



Créer un cadre français

de responsabilisation des réseaux sociaux :

agir en France avec une ambition européenne

Rapport de la mission

« Régulation des réseaux sociaux – Expérimentation Facebook »

Remis au Secrétaire d'État en charge du numérique

Mai 2019

Synthèse

Les réseaux sociaux permettent à tout citoyen de publier les contenus de son choix et de les partager avec d'autres utilisateurs du réseau. Ils ont ainsi révolutionné l'industrie des médias et les modes de communication en offrant aux citoyens et à la société civile un support d'expression directe. Le recours aux médias classiques n'est plus obligatoire pour communiquer publiquement. La possibilité pour les citoyens d'exercer leurs libertés d'expression, de communication, et de s'informer sont donc considérablement accrues par ces services.

Néanmoins, les capacités offertes par les réseaux sociaux suscitent des abus inacceptables de ces libertés. Ces abus sont le fait d'individus isolés ou de groupes organisés auxquels les grands réseaux sociaux, Facebook, YouTube, Twitter ou Snap, pour ne citer qu'eux, n'apportent pas de réponse pleinement satisfaisante à ce jour. Or, par l'ordonnancement qu'ils réalisent des contenus publiés et par leur politique de modération, les réseaux sociaux sont en mesure d'agir directement sur ces abus les plus manifestes pour les prévenir ou y répondre et limiter ainsi les dommages en termes de cohésion sociale.

Une intervention publique pour que les plus grands acteurs adoptent une attitude plus responsable et plus protectrice de la cohésion de nos sociétés apparaît donc légitime. Compte tenu des enjeux de libertés publiques, cette intervention doit faire l'objet de précautions particulières. Elle doit (i) respecter la diversité des modèles de réseaux sociaux, qui forment un ensemble particulièrement hétéroclite, (ii) faire preuve de transparence, notamment en y associant la société civile, (iii) viser un objectif d'intervention minimum conformément aux principes de nécessité et de proportionnalité et (iv) s'en remettre aux juridictions pour la qualification de la licéité des contenus pris individuellement.

La démarche actuelle d'autorégulation des réseaux sociaux est intéressante en ce qu'elle démontre que les plateformes peuvent faire partie de la solution aux problèmes constatés. Elles ont inventé des réponses variées et agiles : retrait, moindre exposition, rappel à la règle commune, pédagogie, accompagnement des victimes. Toutefois, l'autorégulation est toujours en développement. Elle se contente trop souvent de proposer une réponse *ex-post* (après l'apparition du dommage). Elle manque de crédibilité, du fait de l'asymétrie extrême d'information, provoquant un sentiment de « *story telling* » qui suscite une suspicion sur la réalité de l'action de la plateforme.

L'intervention publique doit reposer sur un équilibre entre une politique répressive, indispensable pour lutter efficacement contre les auteurs des abus, et une logique de responsabilisation accrue des réseaux sociaux fondée sur une régulation *ex ante*, pour capitaliser sur la logique de prévention et la capacité d'autorégulation des plateformes.

Compte tenu de l'unicité et de l'ubiquité des réseaux sociaux, qui dépassent les frontières des États membres – un service unique accessible sur plusieurs géographies – cette régulation *ex ante* doit s'inscrire dans une dynamique et un cadre européens. Néanmoins, la règle actuelle dite du pays d'*installation*, selon laquelle seul le pays qui accueille le siège du réseau social peut intervenir pour réguler ce réseau, s'avère inefficace. Par les excès qu'ils permettent, les réseaux sociaux créent des dommages dans d'autres États membres, et difficilement observables par l'État membre d'*installation*.

Toute initiative française devra donc avoir pour ambition d'inverser la logique européenne actuelle au profit d'une logique du pays de *destination*, selon laquelle la plateforme est responsable devant l'État membre où le dommage se produit, pour renforcer la capacité de chacun des États membres à maîtriser les conséquences de la globalisation. Cet objectif doit être pris en compte dès la conception d'une fonction de régulation des réseaux sociaux pour que le dispositif apparaisse pertinent aux yeux de nos partenaires européens, alors même que la politique en matière de régulation des industries de médias diffère sensiblement d'un pays à l'autre.

La mise en place d'une régulation *ex ante* devrait respecter trois conditions : (i) suivre une logique de conformité selon laquelle le régulateur supervise la bonne mise en œuvre de mesures préventives ou correctrices, sans se focaliser sur la matérialisation des risques ni chercher à réglementer lui-même le service fourni, (ii) se concentrer sur les acteurs systémiques capables de créer des dommages conséquents dans nos sociétés sans créer de barrière à l'entrée pour de nouveaux acteurs européens, (iii) rester agile pour affronter les enjeux futurs dans un environnement numérique particulièrement évolutif. Le dispositif législatif devra donc viser à créer une capacité institutionnelle à réguler et non une régulation figée sur les problèmes actuels.

Cette régulation pourrait reposer sur les cinq piliers suivants :

Premier pilier : Une politique publique de régulation garante des libertés individuelles et de la liberté d'entreprendre des plateformes.

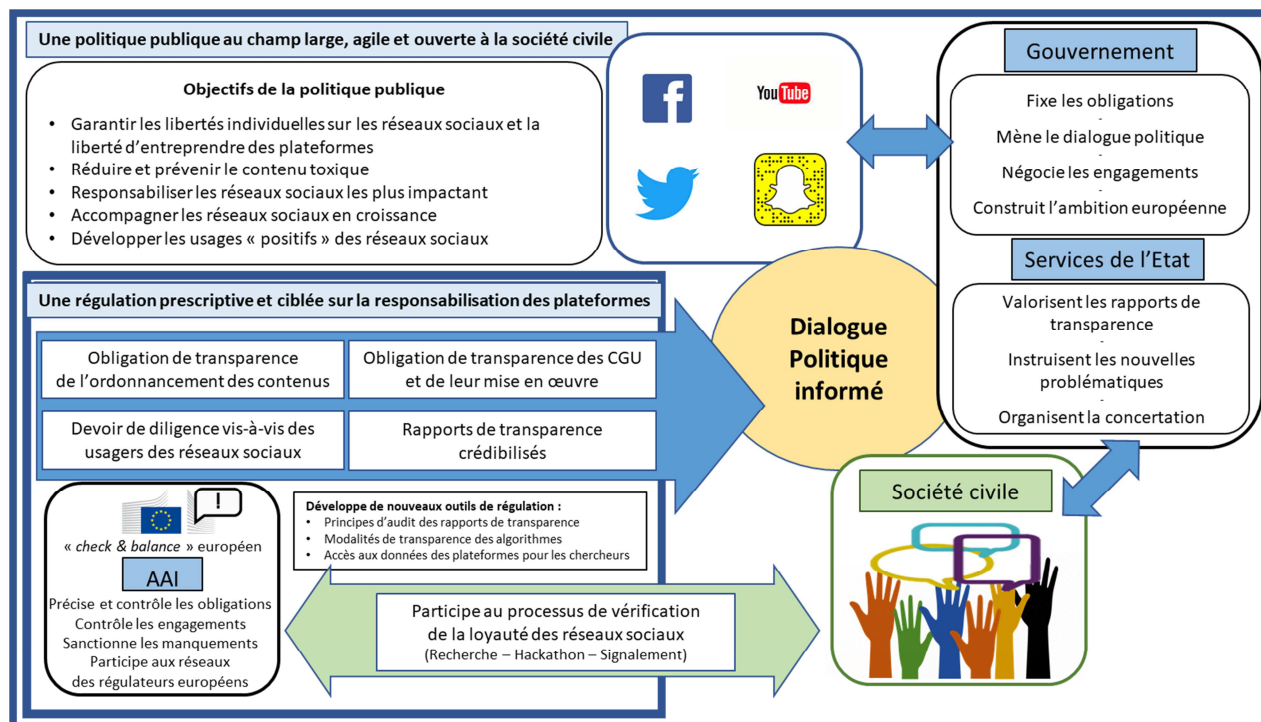
Deuxième pilier : Une régulation prescriptive et ciblée sur la responsabilisation des réseaux sociaux mise en œuvre par une autorité administrative indépendante, reposant sur trois obligations incombant aux plateformes :

- transparence de la fonction d'ordonnancement des contenus ;
- transparence de la fonction de mise en œuvre des CGU et de modération des contenus ;
- un devoir de diligence vis-à-vis de ses utilisateurs.

Troisième pilier : Un dialogue politique informé entre les acteurs, le gouvernement, le législateur, le régulateur et la société civile.

Quatrième pilier : Une autorité administrative indépendante partenaire des autres branches de l'État et ouverte sur la société civile.

Cinquième pilier : Une ambition européenne pour renforcer la capacité des Etats membres à agir face à des plateformes globales, et réduire le risque politique lié à la mise en œuvre dans chaque Etat membre.



Avant-propos

La mission a eu pour objet de préfigurer un cadre général de régulation des réseaux sociaux, en partant de la lutte contre la haine en ligne et en s'appuyant sur la collaboration volontaire, et hors de tout cadre juridique, de la société Facebook¹ (voir lettre de mission en annexe).

La finalité de cette expérimentation consiste à imaginer comment une nouvelle fonction de régulation des réseaux sociaux pourrait être mise en place pour compléter les instruments actuels et permettre de mieux atteindre les objectifs de politique publique en termes de conciliation des libertés publiques et de sauvegarde de l'ordre public sur les réseaux sociaux. Les échanges avec Facebook ont ainsi été centrés sur les contenus haineux, mais la réflexion de la mission peut s'appliquer à l'ensemble des problématiques soulevées par la publication de contenus sur les réseaux sociaux.

Cette mission² interministérielle a mobilisé sept experts de haut niveau et trois rapporteurs permanents provenant de services des ministères de la Culture, de l'Intérieur, de la Justice, de l'Économie, des services du Premier ministre – DILCRAH³ et DINSIC⁴ – et d'autorités administratives indépendantes – ARCEP⁵ et CSA⁶.

Durant les mois de janvier et de février, des échanges ont eu lieu entre la mission et la société Facebook à l'occasion de plusieurs journées de travail à Paris, Dublin (où se situe le siège européen de Facebook) et Barcelone (où se situe l'un des centres de modération). Les représentants de la société ont présenté sa politique de modération des contenus haineux, son organisation, les moyens qu'elle y consacre et ses processus internes. Des réunions de travail d'approfondissement ont eu lieu en mars sur des sujets spécifiques tels que le recours aux algorithmes de détection de contenus haineux dans la fonction de modération ou les principes de base des algorithmes d'ordonnancement des contenus pour l'utilisateur de Facebook.

La mission a bénéficié d'un accueil ouvert de la société Facebook. Néanmoins, elle n'a pas accédé à une information détaillée, ni à des éléments confidentiels en raison des délais très courts, de l'absence de cadre juridique formel, et des limites de la politique de transparence de la société Facebook. Pour autant, cet état de fait est sans conséquence pour les travaux de la mission dont l'objectif n'est pas d'évaluer la pertinence des mécanismes de Facebook mais d'imaginer des « règles du jeu⁷ », qui pourraient être instituées par le législateur pour créer un cadre de régulation dans la durée d'acteurs globaux établis à l'étranger tels que la société Facebook.

Dès lors, le rapport ne détaille pas les mécanismes de modération de Facebook en matière de lutte contre la diffusion de contenus haineux en ligne. Néanmoins, le lecteur peut se référer aux documents publiés par Facebook afin de mieux appréhender les mécanismes de modération du réseau social et notamment les standards de la communauté, le rapport de transparence sur la gestion des contenus sur sa plateforme, et

¹ A la suite un accord entre son président-directeur général, Mark Zuckerberg, et le Président de la République annoncé au Forum sur la gouvernance de l'Internet en novembre 2018.

² La composition de la mission figure en annexe du présent rapport.

³ Délégation interministérielle à la lutte contre le racisme, l'antisémitisme et la haine anti-LGBT.

⁴ Direction interministérielle du numérique et du système d'information et de communication de l'État.

⁵ Autorité de régulation des communications électroniques et des postes.

⁶ Conseil supérieur de l'audiovisuel.

⁷ A ce titre, la mission est convaincue que le pouvoir du régulateur résidera dans sa capacité à exiger la communication de toute information nécessaire à l'accomplissement de sa mission dans un cadre juridique opposable.

le [blog « hard questions »](#) sur lequel Facebook publie régulièrement des réflexions sur le sujet de la modération (en anglais).

La mission a également présenté sa démarche aux associations engagées dans la lutte contre les discours haineux lors d'un séminaire et d'un atelier citoyens organisés par le Conseil national du numérique les 14 et 15 février 2019.

La mission a complété ses travaux par un voyage d'étude à Berlin pour mieux comprendre l'expérience de la loi allemande NetzDG, une mission à Londres et une série de réunions avec des acteurs publics (Inria, Plateforme Pharos, Centre de lutte contre les criminalités numérique, Secrétariat général aux affaires européennes, Conseil national du numérique, Direction générale des entreprises, Direction générale du Trésor), et des acteurs privés ou associatifs (Reporters sans frontières, le CERRE, la Quadrature du net, Webedia, Netino, Snap, Google, et Twitter).

Ce rapport formule des propositions qui, si elles sont retenues, devront être approfondies. Il (i) identifie les caractéristiques essentielles des services de réseaux sociaux, (ii) analyse les dynamiques de politique publique vis-à-vis de ces services, (iii) propose de structurer une nouvelle fonction de régulation autour de cinq piliers et (iv) présente un focus sur le concept de transparence des algorithmes et de sa mise en œuvre.

Compte tenu des contraintes de calendrier, plusieurs sujets n'ont pas été examinés. En particulier, la mission n'a pas réalisé d'étude d'impact concurrentiel du schéma de régulation proposé sur les autres offres de services de réseaux sociaux. Or, le dispositif de régulation devra veiller à ne pas créer une barrière à l'entrée insurmontable pour des acteurs de taille intermédiaire ou de nouveaux entrants, et, par voie de conséquence, de favoriser indûment les acteurs en position hégémonique en consacrant, par l'édiction de barrières réglementaires.

Par ailleurs, la mission a centré sa réflexion sur les contenus publics, alors même que les contenus haineux sont présents aussi au sein des groupes privés ou fermés des réseaux sociaux, et qu'il existe actuellement une tendance à la diffusion de contenus dans ces groupes restreints et sur des services de messagerie. Il est plus complexe d'intervenir sur ces environnements dont les échanges peuvent être couverts par le secret des correspondances privées et surtout chiffrés « de bout en bout », rendant illusoire toute modération par la plateforme elle-même dès lors qu'elle n'a aucune visibilité sur le contenu échangé entre les utilisateurs.

Enfin, le rapport ne traite pas des modalités d'intervention pour des réseaux sociaux « non-coopératifs », qui ne répondent pas à une rationalité économique classique qu'ils soient militants extrémistes (tels que 4chan ou 8chan) ou contrôlés, directement ou indirectement par un État souverain poursuivant des objectifs politiques.

I – Les services de réseau social

En permettant à tout un chacun de publier des contenus et de les partager avec d'autres utilisateurs, les réseaux sociaux ont révolutionné l'industrie des médias et les modes de communication en offrant aux citoyens et à la société civile un support d'expression directe. Néanmoins, les possibilités offertes par les services de réseau social suscitent des abus inacceptables de la part d'individus isolés ou de groupes organisés auxquels les acteurs n'apportent pas de réponse pleinement satisfaisante voire y contribuent *via* leur fonction d'ordonnement des contenus.

1.1 Les services de réseau social ont tous pour objet le partage et la diffusion de contenus au public en ligne mais constituent néanmoins un ensemble hétéroclite

Les services de réseau social se définissent par la capacité de diffuser des contenus produits par leurs utilisateurs à destination de tout ou partie des autres utilisateurs de ce réseau.

Les services de réseau social sont fournis par différents types d'acteurs qui se distinguent par leur statut juridique, leur nature, leur cible, leur modèle économique, le type de contenus publiés⁸ et leurs modalités de diffusion. Il se décompose en deux catégories :

- les services de réseaux sociaux proposés à titre accessoire : les forums de discussions thématiques ou libres des sites web (jeuxvideo.com et les espaces de commentaires sur les sites de presse lemonde.fr, lefigaro.fr), constituent la forme basique d'un réseau social : l'organisation des contenus y est sommaire (principalement par ordre chronologique) et, lorsqu'ils font l'objet d'une monétisation par la publicité, celle-ci n'est généralement pas mélangée aux contenus des utilisateurs ; elle est adjacente ;
- les services de réseaux sociaux proposés à titre principal : les plateformes de réseaux sociaux, qui peuvent être généralistes comme Facebook, Twitter ou structurées autour d'un type ou un format de contenu comme YouTube (vidéos), Pinterest (photos), TikTok (courtes vidéos), Snapchat (courtes vidéos et photos).

Tous ces services proposent tout ou partie des contenus et fonctionnalités suivants : *user-generated contents* (UGC) comme contenu principal, contenus promotionnels⁹, contenus émanant d'éditeurs professionnels, contenus accessibles à tous et/ou contenus faisant l'objet d'une diffusion restreinte à un groupe de personnes, comptes individuels avec ou sans pseudonyme, espaces de discussions attachés à une communauté ou à un événement.

La fonction d'ordonnement des contenus sur un service de réseau social peut être individualisée (c'est-à-dire propre à chaque utilisateur). Elle est plus ou moins sophistiquée selon le volume de contenus publiés.

Les plateformes combinent de plus en plus au service de réseau social, un service de messagerie privée : Direct message pour Twitter, Messages pour YouTube, Messenger (et WhatsApp) pour Facebook. Dans certains cas comme pour Snap, le service de diffusion d'un contenu à un groupe fermé peut être le service proposé par défaut mais l'utilisateur peut à tout moment ouvrir la diffusion à l'ensemble du réseau.

⁸ Le contenu peut être un texte, un lien hypertexte, une image, un son, une vidéo (parfois en temps réel), un programme informatique et/ou une combinaison quelconque de ces six éléments.

⁹ Il peut s'agir de publicité classique, de placement de produit, de contenus sponsorisés, etc.

Les modalités de monétisation sont extrêmement variables d'un service à l'autre : publicités indépendantes des utilisateurs et des contenus partagés, ciblées sur le contenu auquel elles sont attachées, ciblées sur la personne qui consulte le contenu. Les recettes peuvent être partagées avec ceux qui publient le contenu comme dans le cas de YouTube ou Snap.

Au-delà de la distinction par type d'acteurs et par fonctionnalité, la taille du réseau social constitue un élément central à prendre compte : 2 milliards d'utilisateurs pour Facebook, présents dans le monde entier, quelques milliers de personnes sur certains forums de discussion.

Enfin, les modèles et les usages ne sont pas figés et évoluent rapidement : les services génèrent de nouveaux usages de même que les comportements des utilisateurs incitent les acteurs à adapter en permanence leurs services.

Une première tentative de définition juridique d'un service de réseau social a été introduite dans la proposition de loi de la députée Avia, visant à lutter contre la haine sur internet. Elle isole, dans la définition des opérateurs de plateforme en ligne posée à l'article L111-7 du code de la consommation, les opérateurs de plateformes en ligne « *proposant un service de communication au public en ligne reposant sur la mise en relation de plusieurs parties en vue du partage de contenus publics* »¹⁰. Au niveau européen, la directive SMA introduit par voie incidente la notion de « services de médias sociaux ». Elle définit dans son article 1 les « services de médias audiovisuels » et précise dans les considérants (4) et (5) que les services de médias audiovisuels sociaux sont ceux dont les contenus sont créés par les utilisateurs¹¹.

Un réseau social peut être défini comme un service en ligne permettant à ses utilisateurs de publier les contenus de leur choix et de les rendre ainsi accessibles à tout ou partie des autres utilisateurs de ce service.

1.2 Les réseaux sociaux ont révolutionné l'industrie des médias et les modes de communication en offrant aux citoyens et à la société civile un support d'expression directe. Ils constituent en cela une formidable avancée pour la liberté d'expression

Disposant d'une capacité à héberger et diffuser des contenus en masse pour un coût marginal très faible, les réseaux sociaux constituent une *nouvelle forme de média d'expression directe, sans présélection des contenus un à un, ni intermédiation journalistique*. Le réseau social permet l'échange de contenus qu'il n'a pas créés, sous réserve du respect des règles de publication édictées par le réseau social (cf. infra). Cette absence de création et de sélection contenu par contenu, qui distingue les réseaux sociaux des médias d'information traditionnels, *permet à tout un chacun d'une part de s'exprimer, de rendre publiques et de diffuser ses opinions ou les contenus de son choix et d'accéder à de nouvelles sources d'information*. La capacité d'un individu, d'une association ou des acteurs privés ou publics à s'exprimer dans le débat public n'est plus tributaire des choix éditoriaux des médias traditionnels.

¹⁰ Elle se distingue de la définition posée par le législateur allemand : la NetzDG allemande, adopté en 2017 définit les réseaux sociaux des plateformes internet, à but lucratif, destinées à permettre aux utilisateurs de partager tout contenu avec d'autres utilisateurs, ou de rendre ce contenu accessible au public. Sont exclues les plateformes gérant du contenu éditorial ou journalistique. La loi n° 2018-1202 du 22 décembre 2018 relative à la lutte contre la manipulation de l'information vise tous les opérateurs de plateformes en ligne au sens du code de la consommation.

¹¹ Lorsque le partage de vidéos en constitue une « fonctionnalité essentielle », ces services entrent dans la catégorie des « plateformes de partage de vidéos », distincte de celle des « services de médias audiovisuels », et soumise à une régulation principalement destinée à protéger les jeunes publics et à lutter contre la diffusion de contenus haineux.

Ces réseaux créent *de nouvelles formes de relations sociales*, dépassent les contraintes géographiques (voire linguistiques grâce aux outils de traduction) et bousculent à la fois les structures sociales historiques et la primauté de l'organisation territoriale des Etats et de nos sociétés. De nouvelles formes d'associations « numériques » de citoyens, des communautés dématérialisées mais bien réelles, sont apparues ainsi pour partager des informations, un centre d'intérêt ou se fédérer autour d'une cause à défendre.

Les opportunités offertes par ces nouveaux vecteurs de communication se traduisent dans les usages. Les réseaux sociaux occupent une place désormais centrale dans l'accès à l'information et sa diffusion. Un tiers des Français et la moitié des 18-24 ans s'informent *via* les réseaux sociaux et les plateformes de partage de vidéos représentent la moitié de la consommation de vidéos d'information sur internet.¹²

1.3 Les services de réseaux sociaux définissent un ordonnancement des contenus et exercent ainsi une *forme d'éditorialisation de fait et non de droit*, généralement non observable et non transparente

L'ensemble des contenus publiés sur un réseau social ne peut être présenté à l'utilisateur sans ordonnancement. Le volume de contenus publiés implique nécessairement que la plateforme définisse un ordre d'apparition, opère une sélection, tout en laissant à l'utilisateur la possibilité d'aller chercher, à son initiative, un contenu spécifique. Les contenus qu'il consultera *effectivement* dépendront en premier lieu de l'agencement de son interface et du recours à des règles algorithmiques pour hiérarchiser et individualiser la présentation des différents contenus. Contrairement à des médias traditionnels, l'ordonnancement des contenus sur les services de réseau social est généralement individualisé (sauf dans le cas des forums) et chacun voit le résultat de l'individualisation lorsqu'il accède au service. En revanche, les effets globaux de cet ordonnancement sur l'ensemble des utilisateurs ne sont pas observables.

De plus, les acteurs fournissant des services de réseau social ne révèlent pas toujours les critères selon lesquels les contenus sont mis en avant. Ces critères peuvent être nombreux et leur poids varie selon la finalité du service (intérêt supposé du contenu, identité de l'auteur, existence d'une rémunération, préférences et usages de l'utilisateur etc.). Plus généralement, la fonction d'ordonnancement confère aux acteurs fournissant un service réseau social la capacité d'*accélérer* la diffusion de certains contenus, ou au contraire, d'en *ralentir* la propagation.

Le constat de l'existence de cette fonction d'ordonnancement des contenus, qui constitue une forme d'*éditorialisation de fait* ne saurait remettre en cause le statut juridique de ces acteurs, ni conduire, à les requalifier en éditeurs dès lors que la majorité des services de réseau social n'opère pas de sélection préalable à la publication des contenus.

L'existence de cette fonction de structuration de l'information joue un rôle essentiel dans la diffusion de contenus et dans la capacité des réseaux sociaux à prévenir ou accentuer des dommages en matière de cohésion sociale.

¹² Reuters Institute, *Digital News Report 2018*.

1.4 Les libertés de communication et d'expression publiques offertes par les acteurs proposant un service de réseau social entraînent des abus inacceptables de la part d'individus isolés ou de groupes organisés auxquels les réseaux sociaux n'apportent pas de réponse suffisante

Que les contenus publiés sur les réseaux sociaux soient payants, gratuits ou rémunérés, la majorité d'entre eux ne pose aucune difficulté¹³. En revanche, fruits de cette capacité de communication et d'expression à grande échelle, couplés d'un sentiment de relatif anonymat et d'impunité, les réseaux sociaux sont aussi le lieu d'échanges de contenus ou comportements inacceptables : contenus incitant à la haine, contenus terroristes, pédopornographiques, harcèlement en ligne, usurpation d'identité, etc. On observe également des impacts sur le lien social à travers la diffusion de fausses informations et de rumeurs infondées, des tentatives de manipulation frauduleuse de l'opinion publique du fait d'individus ou de groupes poursuivant des objectifs politiques ou financiers.

La plupart des acteurs ont mis en place des règles d'utilisation (CGU) indiquant les catégories de contenus acceptées sur le service, et des mécanismes de modération lorsque ces règles ne sont pas respectées par les utilisateurs. Toutefois, face au volume de contenus publiés et à l'approche statistique des outils algorithmiques, les acteurs fournissant des services de réseau social sont aujourd'hui incapables de prévenir tout risque d'abus sur leurs services. En effet, les efforts déployés se révèlent encore largement perfectibles, notamment pour ceux disposant d'une forte audience. De plus, peu d'informations sont rendues publiques s'agissant des modalités de définition et de mise en œuvre des CGU et de la fonction de modération.

La lutte contre la diffusion de contenus dommageables pour les utilisateurs et la cohésion sociale sur les services de réseau social implique de s'intéresser à la fois aux mécanismes d'élaboration des règles d'utilisation, aux mécanismes de modération *a posteriori* et, potentiellement, à leur fonction d'ordonnancement notamment lorsque celle-ci implique une individualisation des traitements.

Si les dérives proviennent des utilisateurs, le rôle des réseaux sociaux dans la présentation et la mise en avant sélective des contenus, l'insuffisance constatée de leurs mécanismes de modération et le manque de transparence du fonctionnement de leur plateforme justifient une intervention des pouvoirs publics, nonobstant les efforts consentis par certains acteurs.

L'élaboration de politiques publiques visant à prévenir les abus et les détournements d'usages des réseaux sociaux apparaît donc nécessaire mais devra faire l'objet de précautions particulières à plusieurs égards.

D'une part, il conviendra de tenir compte de la diversité des acteurs proposant ce type de service et de se concentrer, au moins dans un premier temps, sur les plus structurants en matière de cohésion sociale.

D'autre part, toute intervention étatique devra être strictement nécessaire, proportionnée et transparente dès lors qu'elle a trait à des libertés publiques aussi fortes que la liberté d'expression et la liberté de communication.

¹³ A titre d'exemple, sur 10 000 vues de contenus sur Facebook, entre 23 et 27 contiendraient des scènes de violence explicite (Rapport de transparence de Facebook, chiffres pour le troisième trimestre 2018).

II – Créer une nouvelle dynamique de politique publique

L'insuffisance et le manque de crédibilité des démarches d'autorégulation développées par les plus grandes plateformes justifient une intervention publique visant à les responsabiliser. Cette intervention doit reposer sur un équilibre entre la répression des auteurs de contenus dommageables et la régulation *ex ante* pragmatique et souple des acteurs fournissant des services de réseaux social, dans un cadre européen à repenser.

2.1 Des plateformes qui s'essaient à une autorégulation dont on peut mesurer les limites

Les échanges avec Facebook, complétés par des discussions avec d'autres acteurs montrent que les plateformes s'efforcent de développer une démarche d'autorégulation.

Dans le cas de Facebook, la mission a pu constater que la plateforme s'est dotée de mécanismes d'autorégulation grâce à des moyens dédiés croissants :

- publication récente d'une version détaillée des « standards de la communauté » ;
- accroissement des moyens humains et développement d'algorithmes de traitements massifs dédiés à la fonction de modération ;
- développement en cours d'outils de modération « distribuée » à la disposition des usagers ;
- structuration interne de la fonction de modération, publication de rapports de transparence ;
- réflexion sur la mise en place de structures de gouvernance ouverte au-delà des seuls représentants de la plateforme à savoir notamment un conseil de surveillance composé d'experts indépendants ayant pour mission de passer en revue les décisions de modération.

Quant à YouTube, les mécanismes d'autorégulation de la plateforme de partage de vidéos de Google intègrent des volets pédagogiques à destination des utilisateurs de la plateforme sur les comportements proscrits, ou à destination des victimes de comportements agressifs d'autrui. A titre d'exemple, la plateforme a mobilisé des influenceurs pour faire évoluer le comportement des utilisateurs et notamment des plus jeunes d'entre eux. L'évaluation de telles initiatives *a priori* positives doit cependant venir confirmer ou infirmer leur efficacité.

Par ailleurs, la mission a pu observer que la fonction de modération ne passe pas uniquement par le retrait de contenus considérés comme nocifs, mais qu'il existe toute une palette de réponses possibles selon le type de contenu et le degré de dommage potentiel causé par ce contenu (mise en en « quarantaine », cache avec un message de prévention, déréférencement, mise en garde, etc.).

La vitesse de déploiement et les efforts consentis durant les douze derniers mois par un acteur comme Facebook permettent d'anticiper le bénéfice à capitaliser sur les démarches d'autorégulation déjà mises en œuvre par les plateformes, en les amplifiant et en les crédibilisant.

La capacité d'autorégulation observée chez ces acteurs fournissant un réseau social permet de les positionner comme les éléments clés de la solution aux enjeux de cohésion sociale soulevés par la présence de certains contenus sur ces plateformes.

Cette solution ne peut se résumer au seul retrait des contenus manifestement illicites, mais doit s'enrichir pour agir en amont du dommage (prévention) et répondre à toutes les situations possibles, selon leur degré de gravité et de risque pour les utilisateurs : mise en quarantaine, décélération, démonétisation, rappel au règlement de la communauté, pédagogie ciblée, etc.

Néanmoins, même en anticipant le plein effet de la volonté affichée des acteurs, la mission a pu constater que celle-ci ne pourrait constituer une réponse suffisante aux préoccupations de politique publique :

- **L'asymétrie extrême d'information entre, d'une part, les acteurs proposant un service de réseau social et, d'autre part, les autorités publiques et la société civile, affaiblit considérablement la crédibilité des démarches d'autorégulation**

Ni les pouvoirs publics, ni la société civile ne savent quelle valeur accorder aux déclarations des réseaux sociaux. Ils partagent, à quelques exceptions près, le même niveau d'information qu'un utilisateur. L'ensemble des informations rendues publiques par la plateforme concernant son action d'autorégulation ne peut être corroborée par aucun fait observable. Cette limite est consubstantielle au fonctionnement des principaux services de réseau social en raison de l'individualisation des contenus proposés. La création de compte sur la plateforme ne permet d'en observer qu'une infime fraction. Seule la plateforme est à même de mesurer les effets à l'échelle globale.

Ce défaut de crédibilité est accru par le volume gigantesque de contenus et d'utilisateurs des réseaux sociaux, qui impliquent obligatoirement un traitement algorithmique dans une approche statistique. Faute de pouvoir caractériser l'existence d'une défaillance systémique du réseau social, les autorités publiques comme les représentants de la société civile ne peuvent que pointer des cas individuels de contenus non ou mal modérés. Ces défaillances isolées ne sont néanmoins pas suffisantes pour caractériser une potentielle défaillance systémique.

L'insatisfaction persistante des autorités publiques s'explique notamment par leur incapacité à évaluer concrètement la réalité et la valeur de l'autorégulation mise en œuvre par les réseaux sociaux, faute d'information validée par un tiers de confiance.

- **L'autorégulation demeure trop « aut centrée »**

Sans doute du fait de sa jeunesse, l'autorégulation ne convainc pas car les acteurs proposant un service de réseau social jouent tous les rôles simultanément : ils édictent leurs CGU, décident dans quelle mesure elles les lient, les modifient en tant que de besoin sans procédure publique, en assurent l'interprétation en première et dernière instances, rendent compte de leur mise en œuvre dans la forme et la périodicité qui leur semblent opportunes.

Parce qu'ils sont récents, concurrents et disparates, les services de réseau social développent chacun de leur côté leur propre modèle d'autorégulation. La crédibilité minimum normalement apportée par une approche sectorielle, qui permet un « contrôle par les pairs » (comme la démarche de l'Autorité de régulation professionnelle de la publicité), est ici absente et ne semble pas envisagée.

- **L'autorégulation est agile mais échappe à tout principe de supervision**

Les plateformes de réseaux sociaux sont agiles. Elles se sont développées dans un esprit entrepreneurial qui ignore, parfois à dessein, certaines contraintes réglementaires pour ne pas brider leur capacité à innover. Elles testent en permanence l'efficacité et la pertinence de leur interface utilisateur, de leurs algorithmes, de l'organisation de leur fonction de modération notamment avec des tests A/B¹⁴. Renoncer volontairement, même partiellement, à ce qui a constitué leur principale force relève donc encore de la gageure pour leurs équipes de directions.

¹⁴ Technique consistant à tester une fonctionnalité, un rédactionnel, une interface graphique, un nouvel algorithme, sur deux groupes d'utilisateurs distincts pour en mesurer ses effets.

Aujourd'hui, on peut légitimement considérer que les principaux réseaux sociaux développent une approche d'autorégulation, non pas pour internaliser et assumer pleinement des objectifs d'intérêt général de politiques publiques, mais plutôt pour limiter le risque d'intervention coercitive de la puissance publique et contenir la pression de la société civile afin de ne pas endommager leur réputation. Dans ce contexte, toutes les initiatives prises, quoique pertinentes, manquent de crédibilité et sont difficiles à évaluer.

Aujourd'hui, à la connaissance de la mission, aucun réseau social ne s'est donné de règles réellement opposables en termes d'information des utilisateurs sur les CGU, les mécanismes pour les modifier, et pour associer la société civile ou les autorités publiques à leur évolution, même à titre consultatif. La crédibilité de la démarche d'autorégulation pâtit clairement de cet état de fait.

L'approche consistant à placer le réseau social au cœur du modèle de régulation semble être pertinente. Selon ce modèle, la plateforme de réseau social intègre des objectifs d'intérêt général, modifie son organisation, s'adapte à cet objectif « social » et agit soit en amont, dès la conception de son service pour prévenir les difficultés, débordements et autres détournements d'usage du réseau social, soit en réaction à des comportements inacceptables de ses utilisateurs.

Pour reprendre la formule « *Privacy by design* » du RGPD en matière de données personnelles, on pourrait parler de « *Accountability by design* » pour le traitement des contenus par les réseaux sociaux.

2.2 Inventer une politique publique de responsabilisation des plateformes

Le modèle d'intervention publique ne viserait donc pas à réglementer l'activité, c'est-à-dire à imposer des contraintes fonctionnelles ou techniques sur les services fournis, mais à responsabiliser les acteurs proposant un service de réseau social par une obligation, opposable juridiquement, de mettre en place des moyens et d'en rendre compte. Un tel modèle, en ce qu'il permet de minimiser l'immixtion publique dans le fonctionnement d'une industrie de média dont l'essence est de servir de support à l'expression individuelle, aurait aussi comme vertu de minimiser la portée de la critique tenant au risque de manipulation de l'information par la puissance publique. Cette critique, inhérente à l'essence même de l'industrie ne doit pas dissuader l'intervention publique mais elle appelle des précautions particulières.

Par ailleurs, les modes de régulation plus directs, telles qu'ils existent dans les industries de l'énergie, des transports, des télécommunications, des médias audiovisuels traditionnels ou des jeux en ligne, apparaissent moins adaptés. En effet, ils supposent des activités clairement rattachables à un territoire national donné et donc à la compétence d'un régulateur unique. A l'inverse, les réseaux sociaux peuvent s'inscrire dans des géographies qui dépassent les frontières des Etats. Une discussion sur un contenu exprimé en français incitant à la haine des réfugiés, publié par un utilisateur situé hors de l'Union européenne peut faire l'objet de commentaires, eux aussi potentiellement haineux, par un ensemble d'autres utilisateurs francophones situés dans le monde entier.

Un mode d'intervention reposant sur des mécanismes de co-régulation, qui imposerait une internalisation d'objectifs d'intérêt général, sans en prescrire les modalités, permettrait de limiter l'impact sur les services de réseaux sociaux.

Ce mode d'intervention publique centré, d'une part, sur la création d'un devoir de diligence des réseaux sociaux vis-à-vis de leurs membres et, d'autre part, sur la crédibilisation de l'autorégulation, ne remettrait pas en cause les éléments constitutifs des réseaux sociaux que sont leur unicité, leur ubiquité et leur agilité.

2.3 Une dynamique de politique publique qui doit reposer sur un équilibre entre une politique répressive et une logique de responsabilisation accrue des acteurs par la régulation *ex ante*

Dans plusieurs pays d'Europe, le premier niveau de réponse des politiques publiques aux défaillances constatées sur les réseaux sociaux a consisté à mettre en place ou à renforcer le volet répressif, tant vis-à-vis des auteurs de contenus jugés illégaux que des réseaux sociaux eux-mêmes, qui, parce qu'elles véhiculent et accueillent ses contenus, apparaissent comme responsables au même titre que l'auteur, ou à tout le moins comme « complices » de celui-ci.

La politique répressive, à travers la mise en œuvre d'une politique pénale, est nécessaire en ce qu'elle traduit de manière claire et visible les règles que la société se donne. Elle s'impose aussi sur un plan purement politique dans les situations de troubles manifestes à l'ordre public. La politique répressive n'est efficace que si elle est lisible et appliquée, afin d'éviter tout sentiment d'impunité des auteurs de contenus inacceptables.

La logique répressive, notamment en ce qu'elle ouvre la voie à une reconnaissance du préjudice et à sa réparation est un levier essentiel mais ne permet toutefois pas d'atteindre tous les objectifs de politique publique. Elle trouve ses limites en agissant de fait *ex post* pour sanctionner un comportement illégal reconnu comme tel par une juridiction. Les capacités d'action dévolues aux autorités répressives ou civiles et le temps législatif ne permettent pas aujourd'hui d'anticiper les mutations des réseaux sociaux et les troubles qu'ils sont susceptibles de créer.

Les réseaux sociaux, contrairement aux médias traditionnels, ne sélectionnent pas contenu par contenu ce qui est publié sur le service. Il s'agit là d'une caractéristique constitutive de ces services. L'action répressive à leur encontre soulève donc plusieurs difficultés. Le réseau social se retrouve mis dans une position de censeur *ex post* de la parole de tous ses utilisateurs, essentiellement après signalement (*via* l'interface de la plateforme elle-même) ou notification (régime spécifique de la LCEN) des usagers ou des autorités publiques. En imposant un standard de conformité absolu qui ne tient compte ni du volume de contenus publiés, ni de leur audience, ni de la nature statistique des traitements, l'action répressive contient le risque d'une incitation à la sur-modération qui contrevient à la liberté d'expression, constitutionnellement et conventionnellement protégée.

Par ailleurs, la logique répressive impose aux plateformes de juger eux-mêmes de l'illégalité manifeste d'un contenu. Ces dernières estiment que cette illégalité manifeste est difficile à évaluer au regard des différents niveaux de lecture que constituent la volonté du législateur, la pratique du ministère public – par définition plus sélective en vertu du principe d'opportunité des poursuites, et la jurisprudence des juridictions nationales et internationales reposant sur une mise en balance de la liberté d'expression et de l'impératif d'ordre public. Pour autant, l'édition de « lignes directrices sur les contenus manifestement haineux » par une autorité administrative, fût-elle indépendante, n'apparaît pas comme une solution satisfaisante.

Enfin, le périmètre des contenus qui ne sont pas « manifestement illicites » (« zone grise ») varie suivant les géographies et se prête assez mal à une harmonisation européenne ou internationale, particulièrement sur des contenus relevant d'une éventuelle qualification d'incitation à la haine, en raison des spécificités historiques, culturelles et législatives propres à chaque Etat. D'autant plus que l'action répressive est souvent mise en place ou renforcée après une situation de crise qui suscite une vive émotion de l'opinion et qui appelle et rend possible une réponse politique forte. Sauf cas exceptionnels, ces situations de crise sont locales, et ne dépassent pas les frontières (ou de manière marginale). Les conditions d'une harmonisation supranationale des réponses répressives sont donc très difficiles à réunir.

Il convient de compléter l'action répressive par la mise en place d'un deuxième volet de politique publique visant à rendre les plateformes plus responsables, en établissant des obligations de transparence et de diligence des réseaux sociaux vis-à-vis de leurs membres, en créant des incitations ciblées et lisibles.

2.4 Une articulation européenne à repenser

Dans l'état actuel de la construction européenne, le lien social s'établit en premier lieu au niveau des Etats membres. Quand bien même les services sont globaux, les dommages qui découlent de leur existence se matérialisent généralement à un niveau national. En Europe, du fait de la disparité de langues, les communautés qui se forment sur les réseaux sociaux répondent tout d'abord à des logiques linguistiques et le plus souvent sur des bases nationales ou infranationales.

Il convient d'organiser la responsabilisation des acteurs proposant un service de réseau social au niveau des États membres, qui sont plus directement affectés par des débordements, qu'au niveau européen, qui reste trop « éloigné » des situations de crise et de leurs conséquences en matière d'ordre public et de paix sociale.

Toutefois, l'Union européenne, par l'espace juridique cohérent qu'elle offre, est en position unique pour permettre aux Etats membres d'agir de façon cohérente vis-à-vis d'acteurs globaux et dotés d'une forme d'ubiquité géographique. Face à la puissance des grands réseaux sociaux, l'Union européenne offre la capacité de conjuguer le poids des Etats membres face à ces acteurs économiques en leur imposant la même norme de régulation.

L'Union européenne est aussi dans une position privilégiée pour réduire les facteurs d'échecs ou de politiques publiques de régulation excessive en réduisant le risque politique au niveau de chaque Etat membre pris individuellement. Cette capacité a été démontrée notamment dans le secteur des télécommunications où les politiques de transformation des monopoles publics en industrie concurrentielle ont bénéficié de la capacité européenne à modérer des décisions de régulation parfois excessives, et à pallier des inerties nationales.

Néanmoins, les possibilités de tirer parti de la construction européenne nécessitent une nouvelle ambition :

- le dialogue entre les acteurs, les Etats membres et la Commission européenne sur le suivi de l'autorégulation de contenus illicites semble porter ses fruits au regard des résultats de la quatrième évaluation du code de conduite de l'Union européenne sur la lutte contre les discours de haine illégaux en ligne, publiée en février dernier¹⁵. Néanmoins, des progrès peuvent sans doute encore être faits : cette initiative souffre de la distance trop forte entre l'échelle européenne de mise en œuvre et la localisation des dommages. Par ailleurs, si la nature des engagements prévus dans le cadre de cette démarche est pertinente, le statut même de l'engagement volontaire dépourvu de mécanismes de sanction montre ses limites (Cf. supra) ;
- le cadre européen de régulation s'est construit autour du principe du pays d'origine, confiant le rôle de régulateur au seul Etat membre dans lequel le service a établi son siège social. Ainsi, la capacité de chacun des Etats membres, hormis celui qui accueille le service, à faire face aux éventuels errements d'un acteur global se trouve drastiquement réduite par cette articulation européenne, accroissant ainsi le risque politique dans le pays de destination (dans lequel le dommage se produit). Or les autorités du pays où une défaillance du service a été observée sont les mieux placées pour établir et corriger cette défaillance, *a fortiori* lorsque qu'elle met en jeu la liberté d'expression (par exemple à travers la publication de contenus haineux) appréciée au regard du contexte social, politique, culturel et historique de l'Etat affecté. De plus, le risque politique entre États membres est accru puisqu'un seul État membre bénéficie de l'installation de la plateforme sur son territoire, ce qui réduit son incitation à intervenir vis-à-vis d'une plateforme

¹⁵ http://europa.eu/rapid/press-release_IP-19-805_fr.html

défaillante tandis que les autres Etats membres en subissent les dommages potentiels et restent impuissants à intervenir ;

Cette architecture fondée sur la compétence du pays d'origine se retrouve dans la directive SMA et dans la version du projet de règlement de lutte contre la diffusion de contenus terroristes adoptée par le Conseil des Etats membres en décembre dernier.

La mise en place d'une régulation européenne fondée sur le principe de compétence du pays d'installation déposséderait les Etats membres de leur capacité à agir sur les services de réseaux sociaux présents sur leur territoire.

La mise en place d'une régulation nationale par chaque Etat membre aboutirait à un risque élevé d'incompatibilité des dispositifs nationaux entre eux du fait de l'unicité et de l'ubiquité des services de réseaux sociaux. Ceux-ci seraient alors soumis à des injonctions contradictoires et donc inopérantes. Ces dispositifs nationaux pourraient potentiellement être exposés au à un risque d'inconventionnalité.

A l'inverse, une articulation reposant sur un règlement européen et sur le principe du « pays de destination » pourrait renforcer la capacité de chaque Etat membre à faire face aux difficultés générées par des acteurs globaux comme les plus grands réseaux sociaux tout en réduisant le risque politique pour ces acteurs :

- la responsabilisation de la plateforme vis-à-vis du pays de destination crée une incitation pour chaque réseau social à prévenir les débordements dans chacune des géographies. Cela rétablit la cohérence géographique entre localisation de la fonction de régulation et localisation des dommages ;
- un cadre commun, qui pose des obligations uniformes, défini au niveau européen via un règlement d'application directe, garantit cohérence et uniformité de la norme juridique dans l'ensemble des géographies. Celle-ci est essentielle vis-à-vis d'acteurs globaux et dotés d'une forme d'ubiquité. Chaque Etat membre devient responsable de la mise en œuvre sur son territoire d'une règle commune. Cela permet l'action coordonnée des Etats membres et de leurs autorités de régulation. Cela permet aussi la mise en place de mécanismes d'atténuation du risque politique de chaque Etat membre au niveau européen par divers mécanismes éprouvés : revue par les pairs, mise en place d'un conseil des régulateurs, surveillance directe de la commission européenne et contrôle juridictionnel de la CJUE.

Ces architectures ont déjà été testées et mise en œuvre notamment dans le cas du règlement européen sur la neutralité de l'internet.

En l'absence d'un tel mécanisme et compte tenu des enjeux, les états risquent de se lancer unilatéralement dans des initiatives législatives divergentes en termes de périmètres et d'obligations, au détriment de la logique de marché unique numérique et au bénéfice des acteurs déjà en place qui seraient les seuls capables de soutenir le poids économique d'un empilement des régulations nationales (en renforçant la logique du « *winner takes all* »).

La pertinence de la mise en place d'un dispositif français de régulation des réseaux sociaux doit être analysée au vu de sa capacité à servir, sous la nouvelle mandature européenne, de proposition permettant :

- d'inverser la dynamique européenne actuelle pour sortir de la logique du pays d'installation qui peut fragiliser la souveraineté des États membres et affaiblir leur capacité à affronter la globalisation, et basculer dans une logique du pays de destination afin de renforcer les États membres ;
- de réduire le risque politique ;
- et d'accroître la légitimité de la construction européenne.

Contrairement à l'approche répressive, la mise en place d'un cadre de régulation *ex ante*, suivant une approche de conformité, centrée sur la création d'incitations fortes à la responsabilisation du réseau social et de ses membres ainsi qu'une crédibilisation de l'autorégulation, constitue une opportunité unique de proposer cette réorientation de la dynamique européenne.

2.5 Une politique de régulation reposant sur une approche de conformité à appliquer et concevoir avec pragmatisme et agilité

Dans le secteur financier, les États ont voulu crédibiliser et inscrire dans la durée l'engagement des établissements financiers à contribuer activement à l'atteinte d'objectifs d'intérêt général que constituent la lutte contre le blanchiment d'argent, contre le trafic de stupéfiants ou le financement du terrorisme. Dans cette logique, les superviseurs bancaires ont donc consacré leurs efforts à imposer et à surveiller des obligations de moyens, c'est-à-dire la conformité à certaines règles préventives, et non à la sanction des échecs lorsque les risques combattus se matérialisent, sans préjudice de l'action pénale dans cette circonstance. Ainsi le superviseur bancaire n'intervient pas lorsqu'il constate qu'un établissement financier a servi de canal pour véhiculer un financement utilisé pour des finalités illégales, mais lorsqu'il constate qu'un établissement financier n'a pas mis en œuvre une action de prévention prescrite, et ce indépendamment du fait de savoir si un comportement illégal implique ou non cet établissement financier. Ce mode d'intervention est de nature à créer des incitations ciblées pour que les plateformes participent à l'accomplissement d'un objectif d'intérêt général sans avoir une action normative directe sur le service offert.

Appliqué aux services de réseau social, ce mode d'intervention suppose d'identifier les obligations génériques de nature à créer les bonnes incitations, notamment en accroissant l'efficacité du dialogue politique avec les institutions politiques des États membres, et leurs sociétés civiles. Ceci suppose d'imposer des obligations fortes sur la transparence des fonctions clés non-observables de l'extérieur, telles que la fonction de modération (et des modalités d'élaboration et de mise à jour des conditions générales d'utilisation qui la sous-tendent) ainsi que du recours aux algorithmes de ciblage et d'individualisation des contenus présentés.

Cette approche doit se mettre en place progressivement et pragmatiquement selon la taille des acteurs et de leurs services :

- Seuls les services les plus structurants par leur taille, et donc potentiellement les plus dangereux par leur impact massif en cas d'abus d'usage, devraient être soumis à ces obligations et à un contrôle *ex ante* de conformité du régulateur. Le régulateur devrait ainsi consacrer sa mission principale à cette mission de supervision active des acteurs systémiques.

- Les services de taille intermédiaire devraient se voir reconnaître une présomption de conformité en première intention, bénéficier d'un accompagnement du régulateur qui pourrait les encourager dans leur engagement de responsabilité par l'édition de recommandations, par une action visant à accroître la mutualisation des actifs communs (base de données d'identification de contenus illicites, accès à des jeux de données annotées pour l'apprentissage automatique des algorithmes de modération par exemple) avec les plus grandes plateformes dans une démarche ouverte, transparente et ayant pour effet d'abaisser les barrières à l'entrée. Néanmoins, si le régulateur constate qu'un service de taille intermédiaire est en défaut patent vis-à-vis de ces obligations de moyens et que cela provoque une matérialisation excessive des effets nocifs combattus par la politique publique, celui-ci doit pouvoir demander à l'acteur fournissant ce service de prendre les mesures adéquates et, à défaut de mise en œuvre par la plateforme, pouvoir, par une décision motivée, lui imposer une démarche de conformité opposable *ex ante* à l'image de celle imposée aux services les plus structurants.
- Enfin, concernant les plus petites plateformes, le régulateur ne doit pouvoir agir que par le dialogue et l'édition de recommandation mais sans action coercitive de son fait, sans préjudice de sa capacité à saisir le procureur de tout fait qui pourrait relever d'une qualification pénale.

Les mécanismes de responsabilisation au titre de la responsabilité pénale et civile, notamment au titre de la LCEN, demeurent néanmoins applicables pour l'ensemble des services de réseau social, quelle que soit leur taille. Elles ne relèvent pas de l'action du régulateur mais des instances et mécanismes judiciaires de droit commun.

Le cadre législatif devrait permettre une mise en œuvre graduelle de ces mécanismes et reconnaître à la nouvelle fonction de régulation une capacité à définir et affiner progressivement les obligations imposées dans une démarche agile afin de pouvoir s'adapter rapidement à la mutation des réseaux sociaux. Autrement dit, la loi instaurant ce cadre de régulation doit définir la nature des obligations, sans chercher à en détailler les modalités. A défaut, le schéma de régulation serait aisément contourné par des acteurs « trop » agiles.

III - Structuration de la fonction de régulation en France dans un cadre défini au niveau européen

Il apparaît nécessaire de compléter l'action répressive contre les auteurs publiant des contenus illégaux ou cherchant à manipuler les réseaux sociaux, par une démarche proactive reposant sur le renforcement du dialogue politique entre les pouvoirs publics et les acteurs concernés. Créer les conditions d'un dialogue constructif et régulier avec les plateformes de réseau social doit les faire basculer du côté des solutions en les incitant à adopter un comportement responsable vis-à-vis de leurs utilisateurs et de la société, et à prévenir les abus sur leurs services.

Cette politique de régulation pourrait reposer sur les 5 piliers suivants :

Premier pilier

Une politique publique de régulation aux objectifs larges garante des libertés individuelles et de la liberté d'entreprendre

Pour fédérer les énergies politiques tant au niveau national qu'au niveau européen, et rassembler les institutions politiques et la société civile, ce schéma de régulation doit avoir pour objectifs de défendre l'exercice des droits et des libertés sur les plateformes de réseaux sociaux :

- la liberté d'expression et de communication des utilisateurs, qui sont à ce titre en droit de savoir si la plateforme respecte cette liberté d'expression ;
- le droit des utilisateurs au respect de leur intégrité physique et morale, y compris sur les réseaux sociaux et dans l'espace numérique ;
- la liberté d'entreprendre des réseaux sociaux, y compris le droit à édicter et appliquer des CGU, à exercer une fonction d'ordonnancement libre de l'information, et à innover (en particulier pour les plus petits acteurs).

D'autres objectifs pourraient également être pris en compte :

- le pluralisme des services de réseaux sociaux qui implique notamment de veiller à ne pas créer de barrières à l'entrée *via* une politique de régulation trop contraignante pour des nouveaux acteurs ;
- la cohésion sociale en encourageant les réseaux sociaux à développer des usages « positifs » de leurs services, i.e. contribuant au renforcement du lien social.

Deuxième pilier

Une régulation prescriptive ciblée sur la responsabilisation des réseaux sociaux et mise en œuvre par une autorité administrative indépendante.

- **Une régulation prescriptive limitée aux plateformes structurantes au niveau de chaque État membre avec deux seuils de déclenchement :**
 - La régulation s'appliquerait d'office pour les services dont le nombre mensuel d'utilisateurs dépasse un certain pourcentage de la population de l'État membre (entre 10 % et 20 %) ;
 - Elle s'appliquerait uniquement sur décision motivée du régulateur en cas de défaillance constatée et persistante pour les services dont le nombre mensuel d'utilisateurs se situerait entre 0 % et 5 % de la population de l'État membre. Il convient de noter que plus ce second seuil d'application est bas plus le test d'impact devra être strict et exigeant pour respecter un principe général de proportionnalité¹⁶ ;
 - En deçà, la régulation n'est pas applicable mais le droit commun de la LCEN demeure et permet de mettre en cause la responsabilité civile et pénale des acteurs en cas de manquements.

- **Des obligations de transparence qui porteraient sur les fonctions clés des réseaux sociaux¹⁷ :**
 - **la « fonction d'ordonnancement des contenus**, c'est-à-dire les modalités de présentation, de hiérarchisation et de ciblage des contenus publiés par les utilisateurs, y compris lorsqu'ils sont promus par la plateforme ou par un tiers en contrepartie d'une rémunération ;
 - **la fonction de mise en œuvre des CGU et de modération des contenus, incluant les** modalités d'élaboration de ces règles communautaires, les procédures, les moyens humains et technologiques mis en œuvre pour assurer le respect de ces CGU et pour lutter contre les contenus illicites. Cette fonction doit être auditable par le régulateur et/ou par un auditeur indépendant de la plateforme. La transparence se traduirait par exemple par :
 - l'obligation pour le service de notifier sa décision à l'auteur du contenu modéré (sauf exceptions légitimes : besoins des autorités publique de poursuite) et l'auteur du signalement (s'il existe) ; mécanisme indépendant et extra-judiciaire de réexamen de la décision de la plateforme (sans préjudice d'un recours juridictionnel) ;
 - des informations sur le recours aux outils de traitement automatisés : quels outils sont utilisés, pour quels types de contenus, avec quelle supervision humaine ? Comment leur efficacité et leur précision est-elle évaluée ?
 - des éléments sur les modalités de coopération avec des « partenaires de confiance » (« trusted flaggers ») : liste, modalités de sélection, « privilèges » attachés à ce statut, données statistiques sur le nombre de signalement examinés, le nombre de contenus détectés proactivement, les suites données (retrait, maintien ou autre), les recours traités, etc.

¹⁶ Ce deuxième seuil pourrait être remplacé par un critère de « malfaisance » afin de doter le régulateur de la capacité de se saisir de tous les réseaux sociaux posant problèmes. Néanmoins, il conviendra de se poser la question de l'économie des moyens du régulateur et de la nécessité de créer un critère de dessaisissement du régulateur vis-à-vis de petits acteurs ne posant plus problème.

¹⁷ Le détail des obligations de transparence présenté ci-dessous est donné à titre illustratif ; il n'a pas nécessairement à figurer de manière exhaustive dans le texte de loi qui organiserait le dispositif de régulation et peut être, pour une large part, renvoyé au pouvoir réglementaire du gouvernement ou du régulateur.

- des statistiques sur l'efficacité de la modération : délais de décision (quelle qu'elle soit), taux de faux positif / négatif, viralité / audience des contenus contraires aux standards avant qu'ils soient retirés (cf. notion de prévalence), etc.

- **Un devoir de diligence du réseau social vis-à-vis de ses utilisateurs**

Par cette obligation qui s'approche du « *duty of care* » anglo-saxon, les réseaux sociaux s'engageraient à assumer une responsabilité vis-à-vis de leurs utilisateurs concernant des abus d'autres membres et des tentatives de manipulation de la plateforme par des tiers.

Une telle obligation de moyen permettrait une intervention de la puissance publique s'il apparaîtrait que la démarche, volontaire à ce jour, des plateformes de veiller à ce que leurs usagers puissent se sentir en confiance, via la création de fonction dites de « *Trust and safety* » (« confiance et sécurité ») ou la fonction de modération manquaient de moyens.

Troisième pilier

Un dialogue politique, transparent, informé, au champ large entre le gouvernement, le régulateur, les acteurs et la société civile.

- Le gouvernement, *via* son pouvoir réglementaire, fixerait les seuils de déclenchement des obligations et définirait les modalités des obligations de transparence des fonctions d'ordonnancement des contenus et de mise en œuvre des CGU ainsi que l'obligation de défendre l'intégrité du réseau social et de ses membres.
- Le gouvernement homologuerait les décisions à caractère réglementaire du régulateur.
- Le gouvernement organiserait le dialogue politique avec les réseaux sociaux en y associant le régulateur et la société civile.
- La portée et l'effectivité du dialogue politique serait enrichie par l'action ciblée du régulateur visant à responsabiliser des réseaux sociaux. Le gouvernement serait alors en mesure de poursuivre son action *via* le dialogue politique sur l'ensemble des questions de société avec les réseaux sociaux en y associant la société civile (monde associatif, territoires, communauté éducative, monde académique)¹⁸.
- Les administrations centrales accompagneraient le gouvernement dans son dialogue politique en tirant partie de la réduction de l'asymétrie d'information entre plateformes et institutions politiques du fait de la régulation prescriptive.
- Les plateformes prendraient, le cas échéant, des engagements volontaires auprès du gouvernement, contrôlables et opposables devant le régulateur (mise en œuvre d'un plan d'actions face à une dérive nouvellement identifiée, amélioration des métriques de transparence pour l'année à venir, etc).

¹⁸ Le contenu des CGU relève de la liberté d'entreprendre des réseaux sociaux mais la transparence apportée par le régulateur permet de faire des CGU l'objet d'un dialogue politique. Facebook reconnaît par exemple avoir modifié la portée de son règlement communautaire pour mieux protéger les réfugiés des discours haineux sous la pression politique.

Quatrième pilier

Une autorité administrative indépendante partenaire des autres branches de l'Etat, et ouverte sur la société civile.

- Dans le schéma de régulation proposé, l'autorité administrative indépendante serait garante de la responsabilisation des réseaux sociaux au bénéfice du gouvernement et de la société civile.
- Elle mettrait en œuvre la régulation coercitive de manière autonome sans être autarcique ni hégémonique. Elle serait le régulateur de la responsabilisation des principaux réseaux sociaux *via* le contrôle des obligations de transparence des fonctions d'ordonnancement et de modération des contenus, et de devoir de diligence leur incombant. Elle ne serait ni le régulateur des réseaux sociaux dans leur globalité, ni le régulateur des contenus qui y sont publiés. Elle ne serait pas compétente pour qualifier les contenus pris individuellement. Elle coopérerait avec les services de l'État placés sous l'autorité du gouvernement et les services judiciaires.
- Elle n'imposerait pas directement de contrainte sur la définition des services de réseaux sociaux, mais elle leur imposerait la transmission et la publication d'informations, dont elle qualifierait la véracité et la pertinence, en associant la société civile (monde associatif, communauté éducative, monde académique). Elle contrôlerait l'effectivité des moyens déployés pour mettre en œuvre le devoir de diligence du réseau social vis-à-vis de ses membres.
- Elle devrait être dotée de larges pouvoirs d'accès à l'information détenue par les plateformes, *via* la possibilité d'utiliser des identités d'emprunt et d'imposer des accès privilégiés aux algorithmes pour vérifier la sincérité de la description publiée par le réseau social¹⁹. Elle ne pourrait se voir opposer le secret des affaires ni la protection des données personnelles, sans préjudice de son obligation de protéger les données qu'elle requiert conformément au RGPD et aux lois sur les secrets commerciaux.
- Elle devrait également être dotée d'un pouvoir de sanction administrative permettant d'imposer (i) des mesures de publicité sur le réseau social à destination de ces utilisateurs incluant ses partenaires commerciaux (notamment les annonceurs), et (ii) des sanctions pécuniaires pouvant atteindre [4 %]²⁰ du chiffre d'affaire mondial de l'acteur qui fournit le réseau social. Ces pouvoirs de sanction ne pourraient être exercés qu'après mise en demeure publique.
- Elle serait dotée du mandat et des compétences juridiques pour mettre en place des passerelles pour permettre des travaux de recherches académiques sur les plateformes à partir de leurs données, dans le respect du RGPD.
- Elle inciterait à la mutualisation des moyens et des savoirs au profit des réseaux sociaux de moindre envergure, contribuant ainsi à l'abaissement des barrières à l'entrée.
- Elle participerait activement au réseau des régulateurs européens et soutiendrait l'action du gouvernement dans la négociation de la politique européenne.

¹⁹ *Via* la mise en place d'accès direct, en temps réel, ciblés et proportionnés, aux systèmes d'information du réseau social via des interfaces (API) dédiées.

²⁰ La mission n'a pas mené de travail spécifique sur le niveau des sanctions pécuniaires à prévoir et reprend les montants prévus par le RGPD.

Cinquième pilier

Une articulation européenne qui renforce la capacité des Etats membres à agir face à des plateformes globales, et réduit les risques de mise en œuvre dans chaque Etat membre

Face à la puissance des plateformes globales, la mission propose que l'Union Européenne puisse organiser la capacité des gouvernements et de leurs sociétés civiles à travailler en réseau en conjuguant leurs forces. Cette architecture renforce les Etats membres dans leur rôle de garants de la cohésion et du lien social dans un monde globalisé.

Cette articulation au niveau européenne devrait reposer sur :

- un règlement européen afin de reconnaître le caractère global des plateformes numériques, de rendre pleinement effective une action coordonnée des autorités nationales face à des acteurs globaux (notamment via des procédures communes, des APIs communes, etc.), de réduire les risques de mise en œuvre dans chaque État membre ;
- une mise en œuvre nationale suivant la règle du pays de destination²¹ pour rendre les plateformes responsables localement devant chaque Etat membre et dans les géographies où elles peuvent créer des dommages ;
- des actions concertées entre autorités nationales et ouvertes à la société civile permettant de rendre plus efficace le contrôle de la transparence des plateformes ;
- des mécanismes européens de réduction des risques de régulation excessive par un Etat membre (« check and balance »), corollaire indispensable de la compétence concurrente de chaque État membre : consultation publique nationale et européenne pour les décisions à caractère réglementaire ou les recommandations, mécanisme de saisine pour avis du régulateur national des données personnelles, coordination et cohérence des décisions des régulateurs nationaux par un organe collégial rassemblant les régulateurs nationaux et la Commission européenne.

²¹ La compétence concurrente des États membres devrait être limitée par des prérequis : (i) seuil exprimé en nombre moyen d'utilisateur mensuel en pourcentage de la population au niveau national pour établir la compétence automatique du régulateur, (ii) seuil plus bas lorsqu'il est couplé au constat d'un dommage manifeste dans le pays de destination, et (iii) pouvoir et montant éventuelle de la sanction proportionnés aux conséquences potentielles au niveau de l'État-membre.

IV – Focus sur la transparence des algorithmes

Des algorithmes de plus en plus complexes

Les utilisateurs des réseaux sociaux sont tous les jours confrontés aux algorithmes : classement des contenus sur le fil d'actualité, insertion de contenus sponsorisés, algorithme de modération des contenus contraires aux CGU, suggestion d'amis, etc.

Lorsque l'on traite de la transparence de ces algorithmes, il faut prendre la définition du mot algorithme au sens large, c'est-à-dire l'ensemble de règles opératoires dont l'application permet de résoudre un problème énoncé au moyen d'un nombre fini d'opérations, et y englober les données proposées en entrée de l'algorithme, celles qui ont pu être utilisées auparavant pour « entraîner » l'algorithme, les données de contexte, etc. De nombreux algorithmes fonctionnent selon des approches statistiques et fournissent donc des réponses probabilistes (par exemple : probabilité qu'un utilisateur apprécie un contenu ou qu'il clique sur un produit. Certains utilisent des techniques d'apprentissage automatique qui consistent à essayer d'imiter des choix humains, par exemple, les choix de modérateurs d'accepter ou rejeter des contenus regroupés dans des corpus de contenus annotés. Ils sont également pour la plupart fortement individualisés. Enfin, les algorithmes sont aussi très évolutifs dans le temps avec des mises à jour qui peuvent être quotidiennes. Tous ces éléments entraînent un changement de paradigme pour l'intelligibilité des algorithmes. Chaque algorithme peut avoir des milliards d'avatars qui ne se comportent pas tout à fait de manière identique selon les utilisateurs et selon les pays. Néanmoins, il est nécessaire pour l'action publique d'en extraire les principes généraux.

Une nécessité de transparence algorithmique

Les algorithmes peuvent être mal employés ou détournés de leur utilité initiale. L'importance qu'ils ont prise sur les plateformes de réseaux sociaux et les dérives qu'ils peuvent causer (mise en avant de propos haineux, modération non effective, ingérence d'un Etat souverain dans le débat public, etc.) ont rendu l'intervention de l'Etat indispensable. Cette intervention passe par la transparence, c'est-à-dire le moyen de rendre intelligible la logique sous-jacente, les grands principes de traitement opérés par les algorithmes. Ce premier niveau, peu interventionniste du point de vue des pouvoirs publics, permet une relative levée de voile sur les boîtes noires que représentent certains algorithmes. Il permet aussi de pointer d'éventuels biais de fonctionnement, dus à des choix (conscients ou inconscients des développeurs, aux programmes informatiques, ou à des biais provenant des données). Surtout, il permet d'alimenter le débat public sur les questions de société que pose l'utilisation massive des algorithmes.

Concrètement, la transparence algorithmique peut prendre différentes formes. Par exemple, pour le citoyen sans compétence technique particulière, la transparence peut se traduire par la publication des critères déterminants qui ont conduit à un résultat le concernant (classement d'information, recommandation, envoi de publicité ciblée, etc.) ou la justification d'une décision particulière (modération d'un contenu posté, absence de réponse suite à un signalement). Un acteur plus expert sera davantage intéressé par des mesures de performance (tels que les taux de faux positif ou de faux négatif dans la modération), ou par des explications relatives à l'architecture des traitements (sous forme d'arbres de décision ou d'autres représentations graphiques mettant en lumière les données prises en compte par l'algorithme et leur influence sur les résultats). Le monde académique sera intéressé par la publication de jeux de données de référence, pour comprendre les choix de modération faits par les plateformes, et sans lesquels il est impossible de reproduire les résultats d'un algorithme d'apprentissage.

La transparence ne peut pas être purement déclarative. Vérifier la loyauté des algorithmes utilisés par les entreprises est une tâche particulièrement complexe. Le régulateur devra avoir les moyens de le faire en s'appuyant sur des mesures statistiques, la mise à disposition d'API de test, des mécanismes de certification par des tiers ou de *compliance*.

Une action publique en deux niveaux

La transparence ne peut être effective que si elle résulte d'un dialogue régulier avec les opérateurs et d'un processus de tâtonnement qui imite celui que suit le développement de ces algorithmes. Dès lors, le législateur doit établir les grands principes de l'action des pouvoirs publics pour imposer la transparence des algorithmes tout en laissant une marge d'appréciation au vu des circonstances de l'espèce au régulateur en charge d'appliquer la loi. En ce sens, le législateur devra définir les principes que le pouvoir réglementaire déclinera en s'adaptant aux contextes particuliers. En pratique, le régulateur devra procéder aux mises en balance et trouver le juste équilibre entre le principe de transparence et la protection du secret des affaires, ou encore fixer des obligations générales d'intelligibilité des algorithmes visés. En revanche, la loi ne doit pas définir de métrique particulière ni imposer des processus spécifiques de mise en œuvre. A titre d'exemple, l'article 14 de la loi relative à la lutte contre la manipulation de l'information n'est pas compris par l'écosystème qui le trouvent soit trop soit pas assez spécifique (que se passe-t-il si l'algorithme est personnalisé ? que se passe-t-il si l'algorithme est mis à jour ? que signifie "part d'accès direct" dans le contexte d'un algorithme de classement ?).

La loi ayant donné la capacité au régulateur d'extraire de l'information, celui-ci pourra alors étudier au cas par cas les spécificités de chaque algorithme pour donner corps à la transparence. C'est cette démarche - par définition expérimentale, de co-construction et de tâtonnement - qui permettra au régulateur de développer ses instruments de politique publique au niveau réglementaire et de manière incrémentale. Dans un second temps, cela pourrait permettre de faire émerger d'éventuelles synergies pour définir des principes plus larges régissant la transparence des algorithmes.

Au-delà des enjeux techniques et juridiques autour des algorithmes et de leur transparence, le but de la régulation sera de porter les questions d'éthique et les choix moraux et politiques que soulèvent les algorithmes dans le débat public, de faire apparaître clairement la « politique algorithmique ». Ainsi, la connaissance accumulée du régulateur et les données qu'il fera remonter devront alimenter au mieux les débats sociétaux et académiques en répondant aux questions sur l'ensemble des aspects sous-jacents au concept de transparence, par exemple sur :

- Les données :
 - Quelles sont les données utilisées pour entraîner les algorithmes ? comment sont-elles collectées ? sont-elles personnelles ?
 - Quelles sont les données servant de paramètres d'entrée de l'algorithme ?
 - Ces données présentent-elles des biais ?
- Le modèle :
 - Quel est le flot de traitement suivi ? Quelles briques algorithmiques sont employées ?
 - Quels mécanismes de supervision/surveillance de l'apprentissage des algorithmes sont utilisés ?
 - Quelle individualisation est réalisée ?
 - L'algorithme reproduit-il des biais ? Peut-il être utilisé à mauvais escient ? Quelles sont des dérives possibles à éviter ?
- Les inférences :
 - Quelles sont les taux de faux positifs et de faux négatifs ?
 - Quelles métriques peuvent rendre compte de la performance de l'algorithme ?
 - Quel intervalle de confiance peut-on accorder au résultat ?
 - Quelles procédures sont utilisées pour corriger les erreurs ?

Annexe 1

Lettre de mission



MINISTÈRE DE L'ÉCONOMIE
ET DES FINANCES

SECRETARIAT D'ÉTAT
CHARGÉ DU NUMERIQUE

MINISTÈRE DE L'ACTION
ET DES COMPTES PUBLICS

Monsieur Frédéric Potier,
Préfet en mission de service public,
DILCRAH

Monsieur Serge Abiteboul,
Chercheur Inria
Membre du collège de l'Arcep

A Paris, le 11 MARS 2019

Objet : Lettre de mission - mission Facebook

Messieurs,

Les réseaux sociaux occupent une place désormais essentielle dans notre société en offrant à ses utilisateurs des espaces et des outils performants d'échanges d'idées, de débats, et de partage de contenus. Ils constituent à ce titre une formidable opportunité pour l'exercice de la liberté d'expression et de la liberté de communication, fondements de notre société démocratique.

Facteurs de progrès et espace de libertés, les réseaux sociaux sont néanmoins aussi le lieu de diffusion de contenus illicites qui peuvent être vus par un nombre potentiellement très élevé de citoyens. La liberté d'expression et la liberté de communication, comme sur les autres médias, doivent être conciliées avec d'autres principes qui peuvent en limiter l'exercice, notamment le respect de la dignité de la personne humaine.

Les lois de la République s'appliquent évidemment dans l'espace numérique, tant vis-à-vis des utilisateurs que des plateformes. Toutefois, le volume des contenus, la vitesse à laquelle ils se propagent sur les réseaux sociaux et leur impact sur la société justifient la mise en place complémentaire d'une régulation systémique des fonctions de modération sur les réseaux sociaux.

Développés par les réseaux sociaux à leur seule initiative, les dispositifs de modération existants relèvent de pratiques d'autorégulation qui, bien que pertinentes à certains égards, n'offrent pas de garantie suffisante à l'exercice des droits fondamentaux de nos concitoyens.

139 rue de Bercy – 75572 Paris Cedex 12

Ces initiatives privées ne peuvent plus faire l'économie d'un accompagnement par la puissance publique. C'est avec cette conviction que j'ai souhaité qu'une mission de réflexion et de propositions portant sur les dispositifs de modération des contenus sur les réseaux sociaux vous soit confiée. Les enseignements de cette mission viendront compléter les conclusions du rapport de la mission confiée par le Gouvernement à Madame Laetitia Avia, Monsieur Karim Amellal et Monsieur Gil Taieb¹.

C'est dans ce contexte que la société Facebook a accepté, dans une démarche expérimentale, de travailler avec votre mission d'expertise et de vous présenter l'état de sa fonction de modération et de ses perspectives de développement, avec une attention particulière concernant la lutte contre la diffusion de contenus incitant à la haine.

Sur la base de cette collaboration inédite avec un acteur privé, vous réaliserez une évaluation de ces mécanismes d'autorégulation mis en œuvre par Facebook. Vous étudierez en particulier les traitements algorithmiques développés et utilisés par Facebook à cet effet.

Vous proposerez des recommandations dont vous aurez préalablement analysé les risques et opportunités, et en particulier les précautions à envisager pour construire un cadre de régulation national, qui puisse se décliner à une échelle plus vaste, notamment européenne, compte tenu du caractère global des réseaux sociaux. Cette expérimentation doit être perçue comme une première étape permettant de réfléchir très concrètement aux meilleures voies pour garantir que l'ensemble des réseaux sociaux, et non pas seulement Facebook, appliquent un haut niveau d'exigence et de qualité dans la modération des contenus qu'ils hébergent.

Pour mener cette mission, vous consulterez l'ensemble des parties prenantes et tiendrez compte des éventuelles initiatives existantes dans d'autres pays.

Vous vous appuyerez sur une équipe d'experts dont la composition figure en annexe et sur trois rapporteurs mis à disposition pendant toute la durée de la mission. Avec votre accord, M. Benoit Loutrel en sera le rapporteur général.

Vous rendrez compte périodiquement de l'avancement de vos travaux à un comité de pilotage réunissant les cabinets des ministres concernés et présidé par mon représentant.

Les enseignements tirés de cette expérimentation serviront à nourrir les travaux réglementaires nationaux et européens sur ce sujet. Je souhaite ainsi pouvoir disposer de votre rapport au plus tard le 30 juin 2019.



Mounir MAHJoubi

¹ Rapport remis au Premier ministre le 20 septembre 2018 et intitulé « Renforcer la lutte contre le racisme et l'antisémitisme sur Internet ».

Annexe 2

Les membres de la mission

Serge Abiteboul

Directeur de recherche en informatique à l'Inria et à l'École Normale Supérieure de Paris, membre du Collège de l'Arcep, Serge Abiteboul est un spécialiste en bases de données, d'information et de connaissances. Il a été professeur (Collège de France, Stanford, Oxford University, etc.), membre du CNNum, et startuper. Il est membre de l'Académie des Sciences. Il est également blogueur et écrivain.

Frédéric Potier

Après un parcours au sein du Ministère de l'Intérieur et en cabinet ministériel, Frédéric Potier a été nommé en mai 2017 préfet en mission de service public et délégué interministériel à la lutte contre le racisme, l'antisémitisme et la haine anti-LGBT (DILCRAH). A ce poste, il pilote les politiques publiques de l'Etat de lutte contre certaines discriminations et les actes et les discours de haine.

Côme Berbain

Ingénieur du corps des Mines, docteur en cryptographie, Côme Berbain est le directeur des technologies numériques de l'État. Son parcours alterne entre entités privées (Orange, Trusted Logic) et publiques (ministère de la défense, ANSSI) dans les domaines de la transformation numérique et de la cybersécurité. Il a été conseiller au cabinet du secrétaire d'Etat chargé du numérique en 2017 et 2018.

Jean-Baptiste Gourdin

Diplômé de l'Institut d'études politiques de Paris et de l'ENA, Jean-Baptiste Gourdin est chef de service, adjoint au directeur général à la Direction générale des médias et des industries culturelles (DGMIC). Magistrat de la Cour des comptes, il a notamment été directeur de cabinet du Président du CSA et coordinateur de la mission « Acte 2 de l'exception culturelle ».

Jacques Martinon

Magistrat judiciaire, Jacques Martinon a débuté sa carrière en tant que juge d'instruction. Depuis 2016, il a rejoint la direction des affaires criminelles et des grâces (DACG) et est chef de la mission de lutte contre la cybercriminalité au Ministère de la Justice. Contributeur à la revue stratégique de cyberdéfense du SGDSN, il est intervenu comme formateur à la session nationale commune INHESJ/IHEDN « souveraineté numérique et cybersécurité ».

Gilles Schwoerer

Gilles Schwoerer est officier de gendarmerie et a travaillé dans une grande variété de secteurs régaliens (Armée de Terre, Gendarmerie départementale, gendarmeries spécialisées en lien avec l'aéronautique) avant de rejoindre le centre de lutte contre les criminalités numériques (C3N) au sein du Pôle judiciaire de la Gendarmerie en tant qu'adjoint au chef du C3N.

Aude Signourel

Après 7 ans au Ministère de la Justice et 3 ans au cabinet du préfet de la Seine-Saint-Denis, elle a rejoint la Direction des libertés publiques et des affaires juridiques au Ministère de l'Intérieur. Depuis 2017, elle est conseiller juridique de la sous-direction de la lutte contre la cybercriminalité de la direction centrale de la police judiciaire à laquelle appartient la plateforme PHAROS.

Les rapporteurs de la mission

Sacha Desmaris

Juriste diplômée de l'Université Paris 2 Panthéon-Assas, Sacha Desmaris est cheffe du département Economie de l'audiovisuel à la Direction des études, des affaires économiques et de la prospective du Conseil supérieur de l'audiovisuel dans laquelle elle a été chargée de mission de 2016 à 2019, après avoir travaillé quatre ans au secrétariat général du Groupe M6.

Pierre Dubreuil

Ingénieur diplômé de Telecom ParisTech et de l'Ecole Normale Supérieure Paris-Saclay, spécialiste des questions d'apprentissage automatique et co-fondateur d'une startup, Pierre Dubreuil est chargé de mission au sein de l'Autorité de régulation des télécommunications et des postes (Arcep).

Benoît Loutrel

Inspecteur général de l'INSEE, spécialisé en économie industrielle et de la régulation, Benoit Loutrel a été directeur général de l'Autorité de régulation des communications et des postes (Arcep) de 2013 à 2016, après y avoir travaillé de 2004 à 2010. Il a été directeur du programme « numérique » des investissements d'avenir de 2010 à 2013. Il a été directeur des affaires publiques de Google France durant quelques mois en 2017.

