

CE QUE PEUT L'ÉTAT FACE AUX PLATEFORMES

Depuis l'émergence des grandes plateformes du Web dans les années 2000, les États ont majoritairement été perçus comme faisant preuve d'impuissance politique à leur égard. De par leur dimension internationale (ces entreprises exercent leurs activités sur l'ensemble des continents), leur popularité (des services tels que ceux proposés par YouTube comptent plusieurs milliards d'utilisateurs) ou leur richesse (les GAFAM – Google, Apple, Facebook, Amazon, Microsoft – font partie des entreprises les plus profitables au monde), ces grandes firmes seraient « plus fortes » que les États et échapperaient ainsi à toute velléité de régulation.

49

L'année 2016 semble à ce titre avoir marqué un tournant, en ouvrant une nouvelle séquence de l'histoire des relations entre États et plateformes. Aux États-Unis, les auditions par le Congrès des responsables des principaux réseaux sociaux dans le cadre de « l'affaire russe » et de l'affaire Cambridge Analytica¹, puis les actions antitrust engagées par le département de la Justice et la Federal Trade Commission contre Google et Facebook, ont instauré un climat politique propice à la régulation. En Europe, la pression des pouvoirs publics s'est traduite par le vote de plusieurs lois, notamment en France et en Allemagne, incitant les plateformes à contrôler davantage les publications qu'elles hébergent. À l'échelle de l'Union européenne, la Commission a présenté fin 2020 une nouvelle législation en matière de régulation de l'économie numérique

1. La première affaire concerne les opérations de l'Internet Research Agency, agence de communication réputée proche du Kremlin, visant à financer des contenus publicitaires pro-Trump et anti-Clinton sur Facebook lors de la campagne de l'élection présidentielle américaine de 2016. La seconde est liée à la récupération de données personnelles de plusieurs millions d'internautes sur Facebook, à des fins de fichage politique en faveur du camp républicain, dans le cadre de cette même présidentielle.

(*Digital Market Act et Digital Services Act*), qui clôture six années pendant lesquelles la commissaire à la concurrence, Margrethe Vestager, s'est notamment illustrée en infligeant des amendes record aux GAFAM pour pratiques anticoncurrentielles.

50 Ce volontarisme politique s'exprime dans différents domaines : protection de la vie privée et des données personnelles, défense du droit d'auteur, mesures contre la propagande djihadiste, démantèlement des monopoles, ou encore lutte contre les discours de haine, les fausses informations ou le cyber-harcèlement². Dans cet article, je traiterai d'un sujet particulier, celui de la régulation des contenus sur les réseaux sociaux. Face à la montée de la désinformation et des discours de haine en ligne, plusieurs États européens ont entrepris de réglementer les usages de la parole publique sur internet, en imposant de nouvelles contraintes aux géants du numérique. Si ces mesures posent question quant aux menaces qu'elles font peser sur l'exercice de la liberté d'expression en ligne, elles sont également le signe d'un retour de la puissance publique dans le champ de la gouvernance d'internet.

DE « CODE IS LAW » À « LAW IS CODE »

Parmi les recherches en sciences sociales qui s'intéressent à la régulation et à la gouvernance d'internet, les travaux du juriste américain Lawrence Lessig font figure de référence³. Dès les années 1990, ce professeur de droit à Harvard développe une théorie politique du pouvoir normatif du code informatique. Dans les sociétés contemporaines, nous dit Lessig, quatre sources de normes règlent les comportements des individus : les lois, la culture, les marchés et les infrastructures. Si, dans notre vie « physique », les lois disposent du pouvoir de prescription le plus fort sur nos comportements, sur internet, ce sont les infrastructures techniques qui dominent. Ce que les internautes y font est moins le résultat de ce que la loi interdit ou non de faire que le produit de ce que la technologie rend possible. Les controverses autour du téléchargement illégal, au cours des années 2000, en sont une bonne illustration : si la loi interdit cette pratique, le simple fait que des logiciels la rendent

2. Sur ces différents sujets, cf. Romain Badouard, *Les Nouvelles Lois du Web. Modération et censure*, Paris, Seuil, 2020 ; Julie Groffe-Charrier, « La loi est-elle dictée par le code ? », *Dalloz IP/IT*, n° 11, 2020, p. 602-606 ; Joëlle Toledano, *GAFAM : reprenons le pouvoir !*, Paris, Odile Jacob, 2020 ; Félix Tréguer, *L'Utopie déçue. Une contre-histoire d'internet*, Paris, Fayard, 2019.

3. Cf. notamment Lawrence Lessig, *Code: And Other Laws of Cyberspace*, New York (N. Y.), Basic Books, 1999.

facilement accessible pousse des milliers d'individus à contourner les contraintes légales, alors même qu'ils ne les auraient jamais enfreintes dans un magasin, par exemple. « *Code is law* », avait résumé Lessig : sur internet, le code informatique fait office de loi. Par extension abusive, la formule a pu être interprétée comme l'expression d'une forme de déterminisme technique, selon lequel le code informatique serait « plus fort » que la loi, et par extension les plateformes – qui possèdent désormais les principales infrastructures d'échange du Web – seraient « plus fortes » que les pouvoirs publics qui produisent la législation.

Dans la lignée de ces travaux précurseurs, le concept de « gouvernamentalité numérique » a émergé pour exprimer la manière dont les technologies que nous manipulons exercent un pouvoir de régulation sur nos comportements sociaux⁴. S'inspirant des travaux de Michel Foucault sur les rationalités de gouvernement, qui décrivent la manière dont les actions des individus peuvent être orientées et canalisées par un pouvoir politique sans avoir recours à la force⁵, la « gouvernamentalité numérique » désigne la capacité des technologies à rendre conformes nos comportements aux attentes de leurs concepteurs. Les dirigeants et développeurs des grandes plateformes disposent ainsi d'un pouvoir de « faire faire » à grande échelle, puisqu'ils orientent nos habitudes d'information, contraignent nos modes d'expression et encadrent nos échanges, produisant de fait de nouvelles normes sociales.

51

Dans le domaine qui nous intéresse ici, celui de l'expression publique, la gouvernamentalité des plateformes s'exprime de trois manières différentes. D'abord, à travers les dispositifs de prise de parole qu'elles mettent à disposition de leurs usagers, ces plateformes produisent des formats d'expression qui canalisent des manières de dire. La limite des deux cent quatre-vingts caractères sur Twitter ou les « *reaction buttons* » de Facebook (ces émoticônes qui permettent en un clic d'exprimer un sentiment au sujet d'une publication) en sont de bons exemples : les réseaux sociaux nous permettent de nous exprimer, mais toujours d'*une certaine façon*. Ensuite, les algorithmes de tri ou de recommandation des plateformes organisent les contenus et distribuent la visibilité des informations. Ils décident de ce qui est vu, consulté et débattu à un moment donné, en donnant davantage de visibilité à certains types de

4. Romain Badouard, Clément Mabi et Guillaume Sire, « Beyond "Points of Control": Logics of Digital Governmentality », PolicyReview.info, 30 septembre 2016 ; *Réseaux*, n° 225, *L'action publique au prisme de la gouvernamentalité numérique*, Paris, La Découverte, 2021.

5. *Sécurité, territoire, population. Cours au Collège de France, 1977-1978*, Paris, Gallimard-Seuil, 2004.

contenus, et en réduisant au silence un certain nombre d'autres publications. Enfin, les grandes plateformes mettent en œuvre des dispositifs de modération qui articulent participation des internautes *via* la pratique du signalement, évaluation des contenus par des modérateurs professionnels et détection automatique des contenus problématiques par des outils d'intelligence artificielle.

52 Le pouvoir politique des plateformes sur l'expression publique, par le biais des architectures et des infrastructures d'échange qu'elles bâtissent, est donc bien réel. Celles-ci donnent naissance à de nouvelles formes de débat public, dont elles fixent seules les règles. Depuis 2016 en Europe, ce pouvoir est de plus en plus contesté par des États qui entendent lutter contre divers maux du débat en ligne, et notamment la désinformation et les discours de haine. Les initiatives qui en découlent visent principalement à pousser les plateformes à contrôler plus efficacement les contenus qu'elles hébergent, en retirant davantage de publications qui posent problème, et en le faisant plus rapidement. Les mesures ainsi adoptées sont critiquées en raison des menaces qu'elles font peser sur la liberté d'expression des internautes : la pression exercée sur les plateformes les conduirait ainsi à sur-censurer, en retirant des contenus légitimes de peur de faire face à des sanctions financières si elles n'obtempèrent pas dans le temps imparti. Les controverses qui entourent ces initiatives ont pour autant le mérite de soumettre à la critique publique les politiques de modération des plateformes, qui ont pendant de nombreuses années géré le débat en ligne en toute opacité. Elles montrent également que le droit n'a pas dit son dernier mot en matière de régulation, et que, sur internet, la loi aussi peut faire le code.

RÉFORMER LES DISPOSITIFS DE MODÉRATION

Sur le territoire européen, plusieurs visions de la régulation des contenus sur le Web coexistent. Une première approche, parfois qualifiée de « *hard regulation* », tend à imposer aux plateformes des obligations de résultat et/ou des délais à respecter pour le retrait des contenus illégaux, tout en prévoyant des sanctions en cas de non-mise en conformité avec les nouvelles règles. C'est notamment la voie suivie par l'Allemagne avec sa loi de régulation des réseaux sociaux, dite *NetzDG*, entrée en vigueur en janvier 2018, ou par la France, à travers la loi de lutte contre les contenus haineux sur internet, dite *Avia*, finalement censurée par le Conseil constitutionnel en juin 2020. La deuxième approche, que l'on peut qualifier de « *soft regulation* », est par exemple celle adoptée par la

Commission européenne qui, à partir du printemps 2016, a mis en œuvre un code de bonne conduite relatif aux discours haineux illégaux en ligne, consistant à former un partenariat avec les plateformes qui définit des obligations de moyens, sans prévoir néanmoins de sanctions en cas de non-mise en conformité. Entre régulations « *hard* » et « *soft* » coexistent des approches intermédiaires : la loi française de lutte contre la manipulation de l'information, entrée en vigueur en janvier 2019, oblige les plateformes à retirer une fausse nouvelle dans un délai de quarante-huit heures après leur signalement, mais sur décision d'un juge des référés ; au Royaume-Uni, le projet législatif *Online Harms*, initié au printemps 2019, prévoit une simple obligation de moyens, assorti cependant d'un contrôle strict qui, en particulier, menace d'infliger des amendes aux plateformes n'obtempérant pas.

Si ces diverses politiques publiques partagent des objectifs communs, elles mettent donc en œuvre des stratégies de régulation bien différentes. Toutes établissent un certain nombre de prérequis en matière de modération de contenus (obligation de moyens), comme la définition de standards de publication, l'embauche d'une équipe de modérateurs professionnels ou la mise en place de procédures de signalement de contenus problématiques. Mais certaines vont plus loin en imposant un délai à respecter pour le retrait des publications incriminées (obligation de résultat) : vingt-quatre heures dans le cas de la *NetzDG* et de la loi *Avia*, quarante-huit heures dans le cas de la loi de lutte contre les fausses informations. La dernière grande différence tient à la définition de sanctions en cas de non-mise en conformité des plateformes. La *NetzDG* dispose que les plateformes peuvent se voir infliger une amende de 5 millions d'euros si elles échouent à retirer des contenus signalés dans le temps imparti. De la même façon, la loi *Avia*, avant sa censure par le Conseil constitutionnel, prévoyait des amendes pouvant atteindre 20 millions d'euros ou 4 % du chiffre d'affaires annuel mondial total de l'entreprise sanctionnée, le montant le plus élevé étant retenu. Des amendes sont également envisagées dans le projet législatif britannique, sans que leur montant ait encore été fixé.

Au-delà de la définition d'objectifs à atteindre pour les plateformes et de sanctions en cas de non-mise en conformité, les pouvoirs publics européens ont également entrepris une réforme du fonctionnement même des dispositifs de modération. En Allemagne par exemple, la *NetzDG* prévoit vingt et un nouveaux types de contenus prohibés devant faire l'objet d'un retrait des réseaux sociaux. Pour ce faire, elle contraint les plateformes à mettre en place des dispositifs de signalement spécifiques

concernant ces contenus et à publier tous les six mois des rapports sur la mise en application des nouvelles mesures. Ces dispositifs de signalement des contenus prohibés doivent également être aménagés pour donner lieu à davantage de transparence, en avertissant l'internaute qui a signalé, comme celui qui est signalé, de l'avancée de la procédure.

54 Pour répondre à ces exigences, YouTube et Facebook s'engagent dans un processus de mise en conformité qui se traduit par l'instauration de dispositifs de modération à deux étages. Ceux-ci consistent à supprimer de leurs plateformes les contenus qui enfreignent à la fois les standards de publication et la loi allemande, et à bloquer géographiquement les publications qui ne contreviennent qu'aux dispositions de la *NetzDG*. Concrètement, dans le deuxième cas de figure, le contenu mis en cause sera retiré du Facebook allemand, mais sera toujours accessible depuis les autres déclinaisons nationales de la plateforme. Les mesures adoptées dans le cadre de cette loi de régulation des réseaux sociaux impliquent par ailleurs, du côté des plateformes, un certain nombre d'évolutions de leurs dispositifs de modération, qui passent notamment par l'embauche et la formation de modérateurs spécialisés. Quarante-vingts modérateurs chez Facebook et soixante-treize chez YouTube ont ainsi été spécifiquement formés à l'évaluation des contenus prohibés par la *NetzDG*.

RÉGULER PAR LA TRANSPARENCE

Le pouvoir des grandes firmes du numérique sur le débat public ne s'exprime pas uniquement à travers leur capacité à autoriser ou interdire certains types de propos sur leurs plateformes. Il s'exprime aussi, comme nous l'avons évoqué précédemment, par leur capacité à mettre en forme l'expression publique (comment nous nous exprimons) et à organiser la visibilité des informations présentes en ligne (à quoi nous sommes exposés). Ces deux champs d'action font également l'objet de tentatives de régulation de la part des États, mais selon une logique différente, qui relève moins d'une menace de sanction que d'une injonction à la transparence et à l'information des usagers. L'approche n'est pas nouvelle : la loi pour une République numérique, promulguée en 2016 en France, dispose que les « opérateurs de plateforme » sont tenus à une obligation de « loyauté », qui consiste principalement à livrer à tout internaute des informations sur leurs méthodes de classement des informations et sur les liens contractuels (les publicités) qui peuvent influencer ces classements. Concrètement, il s'agit pour les plateformes de communiquer

sur les modes de fonctionnement de leurs algorithmes et d'informer clairement leurs usagers du caractère sponsorisé de certains contenus.

Cette logique est suivie par les différentes législations que nous avons présentées. Que ce soit par la *NetzDG* en Allemagne, par les lois françaises sur la manipulation de l'information et sur les discours de haine, par le projet *Online Harms* au Royaume-Uni ou par le code de bonne conduite de la Commission européenne, les plateformes sont invitées à publier des rapports explicitant le fonctionnement de leurs algorithmes de classement, la manière dont ceux-ci influencent la consommation des informations par les internautes, les moyens techniques mis en œuvre pour détecter les contenus problématiques, ou encore le fonctionnement de leurs régies publicitaires. Dans le sillage de ces législations et des auditions des dirigeants des plateformes devant le Congrès américain, Google comme Facebook ont par ailleurs mis en place de nouvelles mesures : registres des publicités politiques diffusées *via* leurs services, rapports de transparence mis à jour tous les trois mois, dispositifs d'information des internautes sur les contenus publicitaires et le système de recommandation.

55

À cet exercice d'« *accountability* » (principe de responsabilité qui vise à « rendre des comptes ») s'ajoutent diverses formes de contrôle étatique de la « loyauté » des plateformes. En France, la loi sur la manipulation de l'information dispose que la mise en conformité des plateformes vis-à-vis des nouvelles mesures doit être contrôlée par le Conseil supérieur de l'audiovisuel. Ce dernier indique aux firmes concernées les informations auxquelles il souhaite accéder, en exigeant deux fois par an des rapports répondant à ses exigences. En Allemagne, les plateformes doivent également produire tous les six mois des rapports sur leurs activités en matière de modération, dont la supervision est assurée par le ministère de la Justice, par l'intermédiaire d'organisations et d'institutions qu'il mandate. Au Royaume-Uni, le livre blanc publié dans le cadre du projet *Online Harms* prévoit la création d'une autorité publique indépendante, chargée d'établir un code de conduite relatif aux pratiques de modération sur les réseaux sociaux et de vérifier que les plateformes s'y conforment.

Le principal écueil de cette régulation par la transparence a trait à la fiabilité des données fournies par les firmes. En effet, les instances de contrôle ne disposant d'aucun moyen de vérification ou de certification de ces données, elles sont contraintes de « faire confiance » aux rapports qui leur sont transmis. Pour les plateformes, il en résulte la possibilité

d'avoir recours à une forme d'« opacité stratégique »⁶, en donnant accès seulement aux informations qu'elles souhaitent divulguer, tout en laissant dans l'ombre celles qu'elles veulent garder secrètes. Sans possibilité d'audit indépendant des dispositifs de modération, une suspicion demeure quant à l'authenticité des données contrôlées.

56 Le recours à un audit, envisagé par certaines législations présentées comme le projet britannique *Online Harms*, était également au cœur, en France, du rapport de la mission sur la régulation des réseaux sociaux. Annoncée en 2018 par le président Emmanuel Macron lors de la tenue, à Paris, de l'Internet Governance Forum, cette mission interministérielle a permis d'observer des dispositifs de modération dans les locaux de Facebook, afin de formuler, avec la firme américaine, un ensemble de propositions pour une régulation plus « efficace » des contenus. Le rapport remis au printemps 2019 défendait une approche d'« *accountability by design* » – qui vise à introduire un principe de responsabilité des plateformes dès le stade de leur conception technologique –, se traduisant par une obligation de moyens, tout en proposant que les autorités administratives chargées du contrôle des activités des plateformes bénéficient d'un pouvoir d'audit et donc d'un accès direct à leurs données. Finalement, les arbitrages gouvernementaux ont donné raison à l'équipe réunie autour de la députée Laetitia Avia, promouvant une approche plus contraignante pour les plateformes (obligation de résultat et sanctions), ce qui mit prématurément fin à la mission interministérielle.

L'enjeu essentiel que constitue la possibilité de réaliser des audits est revenu sur le devant de la scène politique à la faveur de la présentation par la Commission européenne de son *Digital Services Act*, en décembre 2020. Ce texte prévoit en effet que les « très grandes plateformes » (c'est-à-dire les services utilisés par plus de 10 % des consommateurs européens) soient tenues d'organiser des audits de leurs services une fois par an. Les rapports ainsi publiés par des organismes indépendants seraient rendus publics et les plateformes éventuellement sanctionnées pour les manquements à leurs responsabilités. Si, à l'heure où nous écrivons ces lignes, le *Digital Services Act* doit encore traverser le processus législatif européen avant d'être traduit par les parlements nationaux, il n'en demeure pas moins que les fortes exigences qu'il contient marquent

6. Mike Ananny et Kate Crawford, « Seeing without Knowing: Limitations of the Transparency Ideal and Its Application to Algorithmic Accountability », *New Media & Society*, vol. 20, n° 3, p. 973-989.

une nouvelle étape du retour des pouvoirs publics dans le domaine de la régulation des plateformes.

AGIR SUR LE *DESIGN* DES PLATEFORMES

L'impuissance des États face aux plateformes relève largement du mythe : lorsqu'une volonté politique s'affirme, dans un contexte où les médias et les opinions publiques se saisissent des enjeux de la régulation des géants du numérique, de nouvelles règles peuvent être imposées. Mais, pour être réellement efficace, la régulation des discours de haine et des fausses informations doit s'intégrer à une réflexion plus large sur la conception des plateformes. Parce que ces architectures d'échange sont conçues pour capter l'attention des usagers, leurs algorithmes de tri et de recommandation valorisent les contenus racoleurs, démagogues ou radicaux. La métrique principale à l'aune de laquelle est évaluée la popularité des contenus est par ailleurs celle de la vitesse : plus un hashtag ou une vidéo se diffusent rapidement, et plus ils sont mis en valeur par ces mêmes algorithmes, entraînant de fait un phénomène de « réchauffement médiatique »⁷ peu propice à la réflexion et au débat rationnel.

57

La régulation des plateformes doit donc également se traduire par des mesures visant leur conception technologique : limiter le comptage des « *likes* » ou des retweets par exemple, imposer aux algorithmes des méthodes de fonctionnement plus vertueuses, informer en direct les usagers sur leur activité en matière de publication afin de les inciter à en réduire le rythme⁸, filtrer l'usage de certaines expressions, ou réduire la captation des traces d'usage à des fins publicitaires. Des pistes existent pour ouvrir un nouveau chapitre de cette régulation, celui de l'économie de l'attention, afin de pacifier durablement le débat public en ligne.

7. Dominique Boullier, *Comment sortir de l'emprise des réseaux sociaux*, Paris, Le Passeur, 2020.

8. Cf. à ce sujet les propositions de Dominique Boullier (*ibid.*, p. 215-246).

R É S U M É

Alors que les États sont généralement considérés comme impuissants face aux plateformes, des initiatives en matière de régulation des réseaux sociaux menées en Europe prouvent que le droit n'a pas dit son dernier mot sur la gouvernance d'internet. À travers l'exemple de la lutte contre les discours de haine et les fausses informations, cet article explore les différents leviers d'action des États dans le domaine de la régulation des contenus en ligne.