de la production économique brute mondiale serait le fait des grandes

³⁶ Fiona Scott Morton, "Why "breaking up" big tech probably won't work", *The Washington Post*, July 16, 2019

³⁷ Herbert Hovenkamp, "Antitrust Remedies for Big Tech", *The Regulatory Review*, January 18, 2021

³⁸ U.S. House Judiciary Committee's Subcommittee on Antitrust, Investigation of Competition in Digital Markets, Majority Staff Report and Recommendations, October 2020

entreprises du secteur (Google, Apple, Amazon, Facebook) et d'une poignée d'autres et insisté sur le caractère répréhensible des pratiques adoptées par les grandes plateformes en vue d'obtenir et maintenir leur position de domination, en sacrifiant notamment les droits des utilisateurs sur leurs données. Le rapport a formulé diverses recommandations incluant un renforcement des pouvoirs et des moyens des agences fédérales, un contrôle renforcé des acquisitions stratégiques, la possibilité de séparations structurelles afin d'empêcher les plateformes, d'agir à la fois comme fournisseur de l'architecture et concurrent des utilisateurs professionnels, l'interdiction pour les plateformes de s'auto-référencer, l'interopérabilité des applications et la portabilité des données. Depuis lors, plusieurs projets de texte ont été déposés au Congrès. Un amendement déposé début février 2021 par le sénateur républicain Hawley prévoyait ainsi l'interdiction de principe de toutes les opérations de concentration par des plateformes en ligne en situation de domination sur le marché, sauf à ce que la plateforme concernée démontre que le projet ne restreint pas la concurrence. De son côté, la démocrate Amy Klobuchar a déposé le 4 février une proposition de *Competition and Antitrust Reform Act* qui s'inspire des propositions formulées dans le rapport parlementaire d'octobre 2020 pour prévoir l'amélioration des moyens dont disposent les autorités fédérales compétentes en matière d'antitrust, un allègement des critères pour démontrer le caractère anticoncurrentiel d'une fusion, l'obligation pour les parties à une concentration de démontrer que le rapport bénéfice-risque de l'opération est positif et l'interdiction des « comportements d'exclusion » présentant un « risque appréciable d'atteinte à la concurrence ». Les autorités fédérales, FTC et DOJ pourraient prononcer des sanctions allant jusqu'à 15 % du chiffre d'affaires de la plateforme ou 30 % du chiffre d'affaires de l'activité concernée.

2/ L'imposition d'obligations strictes par les autorités de l'Union Européenne

En Europe, les séparations structurelles font partie des sanctions rendues possibles par l'article 102 du Traité sur le Fonctionnement de l'Union Européenne en présence d'un abus de position dominante. Encore faut-il, cependant, établir un comportement anti-concurrentiel au sens du Traité, ce qui n'est pas si simple et suppose de conduire une longue enquête de concurrence destinée à analyser en profondeur le secteur concerné. De fait, si l'hypothèse du démantèlement est parfois évoquée en Europe, elle est, cependant, généralement repoussée au motif que d'autres mesures plus

appropriées peuvent, dans l'immédiat, permettre de lutter contre les pratiques les plus condamnables des plateformes. Il n'est, dit-on, pas certain que la solution du démantèlement serait de nature à remédier durablement aux problèmes structurels présentés par les marchés numériques. Privant les plus grandes plateformes de la possibilité de collecter les données des utilisateurs de la manière massive qui prévaut actuellement, le démantèlement bousculerait leur modèle d'affaire et pourrait les conduire, afin de préserver leur compétitivité, à modifier leurs prix et leurs conditions d'utilisation au détriment des consommateurs³⁹. L'on peut donc comprendre que la stratégie préconisée dans l'Union Européenne consiste plutôt à surveiller de près le comportement des plateformes dominantes et à intervenir rapidement en présence de pratiques condamnables. En 2019, un rapport commandé par Margrethe Vestager⁴⁰ a recommandé d'imposer aux plateformes de nouvelles obligations, comme l'interopérabilité et la portabilité des données, d'agir plus rapidement, et de surveiller les acquisitions sans toutefois modifier les seuils actuels d'application des règles de concurrence. Une autre étude commandée par le gouvernement britannique, le rapport Furman⁴¹, a préconisé un meilleur contrôle des fusions et acquisitions, le renforcement des pouvoirs de l'autorité de concurrence, et la création d'une nouvelle autorité chargée de réguler le secteur numérique et les pratiques en matière de données. C'est sur ces entrefaits qu'a été décidée l'adoption de nouveaux instruments législatifs, de manière à pallier les insuffisances du droit antitrust actuel qui ne permet ni une surveillance étroite, ni une réaction rapide.

L'Union Européenne a ainsi opté pour l'adoption de nouvelles règles contraignantes s'appliquant *ex ante* aux plateformes les plus puissantes, qui contrôlent à ce titre l'accès aux marchés numériques (*gatekeepers*). Il s'agit ici d'agir sur le comportement des plateformes en leur imposant des obligations dont la simple violation peut donner lieu à sanction sans qu'une enquête de concurrence soit nécessaire. Ces nouvelles obligations sont prévues par la proposition de Règlement *Digital Markets Act*⁴²,

³⁹ Alexandre de Cornière, « Should we break-up Big Techs ? », The TSEconomist,

⁴⁰ J. Cremer, Y. A. de Montjoye et H. Schweitzer, « Competition Policy for the Digital Era », 2019

⁴¹ J. Furman, "Unlocking digital competition : report from the digital competition expert panel", 2019

⁴² Proposition de Règlement relatif aux marchés contestables et équitables dans le secteur numérique (législation sur les marchés numériques) COM (2020) 842 final

dévoilée le 15 décembre 2020, qui s'applique aux entreprises qui contrôlent l'accès aux marchés numériques : les *gatekeepers*. Il s'agit donc là de réguler non seulement les réseaux sociaux mais également toutes les plateformes fournissant, par exemple, des services d'intermédiation, de moteur de recherche, de communication interpersonnelle ou d'hébergement en nuage. Les *gatekeepers* sont qualifiés tels s'ils ont un impact significatif sur le marché intérieur, exploitent un service de plateforme qui sert de passerelle importante pour les professionnels désirant atteindre les utilisateurs et jouissent d'une position bien établie et durable de manière actuelle ou dans un avenir proche. S'agissant des plateformes, celles qui réalisent en Europe un chiffre d'affaires d'au moins 6,5 milliards d'euro et disposent d'une capitalisation d'au moins 65 milliards d'euro tout en comportant au moins 45 millions d'utilisateurs et plus de 10000 utilisateurs professionnels actifs en Europe sont présumés être des *gatekeepers*. S'agissant des réseaux sociaux proprement dit, Facebook peut sans aucun doute être qualifié de *gatekeeper* mais il n'est pas acquis qu'il en aille de même de tous les réseaux sociaux, dès lors que ceux-ci n'ont pas, à proprement parler, d'activité d'intermédiation. Les *gatekeepers* sont soumis à des obligations particulières qui sont fixées directement par le texte (article 5) comme ne pas croiser les données personnelles des utilisateurs avec celles d'autres services sans leur consentement, ne pas imposer leur service d'identification aux utilisateurs professionnels, ou permettre aux utilisateurs professionnels de promouvoir leurs offres sur la plateforme, tout en ayant la possibilité de conclure directement des contrats avec les utilisateurs finaux. D'autres obligations sont susceptibles d'être précisées (article 6) par la Commission en fonction des services de l'entreprise concernées comme le fait de permettre la portabilité des données des utilisateurs, de donner accès aux données relatives à leur activités aux utilisateurs professionnels ou de fournir aux annonceurs sur la plateforme un certain nombre d'informations. Le respect des obligations prévues par le texte fera l'objet d'un contrôle approfondi par la Commission et leur violation pourra donner lieu à des sanctions dont les « séparations structurelles » font partie. Le texte de la proposition de Règlement (article 16) prévoit, en effet, la possibilité d'ordonner des mesures structurelles en présence d'un *gatekeeper* qui a systématiquement enfreint ses obligations et renforcé de ce fait sa position. Les mesures structurelles pourront aller jusqu'à imposer la cession d'une activité, mais ne seront possibles que si aucune mesure corrective comportementale aussi efficace n'est envisageable ou si la mesure structurelle serait moins lourde de conséquences.

pour le *gatekeeper* qu'une mesure corrective comportementale. Il est donc bien clair, dans un tel contexte, qu'aucun démantèlement ne sera décidé à partir du simple constat qu'une entreprise est en situation de domination ou de monopole. Une telle issue n'est perçue, En Europe, que comme une solution de dernière extrémité. Cette stratégie est-elle de nature à réduire de manière effective le pouvoir excessif des plus grandes plateformes ? Elle permettra sans doute de prévenir les comportements les plus préjudiciables aux utilisateurs professionnels ou consommateurs en empêchant les plateformes d'imposer des conditions inacceptables. Elle ne règle pas, toutefois, la question des contenus, qui ne peut être abordée qu'au travers d'une réflexion sur la responsabilité des plateformes.

b/ Vers une responsabilité accrue des plateformes

La prolifération des contenus en ligne toxiques ou dangereux conduit à conclure que le régime d'autorégulation ayant prévalu jusqu'à présent, qui repose sur une large immunité reconnue aux plateformes, n'est pas satisfaisant. Certes, les réseaux sociaux ont, de leur propre initiative, mis en place les mécanismes nécessaires pour lutter contre la désinformation et les contenus haineux. Mais les moyens restent limités : Facebook emploie 35000 modérateurs pour 100 milliards de publications postées chaque jour sur la plateforme. Les algorithmes lui permettent de détecter et retirer très vite près de 99% des contenus à caractère terroriste, mais il n'en va pas de même des contenus haineux⁴³. De fait, l'incitation à agir n'est, pour les plateformes, que limitée compte-tenu de l'absence de responsabilité des plateformes du fait des contenus illicites qu'elles hébergent. Un tel régime d'irresponsabilité, qui a rendu possible l'émergence des grandes plateformes contemporaines, ne convainc plus après les événements du Capitole et une crise sanitaire durant laquelle de nombreuses fausses informations sont devenues virales. D'où l'idée de revenir sur l'immunité large accordée aux plateformes aux Etats-Unis (1) et d'imposer de nouvelles obligations en Europe (2).

⁴³ R. Balenieri, « Mark Zuckerberg met les contenus haineux au cour de sa tournée en Europe », Les Echos, 17 février 2020

1/ Vers la fin de l'immunité des plateformes aux Etats-Unis

Depuis le milieu des années 1990, les plateformes jouissent, aux Etats-Unis, d'une immunité quasi-absolue dont il est probable qu'elle soit, dans un avenir proche, sérieusement limitée.

L'immunité des plateformes aux Etats Unis

Au début des années 1990, les services en ligne publiant des contenus de tiers étaient tantôt qualifiés de simples hébergeurs se contentant de fournir une infrastructure, tantôt considérés comme des éditeurs responsables des contenus postés par leurs utilisateurs. La qualification d'éditeur avait été écartée dans le cas de CompuServe, alors considéré comme une « bibliothèque numérique »⁴⁴, tandis que Prodigy avait été jugé responsable de propos diffamatoires tenus par un utilisateur au motif que la plateforme se vantait d'exercer un contrôle éditorial similaire à celui d'un journal⁴⁵. En 1996, le Congrès vota le *Communications Decency Act*, qui figurait au titre V du *Telecommunications Act of 1996* amendant le *Communication Act of 1934*. Son objet n'était pas d'exonérer les services en ligne de toute responsabilité mais d'organiser la lutte contre les contenus obscènes ou indécents auxquels les mineurs pouvaient être confrontés en ligne. Une grande partie de la loi fut censurée par la Cour Suprême dans l'arrêt *Reno v. ACLU*⁴⁶ au motif qu'elle violait le droit à la liberté d'expression garanti par le Premier Amendement à la Constitution. Ne resta alors en vigueur que le contenu d'un amendement ajouté lors des débats à la Chambre des Représentants, la section 230 du *Communications Act of 1934* (47 U.S.C. §230) dont l'objectif était principalement d'inciter les plateformes à intervenir face aux contenus illicites. Ce texte toujours en vigueur, dont beaucoup considèrent qu'il a fondé l'internet contemporain, prévoit à la fois que les plateformes ne sont en aucun cas responsables du fait des contenus postés par les tiers et qu'elles peuvent se livrer de bonne foi à des activités de modération sans encourir aucune responsabilité.

- *L'immunité des plateformes du fait des contenus postés en ligne par les tiers*

⁴⁴ *Cubby, Inc. v. CompuServe Inc.*, 776 F. Supp. 135 (S.D.N.Y. 1991)

⁴⁵ *Stratton Oakmont Inc v Prodigy Services Co*, 23 Media L. Rep. 1794 (N.Y. Sup. Ct. 1995)

⁴⁶ *Reno v. American Civil Liberties Union*, 521 U.S. 844 (1997)

La section 230(c)(1) du *Communications Act of 1934* prévoit qu'aucun fournisseur ou utilisateur d'un service informatique interactif ne doit être traité comme l'éditeur ou l'auteur d'une information fournie par un tiers. Il en découle que le fournisseur de l'architecture ne peut jamais être rendu responsable du fait des contenus postés par les utilisateurs de l'architecture. La portée de l'immunité est d'autant plus grande que la section 230(e)(3) prévoit expressément que les Etats ou autorités locales ne peuvent imposer une responsabilité en contravention avec ce texte fédéral. La loi bénéficie aux grandes plateformes mais également à tous ceux qui sont amenés à permettre à autrui de publier des contenus en ligne, tels les particuliers qui tiennent un blog où ils autorisent les commentaires. Les juridictions américaines ont interprété le principe posé par la section 230(c)(1) de manière extrêmement large, alors même qu'il n'est pas acquis que telle était l'intention des rédacteurs de la loi. Dès 1997, une Cour d'appel fédérale a jugé⁴⁷ que la section 230 protège les services en ligne de toute responsabilité, y compris lorsque les plateformes peuvent être qualifiées de distributeurs (« *distributors* »), ce qui les conduit normalement à être responsables dès lors qu'il est établi qu'elles avaient conscience du caractère illégal d'un contenu. Depuis lors, la section 230 est interprétée comme consacrant un principe de non intervention de l'Etat : l'Internet serait un espace dont les acteurs s'autorégulent, en l'absence de toute interférence étatique et de toute responsabilité juridique. Les plateformes ne sont donc jamais responsables du fait de contenus illégaux qu'elles hébergent, y compris lorsqu'elles les ont délibérément laissés en ligne tout en connaissant leur caractère illicite. Les tribunaux américains ont donc exclu, sur le fondement de la section 230, la responsabilité d'un site internet d'annonces classées pour avoir hébergé des publicités relatives à la prostitution de jeunes filles mineures⁴⁸ et celle d'un site d'annonces facilitant des ventes d'armes illicites ayant conduit à des actes de violence⁴⁹. Certes, l'immunité n'est pas totale et quelques exceptions sont prévues par la section 230 : le texte ne peut faire obstacle aux poursuites fondées sur la législation pénale fédérale, sur les textes relatifs à la propriété intellectuelle, sur des lois étatiques "compatible avec" la section 230, sur *l'Electronic Communications*

⁴⁷ *Zeran v. America On Line Inc.*, 129 F.3d 327 (4th Cir. 1997)

⁴⁸ *Jane Doe No. 1 v. Backpage.com, LLC*, 817 F.3d 12 (1st Cir. 2016)

⁴⁹ *Daniel v. Armslist, LLC*, 926 N.W.2d 710 (Wis. 2019)

Privacy Act de 1986⁵⁰, et, depuis 2018, sur la législation relative au trafic sexuel⁵¹. En outre, les plateformes sont responsables si elles sont réellement les auteurs des contenus concernés ou ne respectent pas leurs promesses de retrait⁵².

L'immunité très étendue garantie par la section 230 a rendu possible la croissance exponentielle des géants technologiques que nous connaissons aujourd'hui comme Twitter, Facebook, YouTube ou Amazon. Les grandes plateformes ne sont, grâce à cette loi, aucunement responsables du fait des contenus illicites ou toxiques qu'elles hébergent en masse, tout en bénéficiant largement du trafic généré par ces contenus qui leur permet de collecter des données et de vendre des publicités. La taille et le nombre d'utilisateurs des plateformes rend, du, reste, presque inévitable la présence massive de contenus illégaux et répréhensibles. A cet égard, les plateformes ont souvent été accusées de réagir trop mollement et tardivement face aux contenus les plus toxiques, ce qui justifierait de revenir sur leur immunité. Certains experts ajoutent qu'en épargnant aux plateformes de prendre en charge les risques qu'elles créent, la section 230 organise une forme de « subvention » des comportements illicites que les plateformes auraient intérêt à voir proliférer compte-tenu des revenus publicitaires qu'ils suscitent⁵³ et permet la domination informationnelle des puissants et des privilégiés et conduit à réduire au silence les plus faibles et les plus vulnérables⁵⁴.

- *La liberté de modérer*

La section 230(c)(2)(A) du *Communications Act of 1934* ne prévoit pas seulement l'immunité des plateformes du fait des contenus postés par les tiers : elle prévoit également une immunité « du bon samaritain » permettant aux plateformes d'intervenir de bonne foi sur lesdits contenus sans engager leur responsabilité (« *Protection for « Good Samaritan » blocking and screening of offensive material* »).

⁵⁰ Qui protège les courriels, les conversations téléphoniques et les données stockées électroniquement.

⁵¹ *Stop Enabling Sex Traffickers Act* (SESTA) et *Allow States and Victims to Fight Online Sex Trafficking Act* (FOSTA)

⁵² *Barnes v. Yahoo!, Inc.*, 570 F.3d 1096 (9th Cir. 2009)

⁵³ Mary Anne Franks, "Reforming Section 230 and Platform Liability", Stanford Cyber Policy Series (2021); Danielle Keats Citron & Mary Anne Franks, "The Internet as a Speech Machine and Other Myths Confounding Section 230 Reform," University of Chicago Legal Forum 3 (2020): 70.

⁵⁴ Mary Ann Frank, "Reforming Section 230 and Platform Liability", préc.

Dans l'espace physique, l'« immunité du bon samaritain » offre une protection à ceux qui tentent, de bonne foi et sans obligation légale, d'aider des personnes en difficultés. Dans l'espace virtuel, l'immunité du bon samaritain garantit aux fournisseurs et aux utilisateurs de services en ligne qu'ils ne seront pas tenus pour responsables de toute action entreprise de bonne foi pour retirer ou restreindre l'accès à des contenus que le fournisseur ou l'utilisateur considère comme répréhensible (« *objectionable* »), que ce contenu soit, ou non, constitutionnellement protégé. Le texte donne ainsi des exemples de contenus indésirables : il peut s'agir d'éléments jugés obscènes, lubriques, lascifs, dégoûtants, excessivement violents, constituant une forme de harcèlement ou, d'une manière ou d'une autre, répréhensible (*obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable* material). L'on rappelle souvent que le principe ici posé par la section 230(c)(2) ne fait que rappeler un droit garanti par le Premier Amendement à la Constitution, qui permet aux plateformes, en tant qu'entités privées, de choisir les contenus qu'elles entendent héberger ou publier, donc de modérer les contenus à leur discrétion. Le texte ne viendrait ici que rappeler un droit constitutionnel en en précisant l'implication, à savoir l'impossibilité d'engager la responsabilité d'une plateforme se livrant à la modération.

Bien qu'agissant dans leur droit, les plateformes qui pratiquent la modération font l'objet de critiques contradictoires. D'un côté, on leur reproche de ne pas suffisamment intervenir pour supprimer les contenus les plus contestables. De l'autre, elles sont accusées d'avoir des pratiques de modération peu transparentes, discriminatoires, voire attentatoires à la liberté d'expression. Sur le premier point, l'inaction des réseaux sociaux a fini par conduire les annonceurs eux-mêmes à renoncer à leurs campagnes de publicité en ligne afin d'inciter les plateformes à agir. En juin 2020, la campagne *#StopHateforProfit* a été lancée en réaction au refus de Facebook de supprimer une publication de Donald Trump, dans laquelle celui-ci menaçait les manifestants mobilisés à la suite de la mort de George Floyd d'une répression violente en reprenant les mots célèbres d'un policier raciste dans les années 1960 (« *When the looting starts, the shooting starts* »). Beaucoup d'entreprises, près de 1000 au total, dont Coca-Cola, Unilever, The North Face, Verizon et d'autres, ont, dans la foulée, suspendu leurs dépenses publicitaires sur Facebook et d'autres réseaux sociaux. Si la pression sur les plateformes pour que celles-ci luttent effectivement contre les contenus illicites et dangereux est si forte, c'est principalement parce que l'Etat n'a, en la matière, qu'une

capacité d'action limitée. La protection très étendue accordée à la liberté d'expression par le Premier Amendement à la Constitution empêche, précisément l'adoption de réglementations, fédérales ou étatiques, imposant davantage de restrictions que celles qui sont pour l'heure admises par la jurisprudence constitutionnelle, en particulier si ces contenus concernent le débat public. S'il est possible d'adopter des textes encadrant les discours les plus répréhensibles (pornographie infantile, diffamation, l'incitation à la violence ou à un acte criminel qui crée un risque réel et imminent d'action illégale etc), il convient, pour le reste, de faire prévaloir le « marché des idées » Dans un tel contexte, la lutte, nécessaire et effective, contre les discours de haine, le complotisme et la désinformation repose sur les épaules des plateformes : l'on dit parfois, à ce sujet, que la section 230, en consacrant l'immunité du « bon samaritain », a privatisé le droit de censurer faute, pour l'Etat, de pouvoir juridiquement l'exercer⁵⁵.

Sur le second point, et paradoxalement, les décisions de modération des plateformes et la manière dont elles sont mises en œuvre font l'objet d'intenses critiques. Les plus conservateurs s'estiment victimes de biais, voient dans les pratiques de modération des atteintes inacceptables à la liberté d'expression et réclament la consécration d'un principe de neutralité dans l'intervention des plateformes. Certains ont introduit des actions contestant la suspension de leur compte, à l'image de l'activiste d'extrême-droite Laura Loomer et du groupe conservateur Freedom Watch contre Google, Facebook, Twitter et Apple⁵⁶. Les recours sont fréquemment fondés sur le Premier Amendement à la Constitution, qui protège la liberté d'expression, et régulièrement rejetés. En effet, et même si la Cour Suprême a consacré le droit, pour les citoyens américains, d'échanger et débattre sur les réseaux sociaux⁵⁷, ce droit n'est pas opposable à un acteur privé. Le texte du Premier Amendement à la Constitution, en disposant que le Congrès ne peut faire aucune loi (« *shall make no law* ») restreignant la liberté d'expression, protège les citoyens américains contre les atteintes à leur liberté d'expression du fait de l'Etat et plus largement des pouvoirs publics. Il n'est, en revanche, pas opposable à une entreprise privée gérant un réseau social, qui n'est

⁵⁵ Philip Hamburger, "The Constitution can crack Section 230", Wall Street Journal, January 21 2021.

⁵⁶ *Freedom Watch, Inc. v. Google, Inc.*, 368 F. Supp. 3d 30 (D.D.C. 2019). La Cour Suprême fédérale a, le 5 avril 2021, refusé d'examiner le recours contre cette décision.

⁵⁷ *Packingham v North Carolina*, préc.

pas l'Etat⁵⁸. Il pourrait en être autrement, en vertu de la « state action doctrine », si les plateformes pouvaient être considérée comme exerçant de manière exclusive des fonctions traditionnellement réservées à l'Etat. L'on pourrait, à cet égard, affirmer que les réseaux sociaux assument des fonctions étatiques en ce qu'elles administrent des espaces publics (« *public fora* ») ouverts à tous les citoyens et au sein desquels ceux-ci peuvent s'exprimer⁵⁹. Telle ne semble pas être, à ce jour, l'analyse de la Cour Suprême qui a jugé en 2019 que le simple fait mettre des espaces de discussion à la disposition du public ne suffit pas à transformer une entité privée en un acteur étatique soumis aux contraintes du Premier Amendement⁶⁰. Bien que ce précédent n'ait pas concerné un réseau social mais une chaîne du câble, les recours exercés contre les décisions de modération prises par les réseaux sociaux sur le fondement du Premier Amendement sont généralement rejetés.

Il est vrai que ce qu'il est désormais convenu d'appeler la « déplateformisation » de Donald Trump en janvier 2021 a conduit à repenser à nouveaux frais la question de l'invocation du Premier Amendement à l'encontre d'une plateforme. Le débat est intense et porte à nouveau sur le point de savoir si la particularité des grands réseaux sociaux et du rôle qu'ils assument ne permettrait pas de leur appliquer la « *state action doctrine* » et de les contraindre à respecter des garanties constitutionnelles, *a fortiori* en présence de comptes détenus par des représentants de l'Etat. De fait, plusieurs juridictions fédérales ont jugé que les espaces de discussion créés et gérés par des représentants de l'Etat constituent des espaces publics (« *public fora* ») au sein desquels les échanges bénéficient de la protection accordée par le Premier Amendement. Il a, ainsi, été jugé à plusieurs reprises que les responsables publics ne peuvent empêcher les citoyens, en les « bloquant », d'accéder aux contenus qu'ils postent sur les réseaux sociaux et de les commenter⁶¹. Une Cour d'appel fédérale a, par exemple, décidé en

⁵⁸ *Tulsi Now, Inc. v. Google, LLC*, No. 2:19-cv-06444 (C.D. Cal. Mar. 3, 2020). Après avoir vu son recours rejeté, la démocrate Tulsi Gabbard a introduit au Congrès une proposition de loi intitulée « *Break Up Big Tech Act of 2020* », dont l'objet est de revenir sur l'immunité des plateformes lorsque celles-ci se comportent en éditeur et censurent certains utilisateurs.

⁵⁹ En mobilisant, pour les besoins de la cause, une jurisprudence ancienne relative à une ville créée et administrée par une entreprise : *Marsh v. Alabama*, 326, U.S. 501 (1946).

⁶⁰ *Manhattan Community Access Corp. v. Halleck*, 139 S. Ct. 1921, 1930 (2019).

⁶¹ *Davison v. Randall*, 912 F.3d 666 (4th Cir. 2019). Dans une décision rendue en janvier 2021, il a été jugé qu'une élue du Missouri pouvait bloquer des citoyens car elle n'usait son compte que pour sa campagne et non dans des fonctions officielles : *Campbell v. Reisch*, No. 19-2994 (8th Cir. Jan. 27, 2021).

2019 que Donald Trump, dans la mesure où il s'exprimait sur Twitter en tant que Président des Etats-Unis, ne pouvait exclure les utilisateurs américains sans violer leur droit d'accès à un espace public protégé par le Premier Amendement⁶². En août 2020, Trump avait décidé de porter l'affaire devant la Cour Suprême, qui a jugé, le 5 avril 2021, que l'affaire était désormais sans objet, Trump ayant quitté le pouvoir⁶³. Non seulement l'arrêt de la Cour d'appel a été invalidé et ne vaut plus comme précédent mais on ne dispose pas de la position de la Cour Suprême sur le point de savoir si les garanties du Premier Amendement peuvent être invoquées par les utilisateurs de Twitter, en particulier lorsque ceux-ci sont « bloqués » par le Président des Etats-Unis. En revanche, la décision de la Cour Suprême a permis au juge Clarence Thomas de publier une « *concurring opinion* » dans laquelle il doute de la pertinence de la qualification de « *public forum* » au motif que le pouvoir de Donald Trump sur son compte restait limité, Twitter conservant celui de supprimer le compte (ce qui a fini par arriver) : or n'est un « public forum » que ce qui est effectivement contrôlé par le gouvernement. Le juge Thomas en a profité pour insister tout particulièrement sur le pouvoir écrasant des plateformes de décider de ce qui peut figurer en ligne ou pas et sur la nécessité de déterminer les limites à apporter à ce pouvoir incroyable, laissé à des acteurs privés en situation de quasi-monopole. Selon Thomas, le droit des Etats-Unis comporte des mécanismes qui permettraient de limiter ce pouvoir, par exemple recourant à la qualification de « *common carriers* » (opérateurs publics), comme on l'a fait, par le passé dans le secteur des transports ou des télécoms, ou en qualifiant les plateformes de « *public accommodations* » (services publics). Ce faisant, le juge Thomas, qui ne paraît d'ailleurs pas penser que la faculté de modérer constitue, pour les plateformes, une prérogative constitutionnellement protégée, a utilisé l'occasion pour fournir divers arguments juridiques permettant de justifier une régulation des plateformes revenant sur l'immunité et la très grande liberté que leur laisse actuellement la section 230 du *Communications Decency Act*.

En tout état de cause, l'on peut douter, en l'état actuel des choses, que les garanties du Premier Amendement puissent être effectivement invoquées par les utilisateurs des plateformes à l'encontre des décisions de modération. Beaucoup ne le souhaitent

⁶² *Knight First Amendment Inst. at Columbia Univ. v. Trump*, 928 F.3d 226 (2d Cir. 2019)

⁶³ *Biden v Knight First Amendment Inst. at Columbia Univ.*, 593 U. S. (2021)

d'ailleurs pas. Certains soulignent que des plateformes contraintes par le Premier Amendement seraient tenues de laisser en ligne des contenus particulièrement toxiques et deviendraient de véritables cloaques⁶⁴. En outre, c'est aux tribunaux qu'il reviendrait, dans un tel contexte, de décider ce qui peut être ou non dit en ligne, une perspective à propos de laquelle l'on fait remarquer que les neuf juges de la Cour Suprême ne sont pas les mieux placés pour décider de la gestion des plateformes technologiques⁶⁵. Il vaudrait donc mieux écarter l'application des garanties constitutionnelles et permettre aux plateformes de se livrer elles-mêmes à une modération efficace en sanctionnant parfois les contenus et les points de vue les plus inacceptable et en protégeant les plus vulnérables⁶⁶. Dans le même temps, la section 230 pourrait être amendée de manière à imposer de nouveaux devoirs aux gardiens de ces espaces virtuels que sont les réseaux sociaux⁶⁷. L'on pourrait songer, en particulier à imposer des garanties procédurales (« *due process* »)⁶⁸ aux plateformes qui se livrent à la modération, dans la ligne des « principes de Santa Clara sur la transparence et la responsabilité dans la modération des contenus » (*Santa Clara Principles on Transparency and Accountability in Content Moderation*). Ces principes, dont la première version a été rédigée en 2018 par un groupe d'experts et d'universitaires, prévoient, entre autres choses, la divulgation, par les plateformes, du nombre de publications supprimées et de comptes suspendus, la communication des motifs d'une décision de modération à la personne concernée et la possibilité d'exercer un recours contre tout retrait de contenu ou toute suspension de compte.

Vers une réforme de la section 230 du *Communications Act of 1934* ?

Nombreux sont ceux qui proposent une réforme en profondeur de la section 230. L'idée serait à la fois de limiter l'immunité dont bénéficient les plateformes quant aux contenus les plus répréhensibles et de réguler les pratiques de modération. Les perspectives varient souvent selon la place, sur l'échiquier politique, de ceux qui

⁶⁴ Daphne Keller, "Who decides what stay on the Internet?", The Verge, January 12, 2021

⁶⁵ Genevieve Lakier, "The Great Free-Speech Reversal", The Atlantic, January 27, 2021

⁶⁶ Mary Anne Franks, "The Free Speech Black Hole: Can the Internet Escape the Gravitational Pull of the First Amendment?", Knight First Amendment Institute at Columbia University, August 21, 2019

⁶⁷ Danielle Citron and Mary Anne Franks, "Cyber Civil Rights in the Time of COVID-19", Harvard Law Review Blog, May 14, 2020

⁶⁸ A. Robertson, "Human rights groups want Facebook to offer users "due process" for takedowns", The Verge, November 14, 2018

appellent à la réforme. Les démocrates sont généralement favorables à une responsabilité accrue des plateformes afin de les inciter à lutter contre les contenus illégaux ou toxiques : au cours de la campagne présidentielle, Joe Biden a déclaré qu'il entendait revenir sur l'immunité de la section 230 en particulier dans les cas où les plateformes hébergent délibérément des informations fausses⁶⁹. De leur côté, les républicains se plaignent des pratiques actuelles de modération qu'ils jugent biaisées et souhaitent limiter, au moins en partie, l'immunité du « bon samaritain » de manière à encadrer davantage la modération en imposant un principe de neutralité. C'est dans cette optique que Donald Trump, à la suite des premiers signalements de ses tweets par Twitter, avait adopté, en mai 2020, un *Executive Order*⁷⁰ destiné à priver d'immunité les plateformes se livrant à des pratiques de modération discriminatoires ou déceptives. Pour ce faire, le texte enjoignait, entre autres choses, à la *Federal Communications Commission* de « clarifier » le texte de la section 230 en précisant notamment les conditions dans lesquelles une plateforme est réputée agir « de bonne foi », de manière à priver d'immunité les plateformes agissant de manière déloyale. La demande de Trump n'a toutefois pas pu être satisfaite avant son départ de la Maison Blanche.

Quoi qu'il en soit, de nombreuses propositions de réforme ont été émises par les experts et les universitaires, qui s'inquiètent non seulement de la prolifération des contenus toxiques, du harcèlement en ligne et des discours de haine mais aussi des difficultés posées par la désinformation et son influence éventuelle sur des élections démocratiques. Par exemple, selon la professeure Danielle Citron⁷¹, l'interprétation trop large donnée par les tribunaux à la section 230 justifierait de modifier le texte pour conditionner l'immunité à des « pratiques raisonnables de modération des contenus », ce qui permettrait sans doute de revenir à l'intention originelle des rédacteurs du texte. Les plateformes ne prenant pas de mesures raisonnables pour lutter contre les utilisations illégales du service en ligne créant un préjudice grave pour les personnes pourraient être qualifiées d'éditeurs et jugée responsables à ce titre. Les plateformes persistant à laisser en ligne des contenus répréhensibles ou dangereux seraient ainsi responsables, le risque étant toutefois qu'elles soient incitées à censurer beaucoup

⁶⁹ R. Lerman, "Social media liability law is likely to be reviewed under Biden", *The Washington Post*, January 18, 2021

⁷⁰ *Executive Order on Preventing Online Censorship*, May, 28, 2020

⁷¹ Danielle Citron and Benjamin Wittes, "The Internet Will Not Break: Denying Bad Samaritans Section 230 Immunity," *Fordham Law Review*, 2017.

trop largement les contenus douteux, au détriment de la liberté d'expression. Dans le prolongement de cette proposition, il a, été proposé de rendre les plateformes responsables lorsqu'elles s'abstiennent d'intervenir alors qu'elles ont connaissance de faits répréhensibles ou lorsqu'elles en tirent profit, voire les sollicitent⁷². L'immunité devrait être réduite pour ne viser que les seules hypothèses dans lesquelles un contenu (« *speech* ») a été mis en ligne par un tiers sans que la plateforme n'encourage, ne sollicite ou n'utilise intentionnellement un contenu ou un comportement illicite pour en tirer profit⁷³. Une idée proche, émise dans un *policy brief* publié par le Stigler Center de l'Université de Chicago⁷⁴, consiste simplement à appliquer aux réseaux sociaux le régime des éditeurs de contenu dans tous les cas où ils font activement la promotion de contenus et gagnent de l'argent grâce à ces contenus, comme le font, par exemple, Facebook ou YouTube. Un rapport également rédigé sous les auspices du Stigler Center s'est, de son côté, prononcé en faveur d'une logique de contrepartie (« *quid pro quo* »)⁷⁵. Les plateformes auraient le choix entre accepter des obligations supplémentaires ou renoncer à la protection offerte par la section 230, en partie ou en totalité. Dans ce cadre, l'on pourrait exiger qu'elles garantissent que leurs algorithmes ne privilégient pas les contenus extrêmes et toxiques aux fins de stimuler l'engagement des utilisateurs, qu'elles soient contraintes à davantage de transparence quant à leurs pratiques publicitaires et leurs politiques de modération, voire qu'elles soient obligées de consacrer un petit pourcentage de leurs revenus annuels à un fonds soutenant la presse traditionnelle. Enfin, un récent rapport⁷⁶ reprend plusieurs de ces suggestions, telle l'idée de conditionner l'immunité à l'adoption de mesures raisonnables pour lutter contre les contenus illégaux ou à des contreparties de la part des plateformes.

⁷² Mary Anne Franks, "Reforming Section 230 and Platform Liability", Stanford Cyber Policy Series (2021); Danielle Keats Citron & Mary Anne Franks, "The Internet as a Speech Machine and Other Myths Confounding Section 230 Reform," University of Chicago Legal Forum 3 (2020): 70.

⁷³ Mary Anne Franks, "Reforming Section 230 and Platform Liability", Stanford Cyber Policy Series (2021).

⁷⁴ L. Zingales, F.M. Lancieri, *Stigler Committee on Digital Platforms: Policy Brief*, September 2019.

⁷⁵ Stigler Center, Committee for the Study of Digital Platforms, *Media Subcommittee Report*, 1 July 2019

⁷⁶ Paul M. Barrett, *Regulating Social Media: The Fight Over Section 230- and Beyond*, NYU Stern, September 2020.

Dans un tel contexte, le législateur fédéral devrait finir, tôt ou tard, par intervenir. Le *Department of Justice* de l'administration Trump avait transmis au Congrès un projet de réforme prévoyant d'écarter l'immunité des « mauvais samaritains » qui facilitent ou sollicitent en connaissance de cause des contenus préjudiciables ou hébergent des contenus particulièrement graves (maltraitance des enfants, terrorisme, cyberharcèlement). Le texte visait également à rendre responsables les plateformes se livrant « de mauvaise foi » à la censure de contenus licites⁷⁷. De nombreux textes ont, en outre, été déposés par des membres du Congrès. Plusieurs propositions de loi républicaines, généralement déposées par le sénateur Hawley ont prévu de revenir sur l'immunité des plateformes lorsqu'elles n'établissent pas que leurs algorithmes et pratiques de modération sont politiquement neutres⁷⁸, ne s'engagent pas à gérer la plateforme de bonne foi dans leurs conditions générales⁷⁹, ou diffusent massivement des publicités ciblées (*targeted advertising*)⁸⁰. Côté démocrate, l'on peut retenir l'initiative de la représentante démocrate Anna Eshoo, qui a déposé une proposition de loi⁸¹, réintroduite en mars 2021, excluant l'immunité en présence d'algorithmes qui promeuvent ou recommandent des contenus relatifs à des violations des droits civils ou à des actes de terrorisme international. Le SAFE Tech Act⁸², déposée en février 2021 par plusieurs parlementaires démocrates (M.R. Warner, M. Hirono et A. Klobuchar), rend les plateformes responsables des contenus sponsorisés et payés par des tiers, comme les publicités. Il permet également d'engager leur responsabilité en cas de cyber-harcèlement, de harcèlement ciblé, de discrimination ou de mort injustifiée, d'engager des poursuites contre les plateformes pour violation des droits de l'homme à l'étranger et d'obtenir des injonctions dans les cas où un "préjudice irréparable" pourrait survenir.

Mais ce sont surtout les propositions de loi bipartisanes qui retiennent l'attention. Une proposition bipartisane datant de juillet 2020, la loi EARN IT ⁸³ prévoit de supprimer

⁷⁷ Department of Justice's review of Section 230 of the Communications Decency Act, 23 septembre 2020

⁷⁸ *Ending Support for Internet Censorship Act*, juin 2019

⁷⁹ The "Limiting Section 230 Immunity to Good Samaritan's Act", juin 2020

⁸⁰ The "Behavioral Advertising Decisions Are Downgrading Services (or BAD ADS) Act », juillet 2020.

⁸¹ *Protecting Americans from Dangerous Algorithms Act*, octobre 2020

⁸² *Safeguarding Against Fraud, Exploitation, Threats, Extremism, and Consumer Harms Act*, février 2021

⁸³ *Eliminating Abusive and Rampant Neglect of Interactive Technologies*, juillet 2020

l'immunité des plateformes en matière d'exploitation sexuelle des enfants, sans toutefois retoucher quoi que ce soit d'autre dans la section 230. *Last but not least*, le *PACT Act*⁸⁴, déposé en juin 2020 a prévu une réforme substantielle et audacieuse, qui mettrait fin à l'immunité accordée par la section 230(c)(1) lorsque la plateforme ne retire pas un contenu jugé illégal par un tribunal. Le texte prévoit que les plateformes comptant plus d'un million d'utilisateurs mensuels ou ayant un chiffre d'affaires annuel supérieur à 25 millions de dollars doivent retirer le contenu dans les 24 heures suivant la notification les informant que le contenu a été jugé illégal. Les plateformes plus petites doivent retirer les contenus "dans un délai raisonnable en fonction de la taille et de la capacité du fournisseur". La proposition de loi PACT oblige par ailleurs les plateformes à créer des mécanismes de traitement des plaintes relatives aux contenus répréhensibles et aux décisions de modération. Le texte exige que les plateformes agissent rapidement suite à un signalement, précisent les raisons de leurs décisions de modération, permettent aux utilisateurs de les contester, et publient des rapports de transparence concernant leur politique de modération. En particulier, les grandes plateformes répondent aux plaintes des utilisateurs dans un délai de 14 jours, tandis que les petites plateformes disposent d'un délai non limité qui dépend de leur taille et de leur capacité. Le respect de ces obligations doit être contrôlé par la *Federal Trade Commission*. La proposition de loi *PACT* a été critiquée⁸⁵ : on lui reproche de ne pas suffisamment préciser la nature des décisions judiciaires ayant jugé un contenu illégal (doivent-elles être irrévocables ?) et de laisser au final les plateformes décider elles-mêmes des contenus à retirer. On souligne également que le texte favoriserait ici les plus grandes plateformes (Facebook, YouTube, Twitter) qui ont les moyens d'employer des modérateurs, au détriment des plus petites. En mars 2021, une nouvelle version du *PACT Act* a été réintroduite au Congrès, dans laquelle le délai donné aux plateformes pour retirer un contenu jugé illégal par une décision judiciaire est porté à quatre jours. Le texte renforce l'obligation de publication, par les plateformes, de rapports de transparence, désormais semestriels, et prévoit la possibilité, pour les autorités fédérales, d'engager des poursuites civiles contre les plateformes. Il est à noter que le *PACT Act* est l'un des rares textes prévoyant des

⁸⁴ *Platform Accountability and Consumer Transparency Act*, juin 2020

⁸⁵ A. Mackey, "The PACT Act's Attempt to Help Internet Users Hold Platforms Accountable Will End Up Hurting Online Speakers", Electronic Frontier Foundation, July 21, 2020

obligations asymétriques avec des obligations sont allégées pour les plateformes ayant moins d'un million d'utilisateurs ou de \$25 millions de chiffre d'affaires.

Il ressort de l'ensemble de ces projets que la section 230 ne devrait pas être supprimée mais sera probablement modifiée à plus ou moins long terme. Cette perspective de réforme fait toutefois l'objet d'un débat intense. Les représentants des grandes plateformes⁸⁶ ont exprimé leurs réserves. Jack Dorsey préfère une stratégie de décentralisation permettant de confier le pouvoir aux utilisateurs à une intervention gouvernementale et vantait, quelques jours après la suspension du compte de Trump, la perspective d'un réseau social conçu sur le modèle de la plateforme Bitcoin. Sundar Pichai, président de Google, a insisté sur le rôle fondateur de la section 230 dans le web ouvert contemporain et sur le risque que son abrogation ou sa modification empêche les plateformes de gérer les contenus de manière responsable en garantissant l'accès à l'information et la liberté d'expression, par exemple en les contraignant à une censure trop poussée. Il promeut donc, plutôt qu'une réforme de cette disposition, la mise en place de pratiques de modération plus transparentes, équitables et efficaces. Mark Zuckerberg a suggéré de ne pas abroger la section 230 mais admet la possibilité d'une modification du texte afin de subordonner le bénéfice de l'immunité à la mise en place, par les plateformes, de bonnes pratiques de modération en vue d'identifier et retirer les contenus illicites, sans pour autant les contraindre à supprimer l'intégralité des publications répréhensibles, un objectif qu'il juge impossible à atteindre. Il a également évoqué la possibilité de faire dépendre les obligations des plateformes de leur taille et de les faire fixer par un organisme tiers garantissant un traitement équitable entre les entreprises. Dans les milieux de la technologie, certaines organisations⁸⁷, comme l'*Electronic Frontier Foundation* (EFF)⁸⁸ militent en faveur du statu quo au nom de la liberté d'expression et de la protection des minorités. Il est également affirmé qu'une réforme pénaliserait l'ensemble du secteur de la Tech, à commencer par les petites entreprises, ou rendrait plus difficile

⁸⁶ Mark Zuckerberg (Facebook), Sundar Pichai (Google) et Jack Dorsey (Twitter) ont délivré des contributions à l'occasion de leur audition au Congrès le 25 mars 2021, consultables sur le site du Congrès.

⁸⁷ On peut citer, parmi d'autres : Common Cause, Fight for the Future, National Center for Lesbian Rights, Wikimedia Foundation, MediaJustice.

⁸⁸ E. Harmon, "Don't Blame Section 230 for Big Tech's Failures. Blame Big Tech", EFF blog, November 16, 2020; J. Kelly, "Section 230 is Good, Actually", EFF Blog, December 3 2020

la suppression des contenus dangereux ayant conduit à l'attaque du Capitole. Mieux vaudrait donc recourir au droit antitrust de manière à assurer la concurrence en empêchant l'apparition de quasi-monopoles au pouvoir écrasant. L'on pourrait également agir sur ce qui est au cœur du modèle d'affaire des plateformes, le « capitalisme de la surveillance » (« *surveillance capitalism* »)⁸⁹, à savoir la collecte et le traitement des données des utilisateurs, que l'on devrait entourer de garanties supplémentaires. L'on pourrait, par ailleurs, se pencher sur les algorithmes que les plateformes utilisent pour administrer les sites et s'assurer que la configuration de ces algorithmes ne conduit pas à « pousser » les contenus les plus répréhensibles. L'on pourrait, enfin, édicter des exceptions à l'immunité de la section 230 en présence de comportements particulièrement graves, à l'image de ce qui a été fait, en 2016, dans le cas du trafic sexuel⁹⁰. Au fond, il est difficile de savoir si les défenseurs de la section 230 souhaitent s'en tenir à la logique du « marché des idées » et d'un Internet sans entrave ou voir les plateformes prendre seule la responsabilité de se livrer à une modération désormais indispensable. Il n'est, dans tous les cas, pas sûr qu'une telle approche permette de régler durablement les difficultés majeures soulevées par la prolifération des contenus répréhensibles.

2/ Une responsabilité relative des plateformes en Europe

Depuis l'adoption, en 2000, de la Directive E-Commerce 2000/31/CE, les règles relatives à la responsabilité des plateformes sont un peu moins favorables aux plateformes qu'elles ne le sont aux Etats-Unis. Il reste que règles européennes prévoient une immunité des plateformes tout en les incitant davantage à la modération. Ce régime issu de la Directive adopté en 2000 devrait toutefois prochainement faire l'objet d'une réforme en profondeur du fait de l'adoption prochaine du Règlement *Digital Services Act* proposé par la Commission le 15 décembre 2020.

Le régime issu de la Directive E-Commerce de 2000

En Europe, comme aux Etats-Unis, les réseaux sociaux bénéficient de règles favorables et ne sont pas civilement ou pénalement responsables du fait des publications des utilisateurs, dans la mesure où ils bénéficient du régime applicable

⁸⁹ S. Zuboff, « The Age of Surveillance Capitalism », Public Affairs, 2019

⁹⁰ La loi "Stop Enabling Sex Traffickers Act" (SESTA) de 2018 a modifié la section 230 en rendant les sites web responsables de l'hébergement délibéré de contenus liés à la traite des êtres humains.

aux hébergeurs⁹¹. La directive 2000/31/CE a en effet prévu, dans son article 14 que l'hébergeur d'un contenu illicite n'en est pas responsable dès lors qu'il n'avait pas effectivement connaissance de l'activité ou de l'information illicites et que, qu'il a agi promptement pour retirer ou rendre l'accès au contenu illicite impossible dès qu'il a eu connaissance de l'illicéité. Cela signifie que la plateforme est supposée agir vite dès lors qu'un contenu est signalé, et qu'elle est incitée à la modération, voir à une surmodération pour éviter d'engager sa responsabilité. L'article 14 §3 prévoit d'ailleurs que la possibilité, pour une juridiction ou une autorité administrative nationale d'exiger de la plateforme une telle intervention. En revanche, l'article 15 de la directive exclut toute obligation générale de surveillance des réseaux qui pèserait sur les plateformes, même s'il est possible aux Etats d'imposer spécifiquement la surveillance de certains contenus (voir en matière de terrorisme). Cette absence d'obligation générale de surveillance doit toutefois être articulée avec l'obligation de retirer ou rendre inaccessible un contenu illicite : la Cour de Justice de l'Union Européenne a par exemple jugé que la plateforme doit retirer ou rendre inaccessible tous les contenus identiques ou équivalents (au sens où le message véhiculé est équivalent) à un contenu qu'une juridiction a jugé diffamatoire⁹².

L'émergence des plateformes d'intermédiation et des réseaux sociaux au cours des dix dernières années a conduit la Commission, consciente que ces règles de responsabilité particulièrement favorables a permis le développement des acteurs d'aujourd'hui, à s'interroger sur une éventuelle réforme des règles applicables aux plateformes, qui ont été conçues pour l'Internet des années 2000. Dans une communication de 2016⁹³, elle a, cependant, insisté que les limites à l'immunité prévues par le droit de l'Union incitaient les plateformes à la modération, avec évidemment à la clé le risque d'une modération trop poussée réalisée en toute opacité par des algorithmes. La Commission⁹⁴ a, depuis lors, insisté sur la nécessité, pour les plateformes de lutter contre les activités les plus condamnables, et adopté des recommandations sur ce point, tant à l'intention des Etats que des plateformes, de

⁹¹ Art. 14 de la Directive « e-commerce » n°2000/31CE sur le commerce électronique, transposée en droit français par la loi n°2004-575 du 21 juin 2004 pour la confiance dans l'économie numérique (art.6 et 6-1).

⁹² CJUE 3 octobre 2019, *Facebook Ireland Limited c/ Eva Glawishnig-Piesczek*, aff. C-18/18.

⁹³ Communication du 26 mai 2016, COM (2016) 288 final.

⁹⁴ Communication du 28 septembre 2017, COM (2017) 555 final.

manière à les inciter à lutter contre les contenus illicites dans le respect des droits fondamentaux⁹⁵. Parallèlement, un code de conduite relatif à la lutte contre les discours de haine en ligne a été adopté en 2016 (*Code of Conduct on Countering Illegal Hate Speech Online*) dont les entreprises signataires (Facebook, Microsoft, YouTube, Twitter, Tik Tok) se sont engagées à lutter contre la prolifération des discours de haine à caractère raciste et xénophobe en ligne. Cette première initiative a été complétée, en 2018, par une Code pratique contre la désinformation en ligne (*Code of Practice on Disinformation*), dont les signataires (Facebook, Google, Twitter, Mozilla, Microsoft, Tik Tok) se sont engagées à respecter certains standards dans la lutte contre la désinformation. La stratégie des autorités européennes a donc, jusqu'à une période récente, privilégiée des approches non contraignantes et fondées sur une démarche volontaire de la part des plateformes.

En France, la Directive « e-commerce » a été transposée par la loi n°2004-575 du 21 juin 2004 pour la confiance dans l'économie numérique. Comme la directive, ce texte prévoit que les plateformes ne peuvent pas voir leur responsabilité civile (ou pénale) engagée du fait des activités ou des informations stockées en ligne « si elles n'avaient pas effectivement connaissance de leur caractère illicite ou de faits et circonstances faisant apparaître ce caractère ou si, dès le moment où elles en ont eu cette connaissance, elles ont agi promptement pour retirer ces données ou en rendre l'accès impossible » (art. 6 I-2 LCEN). Le Conseil Constitutionnel a, à cet égard, jugé que « ces dispositions ne sauraient avoir pour effet d'engager la responsabilité d'un hébergeur qui n'a pas retiré une information dénoncée comme illicite par un tiers si celle-ci ne présente pas manifestement un tel caractère ou si son retrait n'a pas été ordonné par un juge »⁹⁶. Aussi, la responsabilité du fournisseur d'hébergement ne peut être mise en oeuvre qu'à des conditions strictes, dans les cas où il est resté inactif en présence d'un contenu manifestement illicite en ayant connaissance de l'illicéité. Cette connaissance peut découler d'une décision judiciaire, mais il sera plus difficile de l'établir si un simple signalement a été fait à la plateforme, car il ne découle pas forcément d'un signalement qu'un contenu est « manifestement illicite ». En tout cas, la loi prévoit que les plateformes doivent mettre en place des modalités de signalement

⁹⁵ Recommandation (UE) 2018/334 de la Commission du 1er mars 2018 sur les mesures destinées à lutter, de manière efficace, contre les contenus illicites en ligne.

⁹⁶ Décision n° 2004-496 DC du 10 juin 2004 : Loi pour la confiance dans l'économie numérique

des contenus douteux Il est par ailleurs possible à l'autorité administrative de demander aux hébergeurs de retirer, dans un délai de 24h, les contenus à caractère pédopornographique ou terroriste (art. 6-1 LCEN). De son côté, le juge judiciaire peut prescrire toute mesure propre à prévenir ou faire cesser un dommage occasionné par un contenu (art. 6-I 8^e). Surtout, les plateformes doivent transmettre aux autorités publiques les éléments permettant de sanctionner les infractions les plus graves (apologie des crimes contre l'humanité, provocation au terrorisme, incitation à la violence etc) (art. 6-I 7^e). Il reste que les plateformes rechignent parfois à collaborer avec les autorités : tout récemment Twitter a refusé de répondre aux réquisitions du Parquet dans des affaires de harcèlement, ce qui a empêché l'identification des harceleurs⁹⁷.

Durant la période la plus récente, plusieurs lois ont été votées en France aux fins de renforcer la lutte contre les contenus illicites qui prolifèrent sur les réseaux sociaux. La loi du 22 décembre 2018 relative à la manipulation de l'information a prévu la possibilité de saisir le juge en référé, en période électorale, en présence d'une désinformation intentionnelle et de nature à altérer la sincérité du scrutin (article L 163-2 du Code électoral). En outre, dans sa version originelle, la loi dite « Avia » visant à lutter contre les contenus haineux, votée le 13 mai 2020 et promulguée le 24 juin 2020⁹⁸, renforçait significativement les obligations des plateformes. Le texte prévoyait, en effet, que les plateformes devaient retirer ou rendre inaccessibles, dans un délai de 24h suivant notification par tout internaute, certains contenus manifestement illicites comme les incitations à la haine ou les injures à caractère raciste ou anti-religieuses. Les sanctions prévues pouvaient aller jusqu'à une amende de 250 000 euros par manquement, montant pouvant être multiplié par cinq en présence d'une personne morale. La loi imposait, en outre, le retrait, dans le délai d'une heure après notification par l'autorité administrative, des contenus à caractère terroriste ou pédopornographique. La loi plaçait, enfin, les plateformes sous le contrôle du CSA, qui pouvait prononcer une sanction pécuniaire pouvant aller jusqu'à 20 millions d'euros ou 4% du chiffre d'affaires annuel mondial total de la plateforme. La majeure partie des

⁹⁷ « Twitter poursuivi en France pour son manque de coopération dans des affaires de haine en ligne », Le Monde, 3 février 2021

⁹⁸ L. n° 2020-766, 24 juin 2020, visant à lutter contre les contenus haineux sur internet, JO 25 juin.

dispositions de la loi Avia a cependant été censuré par le Conseil Constitutionnel, dans une décision du 18 juin 2020⁹⁹, sur le fondement de l'article 11 de la Déclaration des Droits de l'Homme et du Citoyen. Le Conseil Constitutionnel a admis que la diffusion d'images pornographiques représentant des mineurs et la provocation à des actes de terrorisme ou l'apologie de tels actes constituaient des abus portant gravement atteinte à l'ordre public et aux droits des tiers. Cependant, soulignant que le droit à la liberté d'expression implique la liberté d'accéder aux services de communication en ligne, le Conseil Constitutionnel a censuré les dispositions considérées au motif notamment que la détermination du caractère illicite ou manifestement illicite des contenus en cause relevait de la seule appréciation de l'administration ou des plateformes. Il a, en particulier, relevé que ces dispositions ne peuvent qu'inciter les opérateurs de plateforme en ligne à retirer les contenus qui leur sont signalés et portent donc une atteinte à l'exercice de la liberté d'expression et de communication qui n'est pas nécessaire, adaptée et proportionnée. Compte-tenu de cette décision, la loi du 24 juin 2020 a, pour finir, surtout créé un observatoire de la haine en ligne, prévu que les contenus pouvant engager la responsabilité des plateformes soient « manifestement illicites » comme l'avait jugé le Conseil Constitutionnel et augmenté de 75000 euros à 250000 euros le montant des amendes encourues par les plateformes ne respectant pas leurs obligations.

Les textes en cours : le futur *Digital Services Act* et la loi française confortant le respect des principes de la République

Après avoir longtemps privilégié les normes de *soft law*, les autorités européennes ont modifié leur stratégie au cours de l'année 2020 pour opter en faveur d'une véritable réforme des dispositions de la Directive E-Commerce 2000/31/CE. Elles n'ont pas donné suite, à cet égard, aux propositions de Facebook qui a, en février 2020, publié un White Paper intitulé « *Charting the Way Forward: Online Content Regulation* ». Ce texte appelait à une régulation des plateformes devant les aider à distinguer les discours acceptables de ceux qui ne le sont pas et, idéalement, constituer un standard à la fois global et flexible, qui tienne compte des préférences des utilisateurs. Une telle régulation, disait Facebook, pourrait permettre d'assurer une modération des contenus

⁹⁹ Cons. Const. 18 juin 2020, n°2020-801 DC JO 25 juin.

dans le respect de la liberté d'expression, en prévoyant des obligations de transparence relatives aux stratégies de modération ou aux algorithmes utilisés, des modalités de signalement des contenus répréhensibles et la possibilité de contester les décisions de modération. S'agissant de la responsabilité des plateformes, Facebook suggérait qu'elles ne devraient être tenues que d'une simple « obligation de moyen », donc d'un objectif global de retrait d'un certain pourcentage de contenus illicites, au motif que la modération est particulièrement difficile sur des plateformes mondiales comportant des milliards d'utilisateurs et de publications. La loi pourrait donc établir des normes de prévalence¹⁰⁰ (définie comme la fréquence à laquelle les contenus délétères sont effectivement consultés, même après que les modérateurs ont essayé de les éliminer) et les plateformes feraient, en cas de dépassement des seuils légaux, l'objet d'une surveillance accrue, de plans d'amélioration spécifiques ou, en cas d'échecs systématiques répétés, d'amendes. Il n'y aurait pas lieu, en revanche, de contraindre directement les plateformes à retirer, dans un délai réduit, les contenus illicites ou de les rendre responsables des contenus publiés par les utilisateurs. Bien qu'intéressante, cette approche -qui consiste tout de même à inciter les plateformes à réduire les catégories de contenus qu'elles ont elles-mêmes identifiées comme nuisibles- a été jugée insuffisante par les autorités européennes qui ont souligné la nécessité de mieux connaître les algorithmes utilisés pour administrer et modérer les plateformes¹⁰¹.

Le 15 décembre 2020, la Commission a publié la proposition de Règlement sur les services numériques (*Digital Services Act*) qui vise principalement à moderniser les dispositions de la Directive E-Commerce 2000/31/CE du 8 juin 2000. Le texte n'abandonne pas le principe selon lequel les prestataires de services numériques ne sauraient être responsables des contenus illicites publiés ou transmis par des tiers dès lors qu'ils n'ont pas connaissance de leur caractère illicite ou ont agi promptement pour les retirer ou rendre leur accès impossible lorsqu'ils en avaient connaissance. Est

¹⁰⁰ Facebook publie régulièrement des chiffres de prévalence : en novembre 2020, la plateforme annonçait que la prévalence des contenus haineux était de 10 à 11 vues de contenus haineux pour 10000 vues de contenus. En février 2021, Facebook a annoncé que plus de 97 % des contenus classés comme discours haineux étaient détectés par son logiciel avant d'être signalés par un utilisateur, et que 49 % des contenus d'intimidation et de harcèlement sur son réseau social avaient été traités avant d'être signalés par les utilisateurs durant le quatrième trimestre 2020 (contre 26 % au trimestre précédent).

¹⁰¹ "Brussels pushes back on Zuckerberg pitch", Politico, February 17, 2020.

également repris le principe selon lequel ces prestataires ne sont pas tenus de procéder à une surveillance généralisée des contenus mis en ligne ou de réaliser des vérifications. La proposition de Règlement précise d'ailleurs que les hébergeurs prenant l'initiative de modérer les contenus ou de se livrer à des vérifications relatives aux informations qui y sont publiées n'en deviennent pas, pour autant, responsables. Dans le même temps, l'immunité reconnue aux entreprises technologiques est désormais assortie de nouvelles obligations destinées à renforcer la lutte contre les comportements illicites, à apporter des garanties aux utilisateurs des services numériques et à maîtriser les risques présentés par les très grandes plateformes. Il est intéressant de noter que ces nouvelles obligations sont imposées de manière asymétrique, ce qui signifie qu'elles varient selon les caractéristiques des prestataires de services numériques considérés. Les réseaux sociaux relèvent ainsi de la catégorie des plateformes en ligne mettant en relation des utilisateurs et certains d'entre eux appartiennent également à la catégorie des très grandes plateformes d'au moins 45 millions d'utilisateurs actifs mensuels qui présentent, du fait de leur taille, des risques systémiques. Ils seront, à cet égard, placés sous le contrôle des Coordinateurs nationaux des services numériques qui prononceront les sanctions prévues par les droits nationaux dans les limites de certains plafonds, le *Digital Services Act* renvoyant aux Etats-membres le soin de faire respecter les obligations qu'il édicte.

Les réseaux sociaux seront soumis à des obligations destinées à permettre une modération plus effective et efficace. Ils devront mettre en place des mécanismes accessibles permettant à leurs utilisateurs de signaler la présence de contenus qu'ils jugent illicites. Les signalements provenant des « *trusted flaggers* » (« tiers de confiance ») désigné comme tels par le Coordinateur national des services numériques devront être traités en priorité. Ils devront coopérer avec les autorités compétentes et donner suite aux instructions des autorités judiciaires ou administratives nationales d'agir à l'encontre d'un contenu illicite. Elles devront notamment communiquer aux autorités qui le demandent les informations relatives à un utilisateur et les informer rapidement de tout contenu laissant suspecter la commission passée, présente ou future d'une infraction pénale grave mettant en jeu la vie ou la sécurité des personnes. Enfin, la proposition de Règlement consacre une possibilité de sanction des utilisateurs récalcitrants déjà présente dans les conditions générales des plateformes, tout en encadrant leur marge de manoeuvre: les

internauts publiant fréquemment des contenus illicites pourront faire l'objet d'un avertissement, puis voir leur compte suspendu pour une période raisonnable. De leur côté, utilisateurs bénéficient de garanties renforcées. Ils devront être pleinement informés des pratiques de modération de leurs prestataires, qui devront préciser clairement, dans leurs conditions générales, les éventuelles restrictions pouvant affecter l'utilisation de leurs services. Les plateformes devront agir de manière diligente, objective et proportionnée dans la mise en oeuvre des restrictions prévues et publier chaque année des rapports relatifs à leurs pratiques de modération. Par ailleurs, les décisions de retrait d'un contenu ou de suspension ou de résiliation du compte de l'utilisateur ou de son accès au service devront être motivées et pouvoir être contestées. En cas de désaccord persistant avec la plateforme, les utilisateurs devront pouvoir porter leurs contestations devant un organisme de règlement alternatif des litiges certifié par le Coordinateur national des services numériques.

Les très grandes plateformes d'au moins 45 millions d'utilisateurs sont soumises par la proposition de Règlement à des obligations supplémentaires relativement lourdes compte-tenu des risques systémiques qu'elles engendrent. Elles seront placées sous le contrôle de la Commission, qui pourra procéder à des investigations et des contrôles, adopter des décisions de non-conformité et prononcer des amendes n'excédant pas 6% du chiffre d'affaire de l'exercice antérieur, voire 1% pour certaines infractions. Dans ce cadre, les très grandes plateformes devront régulièrement évaluer ces risques, notamment ceux qui découlent de la diffusion de contenus illicites, envisager les effets négatifs du fonctionnement de leur plateforme sur les droits fondamentaux, et se pencher sur les effets réels ou possibles d'une manipulation intentionnelle de leurs services sur la protection de la santé publique, des mineurs, du débat public, du processus électoral et de la sécurité publique. Elles devront tout particulièrement évaluer l'impact de leurs systèmes de modération, de recommandation, et de ciblage publicitaire, prendre des mesures raisonnables et efficaces visant à atténuer les risques identifiés, désigner un « *compliance officer* » et faire l'objet d'audits annuels. Elles devront se montrer particulièrement transparentes quant à leurs pratiques de publicité en ligne et leurs systèmes de recommandation. Enfin, le texte paraît donner aux autorités européennes et nationales la possibilité de se faire d'obtenir des informations relatives aux algorithmes utilisés : le texte prévoit que ces autorités peuvent obtenir, le cas échéant au profit de chercheurs agréés, « les

données nécessaires pour évaluer les risques et les éventuels préjudices causés par les systèmes de la plateforme, les données concernant l'exactitude, le fonctionnement et les tests des systèmes algorithmiques de modération de contenu, des systèmes de recommandation ou des systèmes de publicité, ou encore les données concernant les processus et les résultats de la modération de contenu ou des systèmes internes de traitement des réclamations au sens du présent règlement » (Considérant 64).

Il est intéressant ici de souligner que le législateur français souhaite adopter par anticipation les règles figurant dans la proposition de Règlement de la Commission. Le projet de loi confortant le respect des principes de la République voté en première lecture le 16 février 2021 par l'Assemblée nationale vise particulièrement les plateformes de partage de contenu et les sites de référencement en reprenant un grand nombre d'obligations prévues par le DSA. L'initiative peut étonner en présence d'un texte européen qui n'a pas été adopté et pourrait encore être modifié. Le législateur français veut manifestement s'assurer que les règles envisagées s'appliqueront chez nous au plus vite : la plupart des dispositions prévues ont vocation à s'appliquer jusqu'à l'entrée en vigueur du DSA et au plus tard jusqu'au 31 décembre 2023, date à laquelle les dispositions du *Digital Services Act* devraient être entrées en vigueur. Le projet de loi prévoit notamment la mise en place de mécanismes de signalement des contenus illégaux et des garanties relatives à la modération des contenus, à commencer par l'information des utilisateurs des décisions de modération les concernant, des raisons les ayant motivées, voire du fait que lesdites décisions ont été prises de manière automatisée (lorsque c'est le cas). Le texte prévoit la possibilité de recours internes et judiciaires contre ces décisions. Le projet de loi confie, par ailleurs, au Conseil Supérieur de l'audiovisuel le soin de contrôler le respect, par les plateformes, de leurs obligations et de prononcer, le cas échéant, des amendes administratives pouvant aller jusqu'à 20 millions d'euros ou 6 % du chiffre d'affaires annuel mondial total de l'exercice précédent. Le texte ajoute, enfin, une possibilité que l'on ne trouve pas spécifiquement dans le *Digital Services Act* : le CSA pourra obtenir que les plateformes « *lui donnent accès aux principes de fonctionnement des outils automatisés auxquels ils ont recours pour répondre à ces obligations, aux paramètres utilisés par ces outils, aux méthodes et aux données utilisées pour l'évaluation et l'amélioration de leur performance ainsi qu'à toute autre information ou donnée lui permettant d'évaluer leur efficacité* » (art. 19 bis). Ce contrôle, par le CSA, de la

configuration des algorithmes constitue une nouveauté bienvenue et indispensable car ce sont, précisément, les algorithmes qui peuvent être amenés à promouvoir les contenus les plus toxiques. Le législateur français prend ainsi acte du fait qu'une simple modération des contenus a posteriori ne peut suffire à lutter contre les contenus répréhensibles et qu'il faut agir, dès l'origine, par le paramétrage des algorithmes.

Si l'on peut s'étonner qu'une loi intervienne au moment même où un texte européen est en cours de négociation au niveau de l'Union Européenne, l'on notera que le législateur français n'est pas le seul. L'Allemagne a récemment réformé la loi NetzDG (*Netzwerkdurchsetzungsgesetz*) du 1^{er} septembre 2017, qui avait fortement inspiré la loi Avia et oblige les plateformes comportant au moins 2 millions d'utilisateurs à supprimer, dans les 24 heures suivant leur signalement par les utilisateurs, les contenus « manifestement illégaux » car réprimés par le Code pénal allemand. Vingt-deux infractions sont visées : discours de haine, diffamation, désinformation, exploitation sexuelle des enfants, incitation à la violence etc. Le délai est porté à 7 jours lorsque l'illégalité n'est pas manifeste. Le texte contraint les plateformes à mettre en place un système de gestion des signalements des utilisateurs et à produire des rapports de transparence semestriels. Les plateformes qui violent leurs obligations « régulièrement et de façon répétée » encourrent une amende pouvant aller jusqu'à 50 millions d'euros. Cette loi a été critiquée pour des motifs assez proches des arguments employés à l'encontre de la loi « Avia » : les plateformes sont ici tenues d'apprécier elles-mêmes l'illicéité d'un contenu et incitée à une forme de « sur-modération ». Pourtant, la réforme en cours consiste surtout à renforcer la loi. La nouvelle version du texte inclut un amendement adopté dans le cadre de la politique de lutte contre l'extrémisme de droite et les crimes de haine (*Gesetzentwurf zur Bekämpfung des Rechtsextremismus und der Hasskriminalität*) qui prévoit que les réseaux sociaux soumis à la NetzDG sont tenus de transmettre à l'Office fédéral de la police criminelle les contenus signalés qu'ils jugent répréhensibles, afin que des poursuites pénales puissent être engagées par les services répressifs compétents. Le texte ajoute, par ailleurs, des exigences supplémentaires dans la loi afin d'améliorer la qualité des rapports de transparence, en obligeant par exemple, les entreprises technologiques à détailler les résultats de leur utilisation de méthodes automatisées pour détecter les contenus illégaux et les modalités de leur politique de modération. Le texte vise également à faciliter les signalements des contenus illicites et à renforcer les

processus d'appel pour permettre aux utilisateurs de contester les décisions de suppression de contenu par un processus d'examen au cas par cas. Les contenus retirés ainsi que les informations permettant d'identifier l'utilisateur et la date et l'heure, doivent être archivés afin de conserver des preuves en vue d'éventuelles poursuites.

La loi allemande a fortement inspiré l'Autriche, qui est, de son côté, en train d'adopter une nouvelle loi sur la responsabilité des plateformes (*Kommunikationsplattformen-Gesetz*), faisant partie d'un ensemble plus vaste visant à lutter contre la haine en ligne. Cette loi vise les plateformes ayant soit plus de 100 000 utilisateurs, soit un revenu annuel supérieur à 500 000 euros, dont il est prévu qu'elles doivent retirer dans les 24 heures les contenus dont l'illégalité est évidente pour un profane en droit. Les autres contenus illégaux doivent être retirés dans les sept jours suivant leur signalement. Si les plateformes ne se conforment pas à leurs obligations, elles pourraient se voir infliger des amendes allant jusqu'à 10 millions d'euros, bien que le régulateur ait une marge de manœuvre pour évaluer l'ampleur de l'amende en tenant compte de facteurs tels que la taille des plateformes et les violations antérieures. L'on ne peut donc que constater ici que les législateurs nationaux se prononcent ici en faveur de textes finalement nettement plus contraignants que le *Digital Services Act*.

c/ La création d'une autorité de régulation

La création d'une autorité de régulation qui aurait pour objet la régulation des plateformes voire, plus largement, du secteur numérique, est discutée de part et d'autre de l'Atlantique. Une telle autorité pourrait avoir pour mission d'assurer une concurrence effective des plateformes grâce aux outils du droit antitrust. Mais s'agissant des réseaux sociaux, elle pourrait avoir un rôle plus approfondi de surveillance et de contrôle de leurs activités, par exemple en contrôlant les algorithmes configurés et utilisés par les plateformes.

Aux Etats-Unis, plusieurs rapports universitaires ont évoqué la création d'une agence fédérale spécifique chargée de réguler les plateformes, même si cette perspective a peu de chance d'aboutir en pratique car elle supposerait une initiative bipartisane qui a peu de chances de voir le jour. Certains experts américains voient cette nouvelle autorité de régulation comme principalement chargée de faire respecter le droit antitrust et la protection des consommateurs. Une étude publiée en août 2020 par la

Harvard Kennedy School¹⁰² a souligné que les agences fédérales américaines existantes ont été conçues pour réguler les marchés de la fin de la révolution industrielle, qui ne sont pas adaptées à des marchés numériques évoluant très rapidement. Elle propose donc la création d'une nouvelle agence fédérale, la *Digital Platform Agency* (DPA) qui veillerait à une concurrence équitable et assurerait à la protection des consommateurs, notamment du point de la protection des données personnelles et de leur portabilité. En revanche, cette nouvelle autorité fédérale n'aurait pas spécifiquement pour mission de s'attaquer aux problèmes liés aux contenus. Sur ce dernier point, d'autres études, envisagent soit l'extension des compétences d'autorités de régulation existantes, tel le *Conseil Supérieur de l'Audiovisuel* (CSA) en France ou la *Federal Trade Commission* (FTC) aux Etats-Unis, soit la création d'une autorité de régulation contrôlant l'ensemble du secteur numérique. Le groupe de travail transatlantique (TWG) constitué par le centre Annenberg de l'Université de Pennsylvanie¹⁰³ a ainsi envisagé la possibilité de créer une nouvelle autorité administrative ou, à défaut, une organisation professionnelle qui constituerait une instance d'auto-régulation, à l'image de la *Financial Industry Regulatory Authority* (FINRA), qui contrôle les sociétés de valeurs mobilières et les courtiers. Quelle qu'en soit la nature, la nouvelle autorité pourrait imposer aux plateformes une plus grande transparence quant à leurs conditions d'utilisation et à leurs pratiques de modération : elle ne pourrait pas, en revanche, véritablement contrôler les contenus compte-tenu des contraintes découlant du Premier Amendement. Elle pourrait, en revanche, contrôler les algorithmes utilisés pour la modération, la recommandation et la hiérarchisation du contenu de manière à s'assurer qu'ils ne promeuvent pas des contenus toxiques. L'autorité de régulation pourrait ainsi vérifier que les algorithmes ne sont pas conçus pour promouvoir des contenus violents, extrêmes ou toxiques. De leur côté, les professeurs Goodman et Kornbluh¹⁰⁴, ont défendu la possibilité de créer une *Digital Democracy Agency* qui serait une autorité indépendante composée

¹⁰² Tom Wheeler, Philip Verveer, and Gene Kimmelman, *New Digital Realities; New Oversight Solutions in the U.S.: The Case for a Digital Platform Agency and a New Approach to Regulatory Oversight*, Shorenstein Center on Media, Politics, and Public Policy, Harvard Kennedy School, August 2020

¹⁰³ Transatlantic High Level Working Group on Content Moderation and Freedom of Expression, "*Freedom and Accountability: A Transatlantic Framework for Moderating Speech Online*" Annenberg Public Policy Center of the University of Pennsylvania, June 2020.

¹⁰⁴ Karen Kornbluh et Ellen P. Goodman, "*Bringing Truth to the Internet*", *Democracy, A journal of ideas*, Summer 2019, n°53

d'experts voire, le cas échéant, un département autonome de la *Federal Trade Commission*. La mission de cette nouvelle autorité serait de protéger les consommateurs, les citoyens et la démocratie par une transparence accrue et une meilleure protection de la vie privée des utilisateurs. Les auteurs envisagent également d'imposer aux plateformes de permettre aux utilisateurs de configurer eux-mêmes les paramètres de leurs flux d'actualités ou les services de référencement qu'ils utilisent et de leur laisser la possibilité de recourir à des applications tierces. Enfin, le professeur Barrett de NYU¹⁰⁵ appelle, lui aussi, à la création d'une autorité administrative indépendante qui serait véritablement adaptée aux spécificités des plateformes numériques.

En Europe, la perspective de la création d'une nouvelle autorité de régulation des plateformes a également été discutée. Certes, au niveau européen, aucune autorité de régulation européenne spécialisée ne paraît, pour l'heure, être sérieusement envisagée pour l'instant. Aux termes des propositions de *Digital Markets Act* et *Digital Services Act*, c'est la Commission qui est chargée de veiller aux respects de leurs obligations par les plus grandes plateformes. Le *Digital Services Act* évoque bien un *European Board for Digital Services* mais il ne s'agira que d'une instance consultative indépendante réunissant les coordinateurs nationaux des services numériques afin d'harmoniser leur action et d'épauler la Commission. En France, un rapport commandé en 2019 par le gouvernement¹⁰⁶ a conclu à la nécessité de disposer d'une autorité administrative indépendante chargée de veiller au respect, par les plateformes, d'obligations de transparence et de diligence mises à la charge des réseaux sociaux en disposant d'un accès privilégié aux algorithmes. Cependant, une telle autorité ne constituerait pas, pour les auteurs du rapport, un régulateur des réseaux sociaux dans leur globalité ou des contenus qui y sont publiés. On ne sait d'ailleurs pas de telles attributions devraient être confiées à une autorité existante ou si une nouvelle autorité devrait être créée. La perspective de la création d'une nouvelle autorité a, en tout cas, été envisagée dans un rapport parlementaire, à ceci près que

¹⁰⁵ Paul M. Barrett, *Regulating Social Media, The Fight Over Section 230- and Beyond*, NYU Stern, septembre 2020

¹⁰⁶ Mission « Régulation des réseaux sociaux- expérimentation Facebook », *Créer un cadre français de responsabilisation des réseaux sociaux : agir en France avec une ambition européenne*, mai 2019.

celui-ci traite avant tout de droit de la concurrence¹⁰⁷. Il est, du reste, désormais acquis que le CSA, qui vient de créer en son sein une « direction des plateformes en ligne », détient désormais la responsabilité de contrôler la manière dont les plateformes sont administrées et modérées, en se faisant, le cas échéant, communiquer différentes informations relatives aux algorithmes utilisés. Il se pourrait d'ailleurs que le CSA soit, dans un avenir proche, transformé en une nouvelle instance à la faveur d'une fusion avec l'Hadopi. En effet, le projet de loi sur l'audiovisuel attendu en 2020 serait devrait être déposé par le gouvernement en avril ou mai 2021 sous le nom de « projet de loi relatif à la protection de l'accès du public aux oeuvres culturelles à l'ère numérique ». Ce texte créerait alors l'Autorité de régulation de la communication audiovisuelle et numérique (Arcom), issue de la fusion du Conseil supérieur de l'audiovisuel (CSA) et de la Haute Autorité pour la diffusion des oeuvres et la protection des droits sur Internet (Hadopi).

¹⁰⁷ V. Valéria Faure-Muntian et D. Daniel Fasquelle, Rapport d'information sur la plateformes numériques, juin 2020.

II. Une nouvelle gouvernance des plateformes

La limitation du pouvoir des plateformes ne passe pas nécessairement par l'intervention de l'Etat et la régulation. Elle peut également reposer sur un renouvellement des modes de gouvernance des plateformes que les auteurs se revendiquant du « constitutionnalisme numérique » étudient et approfondissent¹⁰⁸. Une telle piste s'impose tout particulièrement dans le contexte américain, puisque le Premier Amendement limite très fortement la possibilité d'adopter une réglementation d'ensemble relative aux contenus ou de confier une telle régulation à une autorité gouvernementale. Deux pistes paraissent se dessiner aujourd'hui. La première consiste à impliquer un tiers indépendant dans la gouvernance de la plateforme : tel est l'objet de la création, par Facebook, de l'*Oversight Board* (A). La seconde revient à associer les utilisateurs à la gestion de la plateforme de manière à en démocratiser la gouvernance (B).

a/ Le recours à une instance tierce : l'expérience de l'*Oversight Board*

Dans son rapport déjà cité, le groupe de travail du centre Annenberg¹⁰⁹ évoque la possibilité de créer des « conseils » des réseaux sociaux, dont les membres représenteraient les professionnels, le gouvernement, les organisations de la société civile et le monde universitaire. Indépendant de toute entreprise, un conseil des médias sociaux pourrait formuler des recommandations de politique générale, fixer des standards de modération des contenus, voire même entendre certains appels concernant les décisions de modération des plateformes du secteur. C'est manifestement dans cette perspective que s'inscrit la création, par Facebook, de l'*Oversight Board* (Conseil de Surveillance), instance suggérée en 2018 par Noah Feldman, professeur de droit à Harvard, qui lui avait présenté un projet de « cour suprême ». La composition de cette instance a été annoncée en mai 2020. En octobre 2020, Facebook a autorisé les appels d'un pourcentage aléatoire de cinq pour cent

¹⁰⁸ N. Suzor, "Digital Constitutionalism: Using the Rule of Law to Evaluate the Legitimacy of Governance by Platforms", *Social Media + Society* July-September 2018: 1–11

¹⁰⁹ Transatlantic High Level Working Group on Content Moderation and Freedom of Expression, "*Freedom and Accountability: A Transatlantic Framework for Moderating Speech Online*" Annenberg Public Policy Center of the University of Pennsylvania, June 2020.

des utilisateurs de Facebook et Instagram, ce qui a permis au conseil de commencer ses travaux. Conçu pour trancher les contestations, l'*Oversight Board* de Facebook fait penser à une instance arbitrale, à ceci près que seule l'une des parties, Facebook, en a fixé la composition, les conditions de saisine et la procédure qui y est suivie. Le Conseil a un rôle juridictionnel dans la mesure où il examine les recours et rend des décisions obligatoires pour la plateforme mais il a également un (petit) rôle politique, car il peut formuler des recommandations générales à titre consultatif.

- **Composition et fonctionnement de l'Oversight Board**

Le Conseil de Surveillance est composé de vingt membres recrutés parmi de personnalités particulièrement reconnues comme Helle Thorning-Schmidt ancienne première ministre du Danemark, Catalina Botero Marino, ancienne rapporteuse spéciale pour la liberté d'expression auprès de la Commission interaméricaine des droits de l'homme, Alan Rusbridger, ancien rédacteur en chef du Guardian, et Tawakkol Karman, militante qui a reçu le prix Nobel de la paix en 2011 pour son rôle dans les manifestations du printemps arabe au Yémen. Cinq membres sur vingt sont américains, trois sont européens, les autres membres venant des quatre coins du monde.

L'*Oversight Board* a avant tout pour mission d'examiner les recours formés par les utilisateurs de Facebook ou Instagram contre les décisions de modération¹¹⁰. A l'origine, seules les décisions de suppression de contenu prises par la plateforme pouvaient faire l'objet d'une contestation. Depuis le 13 avril 2021, les décisions de maintenir en ligne certains contenus malgré les signalements des utilisateurs peuvent être soumises au Conseil. Dans ces cas, l'utilisateur concerné a 15 jours, après avoir contesté sans succès la décision auprès de la plateforme, pour former un recours devant le Conseil. En revanche, un utilisateur dont le compte n'est plus actif ne peut émettre une contestation : les décisions de suspension, telle celle relative à Donald Trump, ne peuvent être contestées par la personne concernée. En ce cas, seule la plateforme, qui dispose d'un droit général de saisir le conseil en présence de cas

¹¹⁰Evelyn Douek, "The Facebook Oversight Board Should Review Trump's Suspension", Lawfare Blog, January 11, 2021.

importants et difficiles, peut décider de lui soumettre ce type d'affaire, comme elle l'a fait dans le cas de Donald Trump.

Le Board peut accepter ou refuser d'examiner une affaire. Un comité de trois membres du Conseil est chargé de la sélection des cas ayant un impact et une importance particuliers. Une fois acceptée, chaque affaire est confiée à un comité de cinq membres, dont au moins un membre de la région où se trouve le contenu examiné. Le Conseil doit en principe rendre sa décision dans les 90 jours suivant l'acceptation de l'affaire. Une période de commentaires publics est prévue avant que le conseil ne commence à délibérer sur un cas, afin de permettre à des tiers de partager leurs idées et leurs points de vue. Il n'y a pas d'audience : l'utilisateur ayant saisi le Conseil lui soumet un mémoire écrit et un représentant de Facebook dépose un mémoire expliquant le raisonnement de la plateforme. La décision du comité de cinq membres doit être prise à la majorité, et peut inclure des opinions concordantes et dissidentes. Le projet de décision est alors envoyé à l'ensemble du Conseil, qui peut se prononcer conformément à la position du comité ou décider la convocation d'un nouveau comité pour statuer à nouveau. La question se pose des normes appliquées pour trancher les contestations. Facebook a décidé, depuis l'origine, d'exclure les lois nationales des normes appliquées par le Conseil qui ne se prononcera pas sur la licéité du contenu contesté. Il revient, en revanche, au Conseil de se prononcer en fonction des politiques et des valeurs de Facebook tout en accordant une attention particulière à l'impact de la suppression du contenu du point de vue des normes des droits de l'homme protégeant la liberté d'expression. L'indifférence aux droits nationaux n'est pas totale : la saisine du Board sera irrecevable en présence de contenus illégaux selon le droit national d'un pays en lien avec le contenu considéré, soit le droit applicable à l'utilisateur ayant publié le contenu contesté ou à l'utilisateur l'ayant signalé. Ainsi, un utilisateur français ayant posté un contenu injurieux aux termes de l'article 33 de la loi de 1881 ne pourra saisir l'Oversight Board en cas de suppression. La plateforme se livrera donc, une fois de plus, à une appréciation des droits nationaux pour déterminer la recevabilité des contestations.

Une fois la décision de l'*Oversight Board* adoptée, elle est publiée en ligne avec les principaux éléments du dossier et la motivation. Facebook s'est engagé à mettre en œuvre les décisions dans un délai de 7 jours, sauf impossibilité de s'y conformer, par exemple s'il existe une obligation légale de retirer le contenu considéré. Le réseau social s'est, dans tous les cas, engagé à rendre publiques toutes les mesures prises en réponse aux décisions du Conseil. Il reste que ces décisions ne s'appliquent qu'au cas individuel considéré. Facebook s'est engagé à supprimer « les publications identiques avec un contexte parallèle » en fonction de sa « capacité technique et opérationnelle » mais l'on ignore précisément comment un tel engagement pourra se concrétiser dans un contexte où la modération est, au quotidien, assurée soit par des algorithmes, soit par du personnel souvent peu qualifié et chargé de prendre un grand nombre de décisions rapides sans pouvoir nécessairement bénéficier d'un temps d'analyse. Les décisions du Conseil ne paraissent donc guère lier Facebook au-delà du cas considéré¹¹¹.

- **Les premières décisions de l'*Oversight Board***

Les premières affaires examinées ont été retenues car considérées comme présentant une "importance critique". Dans ces décisions, l'*Oversight Board* a annulé cinq des six décisions de Facebook qui lui étaient soumises et paraît privilégier très nettement la liberté d'expression. L'une concernait une publication qualifiant Emmanuel Macron de diable et appelant au boycott des produits français. La position de Facebook était que la publication ne respectait pas ses standards en matière d'incitation à la violence, mais le Conseil a jugé que Facebook n'avait pas examiné avec précision toutes les informations contextuelles et que le risque de violence n'était pas si grand : ainsi, les normes internationales des droits de l'homme en matière d'expression justifiaient de rétablir la publication. Un autre cas concernant la France était relatif à une vidéo et un texte d'accompagnement publiés dans un groupe de 500 000 utilisateurs qui critiquaient le refus du gouvernement français d'autoriser un traitement à base d'azithromycine et d'hydroxychloroquine, pour lutter contre la covid-19 dont il

¹¹¹ Evelyn Douek, "What Kind of Oversight Board Have You Given Us?", The University of Chicago Law Review Online, 2020

était affirmé qu'il pouvait sauver des vies. Facebook a supprimé la publication au motif qu'elle incitait les utilisateurs à s'auto-médiquer. Mais l'*Oversight Board* a toutefois analysé la publication comme un discours politique qui n'avait pas à être censuré dès lors qu'il ne générait pas un risque, les médicaments concernés n'étant pas en vente libre en France. La seule décision dans laquelle le Conseil a approuvé la suppression est relative à une publication dans laquelle un utilisateur désignait les Azerbaïdjanais en des termes injurieux, même si certains membres estimaient que la suppression par Facebook n'était pas forcément proportionnée en présence de propos émanant de minorités¹¹². Enfin, une décision a mis en évidence les difficultés soulevées par le recours aux algorithmes. Des photos montrant des mamelons féminins dans le cadre d'une campagne de sensibilisation au cancer du sein sur Instagram avaient été supprimées par un algorithme pour violation des règles de la plateforme en matière de nudité, puis rétablies au moment où le Conseil avait choisi de s'en saisir. Le *Board* a souligné que ce cas illustre le manque de surveillance humaine appropriée et demandé qu'à l'avenir les utilisateurs soient informés du fait que la suppression des contenus était le fait d'un algorithme et se voit donner la possibilité de faire appel. Cette décision a, en tout cas, montré les limites de l'intervention de l'*Oversight Board*, qui n'a pas compétence pour statuer sur la configuration des algorithmes, même si Facebook a affirmé après coup avoir modifié ses directives et amélioré ses systèmes de signalement basés sur l'apprentissage automatique.

- **Les limites de l'*Oversight Board***

Il faut très certainement saluer l'ambition de Facebook qui a, en créant cette instance, voulu apporter des éléments de solution à la question aigue de la politique des discours en tentant de l'aborder de manière globale, alors même que les opinions varient fortement d'un lieu à l'autre, d'une culture à l'autre, à propos de ce qui peut se dire ou non sur les réseaux sociaux¹¹³. Pour autant, l'apport que représente l'*Oversight Board* comporte des limites et des incertitudes. Le pouvoir réel du Conseil reste limité. Il peut formuler des recommandations de politique générale, mais seulement à titre

¹¹² K. Klonick, "Inside the making of Facebook's Supreme Court", *The New Yorker*, February 12, 2021

¹¹³ K. Klonick, préc.

consultatif. La portée de ses décisions au-delà du cas individuel considérée reste relative, même si l'on peut espérer que la motivation des décisions et le caractère public de la procédure contribueront à améliorer la transparence dans la gestion des contenus sur le réseau et permettront de faire émerger des « précédents », donc une forme de « jurisprudence » pouvant disposer d'une autorité allant au-delà de Facebook : de fait, le *Board* lui-même a récemment annoncé souhaiter voir d'autres réseaux sociaux lui soumettre leurs cas¹¹⁴. Il reste que la manière dont les décisions du Conseil peuvent et pourront s'articuler avec les droits nationaux n'apparaît pas clairement et que l'on ignore ce que pourrait être la décision d'un juge national saisi d'un recours contre une décision du *Board*. Enfin, et surtout, le Conseil n'a pas accès à des informations précises concernant les algorithmes utilisés par Facebook pour gérer les contenus et déterminer les informations visibles sur le fil d'actualité des utilisateurs. Il n'intervient qu'en bout de course, à propos d'un cas individuel, pour se prononcer sur le point de savoir si une publication doit ou non être retirée, sans pouvoir prendre position sur la manière dont les contenus parviennent aux internautes. Ce point a été particulièrement souligné par différents experts à propos du compte de Donald Trump¹¹⁵ : les décisions de modération de contenu soumises à l'*Oversight Board* ne peuvent pas être séparées des questions de conception de la plateforme qui déterminent les discours qui se répandent sur le réseau, la vitesse de leur propagation, les utilisateurs qui les voient et le contexte dans lesquels ils les voient. Ne pouvant statuer ni sur les décisions définissant les types de contenus pouvant être partagés et leur classement, ni sur celles qui décident des publicités politiques vendues par la plateforme et de leur présentation, le *Board* doit se prononcer sur des décisions individuelles de modération en prenant l'architecture de Facebook comme un donné. Dans le cas de Donald Trump, soulignent ces experts, la dangerosité des propos du Président provenait du fait que de nombreuses informations complotistes avaient été diffusées durant des mois sur Facebook et poussées par les algorithmes de la plateforme. Pour cette raison, disent ces auteurs, l'*Oversight Board* serait bien inspiré de ne pas répondre directement à la question du bien-fondé de la suspension de Donald Trump et de commencer par commander une enquête indépendante sur la

¹¹⁴ T. Hatmaker, "Facebook Oversight Board says other social networks "welcome to join" if project succeeds, Tech Crunch, February 11, 2021

¹¹⁵ J. Jaffer et K. Glenn Bass, "Facebook's « Supreme Court » faces its first major test", New York Times, February 17, 2021.

manière dont la configuration de la plateforme a pu contribuer aux événements du Capitole¹¹⁶. Pour l'heure, un membre du Conseil, Alan Rusbridger a fait part du souhait des membres du Conseil d'accéder aux algorithmes utilisés par Facebook¹¹⁷, de manière à bien comprendre la modération algorithmique, voire à formuler des recommandations sur la configuration des algorithmes. Cependant, même si Facebook a confirmé souhaiter étendre les prérogatives de l'Oversight Board, l'on peut douter que celui-ci parviendra à accéder à des informations relatives à la stratégie et aux modalités de gestion des contenus en ligne, qui constituent le cœur du modèle d'affaire de la plateforme. Facebook a d'ailleurs déclaré¹¹⁸ qu'il n'était pas prévu de soumettre au Conseil les algorithmes de recommandation, d'engagement, ou de classement de contenus. Pour le reste, à supposer même que les membres de l'Oversight Board accèdent aux algorithmes, il n'est pas acquis qu'ils seraient réellement en mesure d'influer sur la manière dont les algorithmes sont développés et configurés. En l'état, on peut redouter que la création de ce Conseil permette surtout de conférer un pouvoir symbolique à une instance tierce qui statue dans un nombre restreint de cas difficiles tout en permettant à Facebook de conserver le pouvoir, nettement plus important, de décider des politiques globales de modération des contenus et de la configuration des algorithmes.

Si l'on peut regretter les pouvoirs limités de l'Oversight Board, il reste qu'un accroissement de son influence conduirait directement à un écueil redoutable : sa légitimité. Le Conseil dispose pour l'heure d'un pouvoir juridictionnel. Il est fréquemment souligné qu'il devrait devenir une véritable instance politique, disposant d'un droit de regard sur les décisions stratégiques en matière de configuration des algorithmes et de pratiques publicitaires. Cependant, cette instance est composée de membres exclusivement nommés par Facebook, selon des critères qui lui appartiennent, et rémunérés, fût-ce indirectement¹¹⁹, par la plateforme. A ce problème d'indépendance s'ajoute la question de la représentativité des membres du Conseil, qui doivent se prononcer sur des affaires provenant du monde entier concernant une

¹¹⁶ J. Jaffer et K. Glenn Bass, préc.

¹¹⁷ A. Hern, "Alan Rusbridger says Oversight Board will ask to see Facebook's algorithm", The Guardian, March 2 2021

¹¹⁸ R. Heilweil "If Mark Zuckerberg won't fix Facebook's algorithms problem, who will?",

¹¹⁹ L'*Oversight Board* est un *trust* juridiquement distinct de Facebook, dans lequel la plateforme a placé 130 millions de dollars afin de financer ses activités.

plateforme comportant près de 2,5 milliards d'utilisateurs où se parlent plus de 100 langues différentes. On a souligné que le Conseil comprend un nombre excessif d'américains (cinq) mais un seul indien alors que l'Inde est le pays comprenant le plus grand nombre d'utilisateurs de Facebook¹²⁰. Du reste, l'Inde compte plus de 22 langues principales et 700 dialectes, une majorité hindoue mais des millions de musulmans ainsi que bouddhistes, de chrétiens, de juifs et de bahaïs. Peut-on, dans ces conditions, espérer concevoir des principes globaux relatifs aux discours en ligne ou s'agira-t-il simplement ici simplement d'appliquer les conceptions occidentales relatives à la liberté d'expression ? Beaucoup redoutent que le Conseil ait été façonné pour faire prévaloir la liberté d'expression conformément à la stratégie originelle de Facebook et au détriment d'autres valeurs, comme la sécurité et la dignité. Les critiques très vives formulées à l'encontre de l'*Oversight Board* ont d'ailleurs conduit à la création, par un groupe d'experts, d'une instance informelle et concurrente, totalement indépendante de Facebook, le "*Real Oversight Board* », où l'on trouve, par exemple, l'ancien président estonien, Toomas Henrik Ilves, la professeure Shoshana Zuboff, et Derrick Johnson, le président du NAACP. Dans un tel contexte, c'est la décision très attendue sur la suspension du compte de Donald Trump qui devrait façonner (ou pas) la crédibilité de l'*Oversight Board* officiel.

b/ L'implication des utilisateurs

Il est tentant d'invoquer le caractère « quasi-étatique » des plus grandes plateformes pour inciter à une gouvernance plus démocratique des plateformes donnant davantage de place à ces « citoyens » que sont les utilisateurs. Certains ont soutenu qu'à partir de 100 millions d'utilisateurs, une plateforme devient « indiscernable de l'État » et ne doit pas rester exclusivement gouvernée par des particuliers, car « les décisions sur une plateforme de cette envergure sont intrinsèquement politiques »¹²¹. La plateforme devrait donc être gouvernée démocratiquement, par exemple par un « conseil d'administration composé d'actionnaires, d'élus, de développeurs ou travailleurs de la plateforme et d'utilisateurs »¹²². Si certains vont jusqu'à envisager froidement la perspective d'une nationalisation¹²³ des grandes plateformes, un tel projet paraît peu

¹²⁰ S. Vaidhyathan, "Facebook and the Folly of Self-Regulation", *Wired*, 5 septembre 2020.

¹²¹ Jason Barrett Prado, "Taxonomizing platforms to scale regulation", *Venture Commune* (substack), 18 novembre 2010.

¹²² Ibid.

¹²³ Ben Tarnoff, « Platforms don't exist » (substack), 22 novembre 2019

réaliste à propos d'entreprises transnationales dont la plupart ont leur siège social aux Etats-Unis. L'on peut, pour la même raison, douter qu'une loi nationale vienne un jour imposer aux grandes plateformes une modification de leur gouvernance et notamment de la composition de leur conseil d'administration. Une telle loi devrait, en effet, être avant tout américaine, voire même californienne compte-tenu du lieu d'établissement des principales plateformes, et l'on peut juger improbable la perspective d'une évolution législative sur ce point, que ce soit du côté du législateur fédérale ou californien. Dans un tel contexte, l'on en est, bien plutôt, réduit à attendre des plateformes elles-mêmes qu'elles prennent l'initiative de transformer leur gouvernance pour y laisser davantage de place aux utilisateurs. Elles y sont ouvertes.

Mark Zuckerberg s'est ainsi déclaré favorable à une organisation plus démocratique ou à un processus de décision laissant davantage de place aux différentes communautés présentes sur la plateforme¹²⁴, de manière à ne pas laisser la place aux seuls intérêts à court-terme des actionnaires. De fait, Facebook a déjà fait des tentatives en ce sens. En 2009, peu après que la plateforme a été critiquée pour avoir discrètement modifié ses conditions d'utilisation afin de pouvoir conserver les données des utilisateurs même après la suppression de leur compte, Facebook a annoncé une "nouvelle approche de la gouvernance du site"¹²⁵. L'idée était de laisser les utilisateurs voter par référendum sur les politiques de Facebook, notamment en matière de gestion des données. Facebook s'engageait à respecter le résultat du référendum dans le cas où 30 % de ses utilisateurs à l'échelle mondiale voteraient. Le premier référendum a porté sur la modification des conditions générales de la plateforme : seuls 0,32 % des utilisateurs ont voté. Le dernier référendum a eu lieu en 2012, à propos de la définition des conditions dans lesquelles Facebook pouvait partager les données des utilisateurs avec d'autres organisations : seuls 668 000 utilisateurs ont voté. Depuis lors, Facebook a renoncé à organiser des référendums. Il est vrai que ces référendums de masse sont difficiles à organiser et que la participation s'est en l'occurrence révélée très faible. D'autres perspectives seraient peut-être envisageables, comme la

¹²⁴ « un processus plus démocratique ou axé sur la communauté, qui reflète les valeurs des gens dans le monde entier », in Ezra Klein, « Mark Zuckerberg on Facebook's hardest year, and what comes next », Vox, 2 avril 2018

¹²⁵ K. Klonick, préc.

désignation, à l'échelle nationale, de représentants des utilisateurs qui pourraient être consultés sur la politique de la plateforme.

De son côté, Twitter promeut le principe d'une décentralisation qui rendrait inutile toute intervention étatique supplémentaire et toute régulation contraignante. Dans un « fil » publié sur Twitter le 14 janvier 2021, quelques jours après la suspension de Donald Trump, Jack Dorsey a promu l'idée d'un réseau social décentralisé en se référant au modèle de décentralisation de la plateforme Bitcoin. Twitter a ainsi récemment adopté plusieurs initiatives destinées à associer davantage les utilisateurs à la gestion de la plateforme. Depuis janvier 2021 les utilisateurs situés aux Etats-Unis peuvent participer à l'expérience Birdwatch¹²⁶, dans laquelle les utilisateurs contribuent au *fact checking*. Il ne s'agit toutefois pas d'un programme de vérification des faits, car il ne s'agit pas ici de faire appel à *fact checkers* qualifiés. Les utilisateurs peuvent simplement identifier les tweets qu'ils jugent trompeurs et présenter les raisons motivant leur position en joignant des liens vers des sources extérieures fiables, une présentation du contexte et une appréciation du dommage que la désinformation est susceptible de causer. Les contributions sont classées par un algorithme en « utiles » et « inutiles », dans l'objectif de mettre en avant celles qui suscitent un "consensus diversifié", c'est-à-dire un accord entre des personnes qui n'ont pas tendance à être d'accord. L'idée est intéressante : elle permet une modération plus transparente (l'algorithme est divulgué), plus efficace (car les utilisateurs réagissent en temps réel), de nature à renforcer la confiance des utilisateurs en répartissant sur l'ensemble de la communauté Twitter la responsabilité de la lutte contre la désinformation. La démarche a toutefois ses limites : il n'est, en effet, pas exclu que l'algorithme donne des résultats faussés, puisqu'il dépend des contributions qui lui parviennent, qui pourraient tout à fait provenir de groupes extrémistes ou complotistes¹²⁷.

Au-delà de Birdwtach, Twitter souhaite développer des fonctionnalités permettant aux utilisateurs de choisir ce qu'ils voient en ligne en sélectionnant eux-mêmes les algorithmes d'affichage et de recommandation qu'ils souhaitent utiliser. Twitter pourrait à la fois créer ses propres algorithmes de classement parmi lesquels les gens

¹²⁶ K. Tiffany, "Who would volunteer to fact-check Twitter?", The Atlantic, March 3, 2021

¹²⁷ *Ibid.*

pourraient choisir et proposer une "place de marché" où les gens sélectionneraient différentes options. Ce projet de réseau social décentralisé, nommé Bluesky, a été lancé en décembre 2019 et est en cours de développement¹²⁸. Les utilisateurs pourraient ainsi façonner leur propre fil d'actualité et les recommandations qu'ils reçoivent en fonction de leurs préférences et partager leurs préférences avec leurs contacts¹²⁹. Plus encore, ils pourraient évaluer le fonctionnement des algorithmes et mesurer l'impact des différents filtres de contenu qui leur sont proposés. L'on pourrait même imaginer la conduite de véritables études d'impact des algorithmes sur le comportement des utilisateurs. Au-delà d'une amélioration de l'expérience des utilisateurs, une telle décentralisation aurait des avantages politiques et juridiques. Elle permettrait d'éviter que le réseau social et les discours qui y sont tenus soient sous le contrôle d'une seule entité. L'intérêt de la décentralisation serait également juridique : même si la section 230 était réformée et les obligations des plateformes alourdies, leur responsabilité serait moins lourde en présence d'une forte décentralisation. Reste à vérifier, dans les faits, qu'un réseau social décentralisé, dans lequel les utilisateurs disposent du choix des algorithmes qu'ils utilisent, ne servirait pas les pires abus tels les crimes sexuels, les discours de haine, les appels à la violence ou le complotisme. Car une telle liberté de choix donnée aux utilisateurs pourrait produire un résultat opposé à celui recherché, un scénario où les internautes se constitueraient des bulles de filtre dans lesquelles les idées les plus fausses ou intolérables se répandraient sans aucune limite. Ce scénario rappelle le propos du conservateur Brendan Carr¹³⁰, nommé par Trump à la FCC, selon qui l'on devrait, au nom du « *user empowerment* » permettre aux utilisateurs de désactiver les fonctions de modération des contenus et de vérification des faits, ce qui permettrait d'obtenir une version « Wild West » des réseaux sociaux et de disposer, selon l'intéressé, d'un contenu exempt de tout biais anti-conservateur. Une perspective plutôt inquiétante!

¹²⁸ L. Matney, « Twitter's decentralized future », TechCrunch, 15 janvier 2021

¹²⁹ J. Kastrenakes, "Twitter's Jack Dorsey wants to build an app store for social media algorithms", The Verge, February 9, 2021

¹³⁰ B. Carr, "A conservative path forward on Big Techs", Newsweek, July 27, 2020

Conclusion

Au final, beaucoup des questions relatives à l'appréhension du pouvoir écrasant des réseaux sociaux et de la police des discours en ligne restent, pour l'heure, en suspens. La première d'entre elle concerne le point de savoir s'il serait possible, à terme, de traiter ces questions de manière globale ou, à tout le moins, commune à l'ensemble des démocraties occidentales qui se reconnaissent autour de valeurs essentielles comme le respect de la liberté d'expression ou le refus de la violence en ligne. Les Etats, ou une majorité d'Etats, pourraient-ils, par exemple s'entendre sur une ligne commune à l'égard des plateformes ? Pourrait-on envisager la création d'une structure internationale destiné à élaborer des standards globaux et à contrôler l'action des plateformes ? Ou sommes-nous en train de voir se dessiner progressivement un Internet régionalisé, au sein duquel de nouveaux blocs pourraient apparaître, à côté de l'Union Européenne et des Etats-Unis ? Une autre question tient aux rôles respectifs des Etats, des plateformes et des utilisateurs dans la détermination de ce qui peut être dit ou non en ligne. Faut-il s'en remettre aux législations nationales et à leur définition de l'illicéité ? Faut-il laisser les plateformes déterminer les types de contenus qu'elles souhaitent exclure du réseau en se détachant des législations en vigueur, que ce soit pour tolérer des contenus illicites en vertu d'une loi nationale ou écarter des contenus licites que la plateforme juge indésirables ? Faut-il, enfin, confier aux utilisateurs la responsabilité de signaler ou vérifier les contenus publiés, au risque de ne pas parvenir à lutter efficacement contre les publications toxiques et virales ? L'apparition, en une décennie, des grandes plateformes a fait émerger un univers inconnu et inédit, dont les meilleurs spécialistes ne soupçonnaient pas qu'il allait changer aussi profondément l'accès et la compréhension de l'information¹³¹. Il reste, désormais, à construire un espace global dans lequel les personnes peuvent communiquer et échanger à partir d'un socle minimal de standards et de valeurs communes.

¹³¹L. Lessig, "La segmentation du monde que provoque Internet est dévastatrice pour la démocratie", France Culture, 22 décembre 2016.

Au sujet de l'auteur :

Florence G'sell est agrégée de droit privé et sciences criminelles et professeur de droit privé à l'Université de Lorraine. Enseignante à Sciences Po. Elle est titulaire de la Chaire Digital, Gouvernance et Souveraineté (Ecole d'Affaires Publiques). Ses travaux portent principalement sur les enjeux juridiques de la transition numérique, qu'elle appréhende notamment d'un point de vue comparatif, en envisageant tant le droit français que le droit des Etats-Unis. Elle a récemment dirigé *Le Big Data et le Droit* (Dalloz, 2020) et écrit *Justice Numérique* (Dalloz, 2021).

Au sujet de la Chaire Digital, Gouvernance et Souveraineté :

La mission de la **Chaire Digital, Gouvernance et Souveraineté** de Sciences Po est de créer un écosystème unique pour rapprocher l'univers des entreprises technologiques du monde de la recherche académique, du monde politique, de la société civile, et des incubateurs de politiques publiques et de régulation du numérique. Ces relations nécessitent un écosystème de recherche, d'innovation et de formation qui soit pluridisciplinaire, international et en prise directe avec la sphère publique.

Portée par l'**École d'Affaires Publiques**, elle est résolument pluridisciplinaire pour penser de façon holistique les transformations économiques, juridiques, sociales ou encore institutionnelles entraînées par le numérique.

La Chaire Digital, Gouvernance et Souveraineté est dirigée par **Florence G'sell**, professeure de droit à l'Université de Lorraine, enseignante à l'École d'Affaires Publiques de Sciences Po.

Les activités de la Chaire sont soutenues par ses partenaires:

