

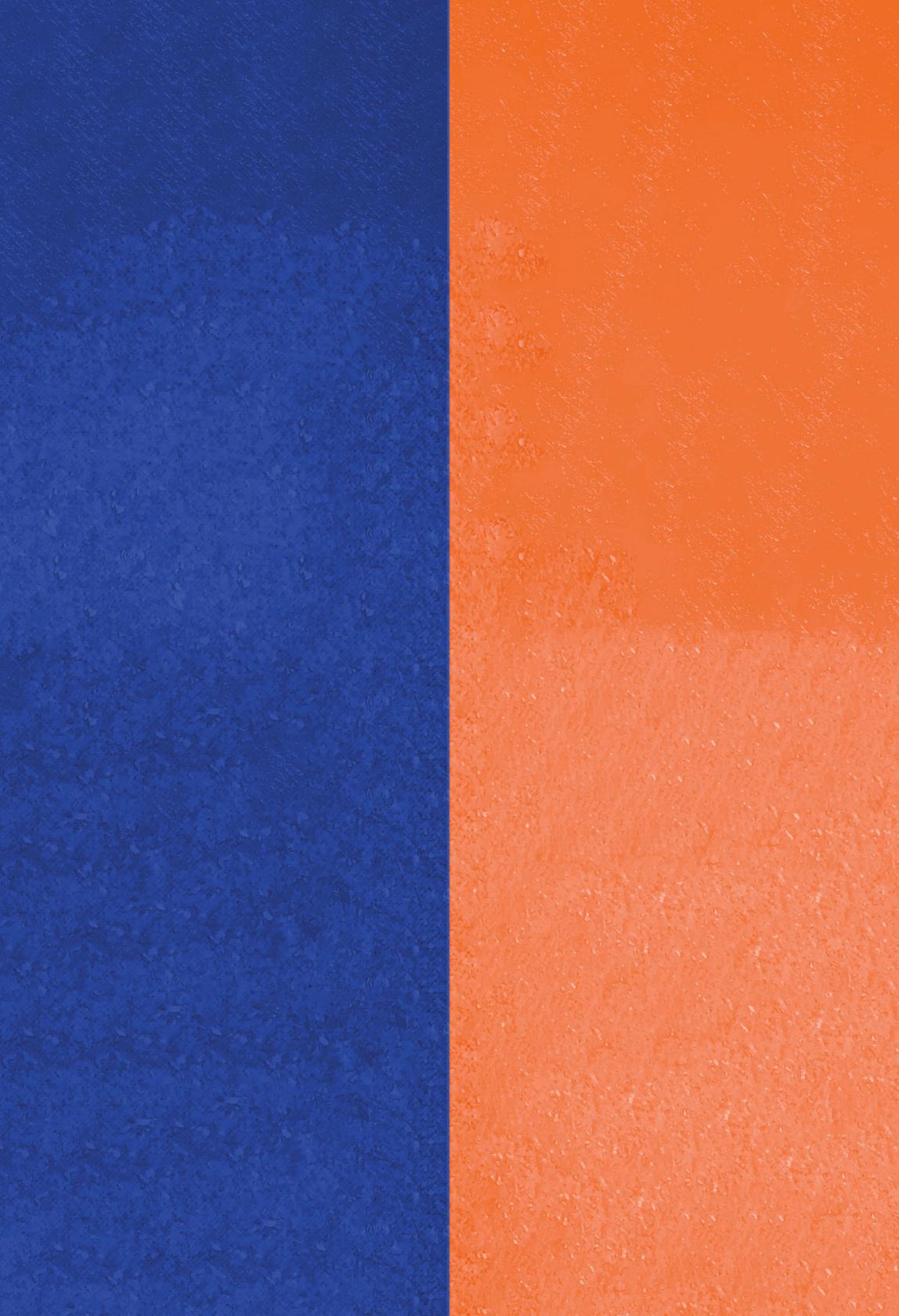
ARTICLE 19



Séries de normes internationales

Surveiller les gardiens

Modération de contenu,
gouvernance et liberté
d'expression



Séries de normes internationales

Surveiller les gardiens

Modération de contenu, gouvernance et liberté d'expression



ARTICLE 19 œuvre pour un monde où chacun, où qu'il soit, puisse s'exprimer librement et s'engager activement dans la vie publique sans crainte de discrimination. Pour ce faire, nous travaillons sur deux libertés imbriquées, qui constituent le fondement de tous nos travaux. La Liberté de parler concerne le droit de tous à s'exprimer et diffuser des opinions, des idées et des informations par tous les moyens, et d'être en désaccord et d'interroger les détenteurs du pouvoir. La liberté de savoir concerne le droit d'exiger et de recevoir des informations par des détenteurs du pouvoir pour garantir la transparence, la bonne gouvernance et le développement durable. Lorsque l'une ou l'autre de ces libertés est menacée par l'incapacité des détenteurs du pouvoir à les protéger de manière adéquate, ARTICLE 19 parle d'une seule voix, par le biais des tribunaux, des organisations internationales et régionales, et de la société civile partout où nous sommes présents.

✉ info@article19.org

🌐 www.article19.org

🐦 [@article19org](https://twitter.com/article19org)

📧 [@article19](https://www.instagram.com/article19)

🌐 www.linkedin.com/company/article19/

📘 facebook.com/article19org

978-1-910793-48-0

©ARTICLE 19, 2023

Ce document est fourni sous la licence 3.0 Creative Commons Attribution-Non-Commercial-ShareAlike.

Vous êtes autorisés à copier, diffuser et afficher ce document et à créer des produits dérivés à condition de :

1. créditer ARTICLE 19 ;
2. ne pas utiliser ce document à des fins commerciales ;
3. diffuser tous les travaux dérivés de cette publication sous une licence identique à celle-ci.

Pour accéder au texte juridique intégral de cette licence, se rendre sur : <http://creativecommons.org/licenses/by-nc-sa/3.0/legalcode>

ARTICLE 19 vous serait reconnaissant de bien vouloir lui adresser une copie de tout contenu reprenant des informations de ce rapport. ARTICLE 19 assume l'entière responsabilité du contenu de ce document.

Toutes les illustrations :
Mariana Coan
www.marianacoan.com

Resume exécutif

Dans cette note d'orientation, ARTICLE 19 présente sa position sur la réglementation de la modération du contenu par les plateformes de médias sociaux en vue de protéger le droit à la liberté d'expression et d'information.

Cette note s'appuie sur nos précédents travaux dans ce domaine, en particulier nos rapports sur la responsabilité des intermédiaires et sur les directives communautaires/conditions de service des entreprises. Nos propositions ont reposé jusque-là sur une réglementation que l'on pourrait qualifier de légère. Nous avons soutenu que les plateformes de médias sociaux devaient continuer de bénéficier d'une exonération de responsabilité pour le contenu de leurs utilisateurs. Cette règle importante n'empêche pas les entreprises d'être tenues responsables de la non-suppression de contenus illégaux. Dans le même temps, ce modèle repose sur le fait que les entreprises appliquent des conditions de service conformes aux normes internationales des droits humains. Nous avons défendu ce modèle parce qu'il protège mieux le droit à la liberté d'expression des utilisateurs : il est important non seulement pour les plateformes de médias sociaux, mais aussi pour un large éventail de prestataires de services d'infrastructure sur Internet, tels que les réseaux de diffusion de contenu et les bureaux d'enregistrement des noms de domaine.

Aujourd'hui, ce paradigme ne semble plus viable face à l'ampleur des grandes plateformes de médias sociaux et à leur incapacité constante à répondre aux critiques qui leur sont adressées – de la gestion par Facebook de la crise des Rohingyas au Myanmar à la position incompréhensible de YouTube sur le « discours de haine » et aux attaques contre les femmes journalistes sur Twitter. Nous croyons également que la transparence devrait être une exigence fondamentale dans toutes les activités des entreprises, et qu'elle doit être accompagnée d'une plus grande responsabilité et d'un engagement en faveur de la protection des droits humains.

Dans cette note d'orientation, ARTICLE 19 examine si le modèle que nous préconisons tient la route face aux critiques adressées aux grandes plateformes de médias sociaux. Nous examinons les avantages et les inconvénients des diverses propositions formulées récemment dans le monde. Enfin, nous présentons nos nouvelles positions sur la réglementation des plateformes.

Recommandations clés

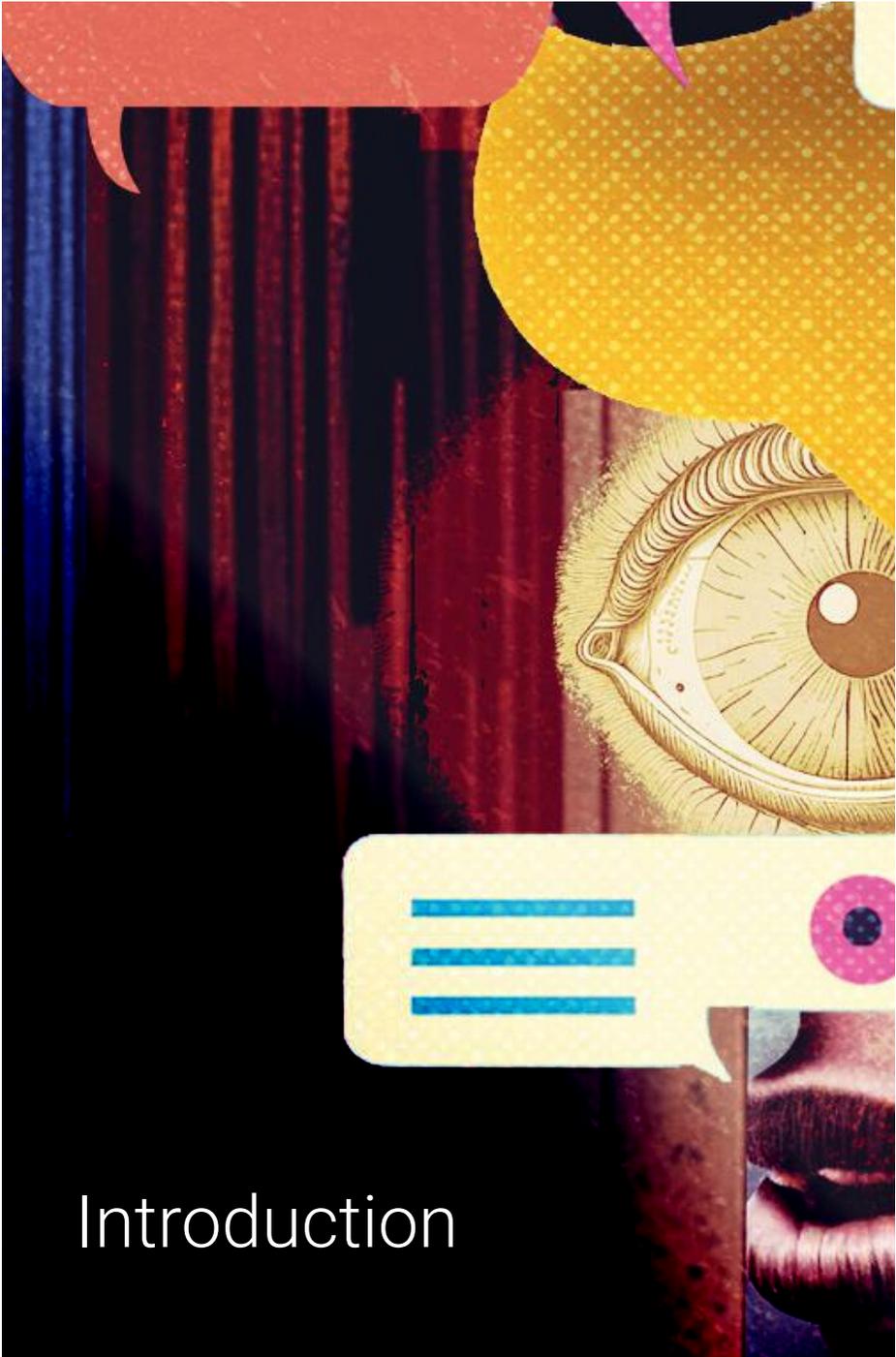
1. Les États devraient s'abstenir d'adopter une réglementation inutile de la modération de contenu en ligne.
2. La transparence, la redevabilité et la protection des droits humains doivent être les principes fondamentaux de tout cadre réglementaire.
3. L'exonération conditionnelle de responsabilité pour le contenu de tiers doit être maintenue, mais sa portée et ses procédures de notification et d'action doivent être clarifiées.
4. La surveillance générale des contenus doit rester interdite.

5. Tout cadre réglementaire doit avoir une portée strictement limitée. La réglementation devrait se concentrer sur le contenu illégal plutôt que sur le « contenu légal, mais préjudiciable ». Les services de messagerie privée et les agences de presse devraient être exclus du champ d'application. Les mesures ne devraient pas avoir d'application extraterritoriale.
6. Les obligations prévues dans tout régime réglementaire doivent être clairement définies. Elles comprennent, en particulier, des obligations de transparence et des obligations internes de procédure régulière.
7. Tout régulateur doit être indépendant, tant en droit que dans la pratique.
8. Tout cadre réglementaire doit être proportionné.
9. Tout cadre réglementaire doit donner accès à des recours effectifs.
10. Les grandes plateformes devraient être contraintes de dégroupier leurs fonctions d'hébergement et de curation de contenu et de garantir leur interopérabilité avec d'autres services.

Table des matières

Introduction	11
Concepts principaux	17
Normes internationales des droits humains applicables	22
Garanties du droit à la liberté d'expression	22
Restrictions du droit à la liberté d'expression	22
Entreprises de médias sociaux et liberté d'expression	23
La protection du droit à la vie privée et à l'anonymat en ligne	28
Dilemmes sur la réglementation de la modération de contenu	31
Arguments en faveur de la réglementation des plateformes	31
Position d'ARTICLE 19 sur les arguments en faveur d'une réglementation des plateformes	35
Concepts et caractéristiques clés de la réglementation de la modération de contenu	45
Caractéristiques positives des propositions législatives sur la réglementation de la modération de contenu	48
Lacunes des propositions réglementaires sur la modération de contenu	49

Recommandations D'ARTICLE 19	57
Recommandation 1	57
Recommandation 2	58
Recommandation 3	59
Recommandation 4	65
Recommandation 5	67
Recommandation 6	69
Recommandation 7	76
Recommandation 8	78
Recommandation 9	80
Recommandation 10	80
Notes	83



Introduction

Introduction

En 1996, John Perry Barlow, l'un des fondateurs d'Internet, publia la Déclaration d'indépendance du cyberspace.¹ Sa vision était celle d'un monde où « chacun, où qu'il soit, peut exprimer ce qu'il croit, quel que soit le degré de singularité de ses croyances, sans devoir craindre d'être forcé de se taire ou de se conformer ». Les gouvernements n'y étaient pas les bienvenus. La gouvernance devait être établie sur une base « éthique », et sur « l'intérêt bien compris de chacun et de la communauté dans son ensemble ». Près de 25 ans plus tard, Internet est vu comme un espace rassemblant un très petit nombre de plateformes de médias sociaux (en particulier Facebook, Google, et Twitter), la censure y sévit, et les gouvernements sont en passe de le réglementer. Que s'est-il passé entre-temps ?

La perception du rôle des plateformes de médias sociaux a également évolué. À leurs débuts, ces plateformes passaient pour une force puissante agissant pour le bien, favorisant la liberté d'expression, facilitant des liens entre les individus et constituant le fer de lance d'une révolution démocratique dans le monde entier. Mais aujourd'hui, elles sont perçues comme un foyer de « discours de haine », de harcèlement, d'intimidations, de théories du complot et de propagande. Elles ne se contentent plus « d'héberger »² simplement des contenus dans une relative neutralité, mais promeuvent activement certains contenus de tiers, voire même produisent leurs propres contenus. Cela s'explique par l'incapacité des principales plateformes de médias sociaux à comprendre les préoccupations de leurs utilisateurs et gouvernements et à y répondre, à commencer par le scandale de Cambridge Analytica, ou l'incapacité de Facebook à supprimer des incitations au génocide des Rohingyas au Myanmar ou encore le combat de YouTube pour supprimer la vidéo de l'attaque terroriste de Christchurch en Nouvelle-Zélande.

Le pouvoir que ces entreprises exercent sur les particuliers et leur position dominante sur plusieurs marchés³ ont été mis en évidence en janvier 2021, lorsque les principales entreprises de médias sociaux – dont Twitter, Facebook et bien d'autres – ont suspendu les comptes de l'ancien président américain Donald Trump, dont les tweets incitant à prendre d'assaut le Capitole au début du mois menaçaient de déclencher des violences.⁴ Toujours en janvier, Twitter a verrouillé le compte officiel de l'ambassade de Chine aux États-Unis après la parution d'un message de soutien au pouvoir chinois dans la région occidentale du Xinjiang, où la Chine est engagée dans la stérilisation forcée des femmes ouïghoures.⁵ Ce n'était pas la première fois, mais il s'agissait d'une remarquable démonstration de force des plateformes de médias sociaux, qui a relancé l'un des débats les plus féroces dans les cercles de la modération de contenu, à savoir qui doit décider de ce qui est ou n'est pas autorisé dans le discours public en ligne.

Il est essentiel que ce débat ait lieu, et les États ont un rôle à jouer. En réalité, une grande partie de la faute a été imputée aux cadres réglementaires de responsabilisation des intermédiaires d'Internet. Dans de nombreux pays, les médias sociaux et autres plateformes numériques ont bénéficié d'une exonération, large ou conditionnelle, de responsabilité pour l'hébergement de contenus illégaux.⁶ Ces dernières années, ces cadres ont été vivement critiqués parce qu'ils donnaient carte blanche à ces entreprises, en les aidant à augmenter leurs profits avec des algorithmes favorisant une participation addictive autour de contenus « extrémistes » et autres contenus « préjudiciables ». La question s'est alors posée de savoir si une réglementation plus poussée était nécessaire pour limiter le pouvoir des médias sociaux dominants, lutter contre les contenus illégaux et autres contenus néfastes et prévoir une plus grande redevabilité démocratique. Plus récemment, les gouvernements ont également proposé de placer les plateformes de médias sociaux sous la tutelle d'organismes réglementaires de type audiovisuel ; ou d'imposer aux plateformes un « devoir de diligence » envers leurs utilisateurs afin d'éviter les « préjudices » causés par le discours d'autres internautes.⁷

Pour les défenseurs de la liberté d'expression, ces propositions soulèvent des questions difficiles, à savoir quels sont les acteurs suffisamment dignes de confiance pour surveiller et contrôler l'expression des utilisateurs. Si l'on considérait autrefois que les plateformes de médias sociaux offraient un haut niveau de protection de la liberté d'expression, ce n'est plus le cas aujourd'hui, car elles ont fini par restreindre progressivement leurs normes communautaires, faisant souvent taire les voix des minorités.⁸ La transparence et le règlement des différends liés à la suppression de contenus n'ont pas permis jusque-là un examen minutieux de leurs actions et n'ont pas fourni de réparation significative à leurs utilisateurs. Enfin, il est suspect qu'un petit nombre de plateformes dominantes soient autorisées à détenir autant de pouvoir sur ce que les gens peuvent voir sans rendre des comptes directement.

Dans le même temps, la perspective d'une nouvelle réglementation des « plateformes » pose de sérieux problèmes. Un grand nombre des propositions actuelles concernent en réalité la réglementation des « contenus » en ligne, c'est-à-dire de l'expression des utilisateurs, les États exigeant effectivement que les entreprises surveillent elles-mêmes les communications humaines et décident quel contenu est « illégal » ou « préjudiciable ».⁹ Cela est profondément problématique car seuls les tribunaux peuvent déterminer l'illégalité d'un contenu, et différents types de contenu pourraient bien nécessiter différents types de réglementation. En effet, les solutions conçues pour lutter contre les contenus pédopornographiques peuvent être inappropriées contre la désinformation ou pour protéger le droit d'auteur.

En tout état de cause, ARTICLE 19 estime que les nouvelles propositions de régulation des plateformes doivent être examinées sous le prisme de la liberté d'expression. À cette fin, les États doivent identifier la méthode la plus adéquate et la moins contraignante pour protéger efficacement chacun des objectifs traditionnellement assignés à la régulation des médias (tels que le pluralisme et la diversité de la liberté d'expression), en tenant compte de l'évolution et des rôles des plateformes numériques dans la promotion et la protection des droits humains – y compris la liberté d'expression – en ligne.

Cette note d'orientation propose ce genre d'analyse à la lumière des normes internationales relatives à la liberté d'expression. Elle s'appuie sur notre position actuelle sur la responsabilité des intermédiaires,¹⁰ sur notre politique relative aux règles internes/directives communautaires des plateformes numériques et à la liberté d'expression,¹¹ notre proposition de nouveau modèle réglementaire multipartite – les Conseils de médias sociaux¹² – et notre travail de plaidoyer auprès des plateformes numériques. Tout en évaluant si nos positions sur ces questions restent viables, nous formulons des recommandations sur les garanties minimales à inclure dans tout cadre régissant les activités des plateformes de médias sociaux.

Cette note d'orientation porte principalement sur les plateformes de médias sociaux « dominantes », que nous appelons parfois « plateformes numériques ». Cependant, nous sommes très attentifs au contexte plus large de notre examen, et nous nous référons à d'autres acteurs de l'écosystème plus large d'Internet, notamment les fournisseurs d'infrastructure (comme les réseaux de diffusion de contenu) et d'autres services numériques (tels que la messagerie privée ou les services du cloud).

Cette note d'orientation comprend quatre parties :

- La première définit certains termes clés, notamment les plateformes de médias sociaux « dominantes », la « modération de contenu » et « l'autorégulation » ;
- La deuxième présente les normes applicables de la protection de la liberté d'expression en ligne qui devraient guider tous les efforts législatifs et politiques dans ce domaine ;
- La troisième expose les principaux arguments en faveur d'une plus grande réglementation des politiques de modération de contenu des plateformes et notre réponse globale à celles-ci ;¹³ et
- Enfin, ARTICLE 19 formule des recommandations sur les garanties fondamentales à inclure dans tout nouveau cadre réglementaire pour protéger les droits humains.

Cette note d'orientation complète d'autres rapports d'ARTICLE 19 sur la réglementation des plateformes, notamment notre proposition de « dégroupage » de la curation de contenu,¹⁴ sur la diversité des médias dans l'écosystème numérique,¹⁵ et sur les obligations de diffusion (must-carry obligations).¹⁶ Nous prévoyons également de publier d'autres positions politiques sur d'autres questions liées à ce sujet complexe.¹⁷



Concepts principaux

Concepts principaux

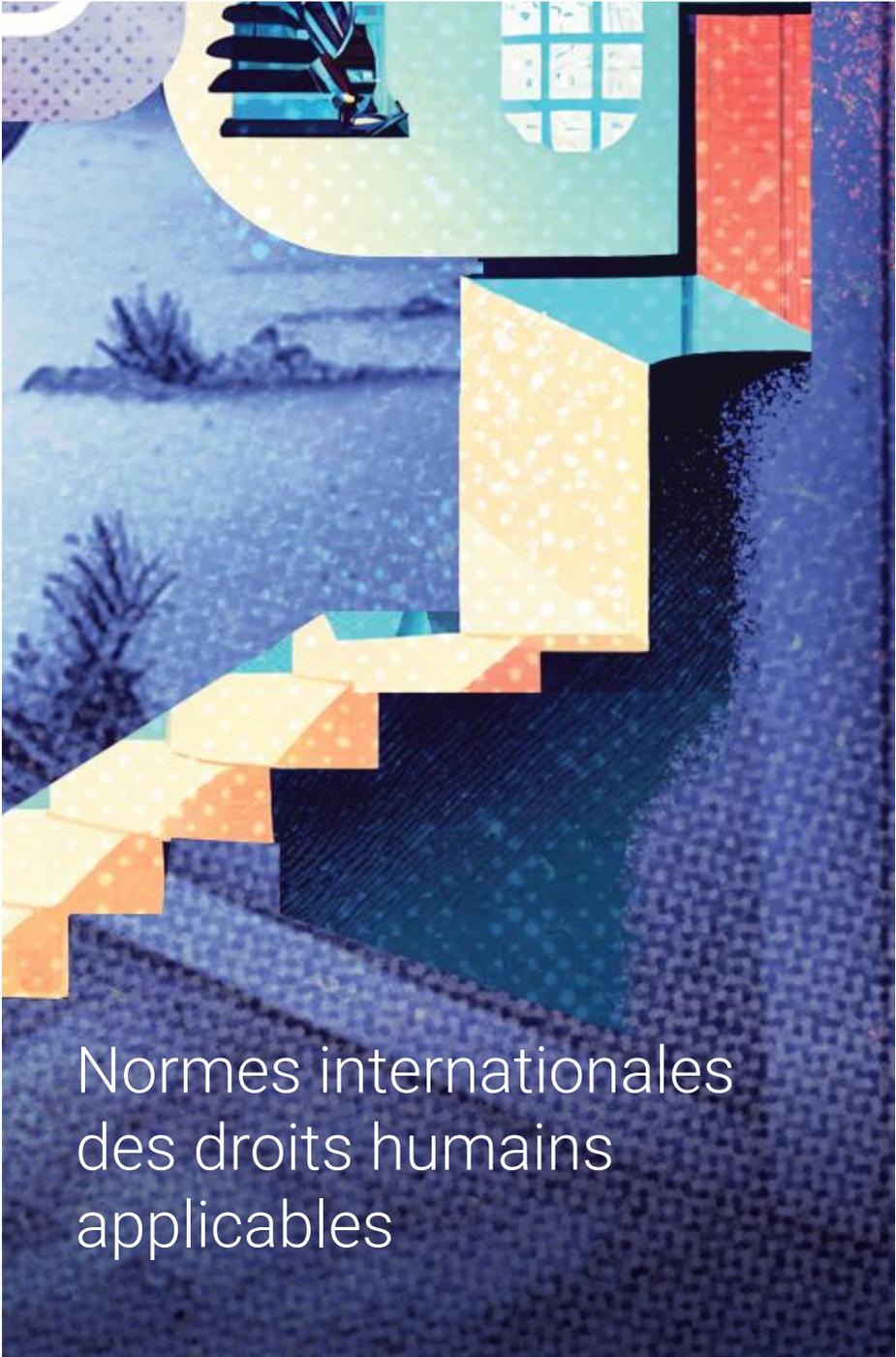
ARTICLE 19 a déjà établi une typologie des intermédiaires d'Internet – fournisseurs de services Internet, fournisseurs d'hébergement, plateformes de médias sociaux et moteurs de recherche.¹⁸ Cette typologie reste toujours pertinente, mais les pratiques des plateformes de médias sociaux ayant évolué, et leurs activités faisant l'objet d'un examen plus attentif, il est important de définir certains concepts clés dans ce débat. Aux fins de cette note d'orientation :

- **Les plateformes de médias sociaux** sont des entreprises qui permettent à des individus et des groupes de se connecter et d'échanger avec d'autres internautes, et de partager des contenus via des réseaux de communication numériques. Elles fournissent des services « d'hébergement » aux utilisateurs, qui vont souvent de pair avec d'autres services tels que la messagerie privée, la « modération de contenu » et la curation des fils d'actualité. De plus, les plateformes de médias sociaux offrent habituellement des espaces publicitaires et des services associés aux annonceurs ainsi que des services de référencement de l'information aux éditeurs d'actualités, mettant en relation ces derniers avec les utilisateurs. Elles peuvent également fournir des interfaces de programmation d'application (API) aux développeurs d'applications, pour programmer et développer des outils qui s'intègrent à la plateforme. De ce fait, certains spécialistes ont provisoirement identifié quatre marchés pertinents sur lesquels les plateformes de médias sociaux ont l'habitude d'opérer : les services de réseautage social, l'affichage publicitaire, le référencement des actualités et les plateformes pour applications.¹⁹

- **Le terme de plateforme numérique** est souvent utilisé ici pour désigner les « plateformes de médias sociaux ». Toutefois, les plateformes numériques couvrent un ensemble d'activités plus large que les médias sociaux. Elles peuvent comprendre des places de marché en ligne, des magasins d'applications, des systèmes de paiement, des moteurs de recherche et des plateformes de l'économie collaborative. Les plateformes numériques ont les caractéristiques suivantes, entre autres : (1) elles reposent sur un modèle commercial largement fondé sur la collecte, le traitement et l'édition de grandes quantités de données ; (2) elles opèrent sur des marchés multifaces ; (3) elles bénéficient des effets de réseau, où la valeur du service tend à augmenter avec le nombre de ses utilisateurs ; et (4) elles reposent sur des technologies de l'information et de la communication pour atteindre leurs utilisateurs instantanément et sans frais.²⁰
- **La modération de contenu** comprend les différents ensembles de mesures et outils utilisés par les plateformes de médias sociaux pour lutter contre les contenus illégaux et garantir le respect de leurs normes communautaires dans le contenu généré par les utilisateurs de leurs services. Cela implique généralement des signalements par les utilisateurs, des signaleurs ou « filtres » de confiance, la suppression de contenu, l'étiquetage, la rétrogradation ou la démonétisation du contenu, ou la désactivation de certaines fonctionnalités.
- **La curation de contenu** sous-entend l'utilisation par les plateformes de médias sociaux de systèmes automatisés pour classer, promouvoir ou rétrograder les contenus dans les fils d'actualité, généralement en fonction des profils des usagers. Le contenu peut être également promu sur des plateformes en échange d'un paiement. Les plateformes peuvent aussi organiser le contenu en utilisant des interstitiels pour prévenir les utilisateurs contre un contenu sensible ou en appliquant certaines étiquettes pour avertir, par exemple, que le contenu est de source fiable ou non.

- **L'autorégulation** est un cadre reposant entièrement sur le respect spontané des règles ; la législation ne joue aucun rôle dans l'application des normes pertinentes. Sa raison d'être est de tenir les membres responsables devant le public, de promouvoir les connaissances parmi ses membres, et d'élaborer et respecter des normes éthiques. Les modèles d'autorégulation reposent d'abord et avant tout sur une compréhension commune par les membres des valeurs et de l'éthique qui sous-tendent leur conduite professionnelle. Dans le contexte des entreprises de médias sociaux, le terme « autorégulation » a été utilisé pour désigner la « régulation en solo » (la réglementation de l'expression par le biais des conditions de service ou des normes communautaires) et la réglementation sectorielle (gamme d'initiatives adoptées par les plateformes de médias sociaux pour gérer certains problèmes, par ex. la « désinformation »). Le terme d'autorégulation est parfois utilisé dans cette note pour décrire aussi les initiatives gouvernementales qui contraignent les entreprises à adopter une série de mesures sous la supervision du gouvernement, mais où le non-respect de ces mesures n'entraîne aucune sanction juridique.²¹
- **La responsabilité des intermédiaires**²² fait référence aux lois selon lesquelles les intermédiaires d'Internet peuvent être tenus légalement responsables du contenu diffusé ou créé par leurs usagers. Dans de nombreux pays, les intermédiaires d'Internet sont exempts de responsabilité (c'est-à-dire non passibles de poursuites) tant qu'ils suppriment le contenu après avoir effectivement pris connaissance de l'illégalité, ou lorsqu'un tribunal leur ordonne de le faire. L'exonération de responsabilité est généralement considérée comme essentielle à la survie des intermédiaires d'Internet, compte tenu de l'ampleur des contenus produits par les utilisateurs.

- **La responsabilité des éditeurs²³** correspond à la responsabilité légale d'une personne pour le contenu qu'elle publie. En pratique, cela signifie que les éditeurs sont soumis aux lois d'application générale, par ex. la loi sur la diffamation, les lois interdisant la publication de secrets défense, etc. Les journaux en constituent un exemple parfait. Ils sont légalement responsables du contenu produit sur leur site ou sous forme imprimée. Cela comprend les lettres aux éditeurs. Dans la pratique, le contenu publié par les journaux, que ce soit en ligne ou hors ligne, est soumis à une révision approfondie et un examen juridique minutieux avant publication. Aujourd'hui, tout le monde peut être éditeur grâce à Internet. Les internautes peuvent donc être également tenus responsables des contenus qu'ils produisent.



Normes internationales des droits humains applicables

Normes internationales des droits humains applicables

Garanties du droit à la liberté d'expression

Le droit à la liberté d'expression est protégé par l'Article 19 de la Déclaration universelle des droits de l'homme (DUDH).²⁴ Il a acquis une force légale dans l'Article 19 du Pacte international relatif aux droits civils et politiques (PIDCP)²⁵ et dans les traités régionaux.²⁶

La portée du droit à la liberté d'expression est large et s'applique à toutes les formes d'expression électronique ou basées sur Internet. Elle contraint les États à garantir à toute personne la liberté de rechercher, de recevoir ou de répandre des informations ou des idées de toute sorte, sans considération de frontières, par tout moyen de son choix. En vertu des normes internationales des droits humains, le cadre juridique réglementant les médias de masse devrait tenir compte des différences entre la presse écrite et audiovisuelle et Internet,²⁷ tandis que la réglementation des secteurs des télécommunications et de l'audiovisuel ne peut être simplement transférée sur Internet.²⁸ Les États devraient adopter une approche adéquate contre les contenus illégaux en ligne et promouvoir l'autorégulation en tant qu'outil efficace contre le discours préjudiciable en ligne.²⁹

Restrictions du droit à la liberté d'expression

En vertu des normes internationales des droits humains, les États peuvent exceptionnellement limiter le droit à la liberté d'expression à condition que ces restrictions soient conformes aux obligations prévues dans le test en trois parties. Les restrictions doivent donc :

- **Être prévues par la loi** : toute loi ou réglementation doit être formulée avec suffisamment de précision pour permettre aux individus d'adapter leur conduite en conséquence ;
- **Poursuivre un but légitime** : figurant parmi les objectifs énumérés de manière exhaustive, comme le respect des droits ou de la réputation d'autrui, la protection de la sécurité nationale ou de l'ordre public, ou de la santé ou moralité publiques ; et
- **Être nécessaires et proportionnées dans une société démocratique** : si une mesure moins intrusive est susceptible d'atteindre le même objectif qu'une mesure plus restrictive, c'est la mesure la moins restrictive qui devra être appliquée.³⁰

De plus, l'Article 20(2) du PIDCP stipule que toute promotion de la haine nationale, raciale ou religieuse qui constitue une incitation à la discrimination, à l'hostilité ou à la violence doit être interdite par la loi. Les mêmes principes s'appliquent aux formes électroniques de communication ou d'expression diffusées sur Internet.³¹

Entreprises de médias sociaux et liberté d'expression

Les organes internationaux des droits de l'homme ont fourni des commentaires sur les relations entre la liberté d'expression et les entreprises de médias sociaux dans plusieurs domaines.

La responsabilité des intermédiaires

Les mandataires spéciaux sur la liberté d'expression ont longtemps affirmé que l'exemption de responsabilité était le moyen le plus efficace de protéger la liberté d'expression en ligne. Ils ont recommandé que les intermédiaires ne soient pas tenus responsables des contenus produits par d'autres lorsqu'ils ne fournissent que des services techniques, et que la responsabilité soit invoquée uniquement dans le cas où l'intermédiaire est spécifiquement intervenu dans le contenu publié en ligne.³²

Les Rapporteurs spéciaux successifs des Nations Unies sur la promotion et la protection du droit à la liberté d'opinion et d'expression (Rapporteurs spéciaux sur la liberté d'expression) ont maintes fois déclaré que la censure ne devait jamais être déléguée à une entité privée, les États ne devant pas utiliser des intermédiaires ou les forcer à entreprendre une censure en leur nom.³³ Par ailleurs, David Kaye a affirmé qu'« une réglementation judicieusement conçue plutôt qu'une réglementation lourde fondée sur un point de vue devrait être la norme. Cette réglementation devrait avoir pour objectif de garantir la transparence de l'entreprise et prévoir des mesures de réparation... ».³⁴ Sans exclure la possibilité d'une réglementation sous certaines conditions, le Rapporteur a rappelé que les États ne devaient chercher à restreindre le contenu qu'en vertu d'une ordonnance d'une autorité judiciaire indépendante et impartiale, et conformément aux garanties de procédure régulière et aux normes de légalité, nécessité et légitimité.³⁵ Il a également noté que les États devaient s'abstenir d'imposer des sanctions disproportionnées – amendes lourdes ou peines d'emprisonnement – aux intermédiaires d'Internet, étant donné leur effet dissuasif important sur la liberté d'expression.³⁶ Son successeur et actuelle Rapporteuse spéciale sur la liberté Irene Khan est parvenue à une conclusion similaire.³⁷

Responsabilité du secteur privé en matière de droits humains

Les recommandations sur les responsabilités des entreprises de médias sociaux en matière de respect des droits humains incluent (mais sans s'y limiter) les instruments suivants :

- **Les Principes directeurs des Nations Unies relatifs aux entreprises et aux droits de l'homme : mise en œuvre du cadre de référence « Protéger, respecter et réparer » des Nations Unies** (Principes directeurs) fournissent une base pour définir le rôle du secteur privé dans la protection des droits humains sur

Internet.³⁸ Ils reconnaissent l'obligation des entreprises de respecter des droits humains, indépendamment des obligations de l'État ou de la mise en œuvre de ces obligations, et recommandent aux entreprises d'adopter plusieurs mesures.³⁹ Il s'agit notamment d'intégrer d'office des garanties des droits humains pour atténuer les impacts négatifs de leurs activités, de créer un effet de levier et d'agir collectivement pour renforcer leur pouvoir vis-à-vis des autorités gouvernementales, et de proposer des recours en cas d'impacts négatifs sur les droits humains.

- **Les mandataires spéciaux pour la liberté d'expression** se sont intéressés au rôle des plateformes de médias sociaux dans la promotion de la liberté d'expression et leur ont recommandé de respecter et promouvoir les Principes directeurs à cet égard. Ils ont recommandé, entre autres, que les entreprises définissent des conditions d'utilisation claires et sans ambiguïté conformes aux normes et principes internationaux des droits humains,⁴⁰ qu'elles produisent des rapports sur la transparence,⁴¹ et offrent des recours efficaces aux utilisateurs affectés en cas de violations.⁴² Ils ont également appelé à plusieurs reprises les États à ne pas contraindre le secteur privé à prendre des mesures qui entravent inutilement ou de manière disproportionnée la liberté d'expression, que ce soit par le biais des lois, de politiques ou de moyens extra-légaux, les entreprises n'étant généralement pas armées pour déterminer l'illégalité d'un contenu.⁴³ Le Rapporteur spécial sur la liberté d'expression a également précisé que les entreprises devaient adopter des approches de la transparence radicalement différentes à toutes les étapes de leurs activités et s'ouvrir à la responsabilisation publique, suggérant que cela pourrait prendre la forme de Conseils de médias sociaux.⁴⁴

- Au niveau régional :
 - Le Commissaire aux droits de l'homme du Conseil de l'Europe a recommandé que les États cessent de compter sur les entreprises d'Internet pour imposer des restrictions contraires à leurs obligations en matière de droits humains,⁴⁵ et que des orientations supplémentaires soient élaborées sur les responsabilités des entreprises commerciales en relation avec leurs activités sur Internet.⁴⁶ De même, le Comité des ministres du Conseil de l'Europe a recommandé que les entreprises de médias sociaux respectent les droits humains et l'état de droit, y compris les garanties procédurales,⁴⁷ et qu'elles soient transparentes dans leur utilisation des techniques de traitement automatisé des données.⁴⁸
 - L'Organisation des États américains (OEA) a publié plusieurs rapports sur la liberté d'Internet où elle a formulé des recommandations similaires.⁴⁹ Par exemple, l'OEA a souligné que les intermédiaires étaient toujours des entités privées, dont les intérêts diffèrent de ceux de l'État, et qu'il n'est pas possible de les obliger à fonctionner comme un tribunal, car cela dépasse le cadre de leur compétence et peut inciter à des abus.⁵⁰
 - La Commission africaine des droits de l'homme et des peuples a récemment adopté la Déclaration de principes révisée sur la liberté d'expression et l'accès à l'information en Afrique.⁵¹ Le Principe 39 de la Déclaration stipule que les États doivent contraindre les intermédiaires d'Internet à veiller à ce que, en modérant et filtrant le contenu en ligne, ils intègrent des garanties des droits humains dans leur processus, adoptent des stratégies d'atténuation pour traiter toutes les restrictions de la liberté d'expression et de l'accès à l'information en ligne, garantissent la transparence dans toutes les demandes de suppression de contenu, incorporent des mécanismes d'appel et offrent des recours efficaces en cas de violation des droits.⁵²

- **Les organisations de la société civile** ont également recommandé que les entreprises de médias sociaux respectent les normes internationales des droits humains. Par exemple, les Principes de Manille sur la responsabilité des intermédiaires précisent les types de mesures qui doivent être prises par les entreprises dans ce domaine.⁵³ En particulier, les pratiques de restriction de contenu doivent respecter les tests de nécessité et de proportionnalité prévus par la législation des droits humains⁵⁴ et fournir aux internautes des mécanismes de gestion des plaintes pour contester les décisions des entreprises.⁵⁵ Le projet de classement des droits numériques (Ranking Digital Rights) évalue la conformité des principales sociétés d'Internet aux indicateurs des droits numériques, dont, entre autres la disponibilité des conditions de service ; les raisons de la restriction du contenu, du compte ou du service ; la notification aux utilisateurs ; un processus pour répondre aux requêtes de tiers ; et diverses activités relatives à la transparence.⁵⁶

Principes spécifiques au contenu

Les rapporteurs spéciaux sur les droits de l'homme ont publié un certain nombre de déclarations conjointes soulignant les responsabilités des États et des entreprises par rapport à des contenus spécifiques, dont « l'extrémisme violent »,⁵⁷ « les fausses nouvelles »,⁵⁸ le harcèlement et les abus en ligne basés sur le genre,⁵⁹ et le « discours de haine ». ⁶⁰ Dans ces déclarations, ils ont appelé les États à ne pas soumettre les plateformes numériques à des injonctions de suppression ou de restriction de contenu, sauf lorsque le contenu est légalement restreint en vertu des normes internationales.⁶¹ Ils ont également appelé les entreprises à s'assurer que leurs utilisateurs peuvent facilement accéder à, et comprendre, leurs politiques et pratiques, les modalités de leur application, et comment elles respectent les directives minimales de procédure régulière ;⁶² et à veiller à ce que ces politiques soient directement liées au droit international relatif aux droits humains.⁶³

Le rôle de l'intelligence artificielle dans la modération de contenu

En août 2018, dans un rapport à l'Assemblée générale des Nations Unies, le Rapporteur spécial sur la liberté d'expression a formulé un certain nombre de recommandations sur le rôle de l'intelligence artificielle (IA) dans la modération de contenu et sur son impact potentiel sur les droits humains.⁶⁴ En particulier, il a recommandé que les États créent un cadre politique et législatif propice à un environnement diversifié et pluraliste de l'information dans le domaine de l'intelligence artificielle.⁶⁵ De telles mesures pourraient inclure la réglementation des monopoles technologiques pour éviter une concentration de l'expertise et du pouvoir de l'IA entre les mains de quelques entreprises dominantes, et une réglementation conçue pour accroître l'interopérabilité des services et des technologies.⁶⁶ Il a recommandé aux entreprises, entre autres, d'indiquer explicitement où et comment les technologies de l'IA sont utilisées sur leurs plateformes, leurs services et applications ;⁶⁷ de publier des données sur les suppressions de contenu, des études de cas, et d'éduquer au profilage commercial et politique ;⁶⁸ et de fournir aux internautes un accès à des recours en cas d'atteinte aux droits humains par les systèmes d'IA.⁶⁹

La protection du droit à la vie privée et à l'anonymat en ligne

Il est essentiel de protéger le droit à la vie privée dans les communications en ligne afin de s'assurer que les individus ont la confiance nécessaire pour exercer librement leur droit à la liberté d'expression.⁷⁰ L'impossibilité de communiquer en toute confidentialité affecte considérablement les droits à la liberté d'expression des individus.

Cela a été reconnu dans plusieurs rapports du Rapporteur spécial sur la liberté d'expression David Kaye, qui s'est inquiété du fait que des États et des acteurs privés surveillent et collectent des informations sur les communications et les activités des internautes. Ces pratiques peuvent constituer une violation du droit à la vie privée des internautes, et au final empêcher la libre circulation de l'information et des idées en ligne.⁷¹

Le Rapporteur spécial sur la liberté d'expression a également recommandé que les États veillent à ce que les individus puissent préserver l'anonymat en ligne et qu'ils s'abstiennent d'adopter des systèmes d'enregistrement sous de vrais noms.⁷² En outre, il a recommandé que les États s'abstiennent d'imposer une identification des usagers comme condition préalable à l'accès aux communications numériques et aux services en ligne, et d'exiger l'enregistrement d'une carte SIM pour les utilisateurs de téléphones mobiles.⁷³ Il a également recommandé que les acteurs commerciaux reconsidèrent leurs propres politiques de restriction du chiffrement et de l'anonymat (y compris par l'usage de pseudonymes).⁷⁴



Dilemmes sur la réglementation de la modération de contenu

Dilemmes sur la réglementation de la modération de contenu

Comme indiqué précédemment, les États se sont mobilisés ces dernières années pour réglementer les activités de modération de contenu des plateformes numériques « dominantes » sur le marché des médias sociaux. Dans la pratique, ces appels à une réglementation des « plateformes » se traduisent souvent par une réglementation du « contenu en ligne », même si la différence est souvent floue dans le discours public.

Dans ce chapitre, ARTICLE 19 examine quelques-unes des raisons majeures invoquées pour une intervention des régulateurs – et, en particulier, pour l’adoption d’un cadre global de réglementation du contenu en ligne. Nous examinerons ensuite si ces motifs résistent à un examen minutieux.

Arguments en faveur de la réglementation des plateformes

Pouvoir des plateformes et position dominante sur le marché

Dans l’ensemble, le principal moteur de la dynamique actuelle en faveur de la réglementation est la position dominante d’un petit nombre de plateformes de médias sociaux sur le marché, qui leur permet de contrôler la manière dont un grand nombre de personnes exercent leur droit à la liberté d’expression en ligne. La taille et les revenus de ces plateformes sont souvent supérieurs à ceux de plusieurs pays.⁷⁵ Pour la liberté d’expression et la démocratie elle-même, la plus grande préoccupation porte sur le fait que seule une poignée d’entreprises mondiales – principalement basées aux États-Unis – décident quelles informations peuvent être vues, mises à disposition et partagées par les utilisateurs.

Dans la pratique, le problème de la concentration de pouvoir a donné lieu à un certain nombre de propositions qui placeraient les plateformes de médias sociaux dominantes sous l'autorité d'un régulateur de type audiovisuel. Ces propositions portent donc essentiellement sur la réglementation du contenu en ligne ; elles ne traitent pas directement du pouvoir de marché.⁷⁶ Dans le même temps, certains éléments de ces propositions tiennent compte de la taille des grandes plateformes, y compris par référence à leur nombre d'utilisateurs et leur chiffre d'affaires.⁷⁷

D'autres propositions impliquant la diffusion de contenu visent plus explicitement à uniformiser les règles du jeu entre les différents secteurs d'activité.⁷⁸ En outre, les autorités de régulation de la concurrence et de l'audiovisuel ont commencé à élaborer leurs propres recommandations sur la manière de gérer leurs défaillances dans la fourniture de services en ligne, notamment en élaborant des interventions réglementaires efficaces et en évitant les conséquences négatives imprévues.⁷⁹

Enfin, il convient de noter qu'un certain nombre de mesures liées à la protection des données – de la portabilité des données au consentement et à la limitation des finalités – ont une incidence sur la compétitivité des plateformes au-delà de la protection des données personnelles des utilisateurs. L'adoption de cadres solides de protection des données, tels que le Règlement général sur la protection des données, est donc essentielle pour contrôler le pouvoir de ces entreprises gigantesques et imposer des limites à leurs modèles commerciaux extractifs.

Si la plupart des nouveaux cadres réglementaires proposés visent plus directement à limiter le pouvoir excessif des grandes plateformes de médias sociaux américaines, ils ont également offert une opportunité de réglementer les petits opérateurs et d'autres services tels que la messagerie, l'hébergement dans le cloud et les services de partage des fichiers. En effet, certains fournisseurs d'infrastructures ont été entraînés dans le débat sur la modération de contenu, en particulier autour de la fourniture de leurs services à des groupes néonazis et d'extrême droite.⁸⁰

Modèle commercial des plateformes

La deuxième préoccupation essentielle derrière la volonté actuelle de renforcer la réglementation des contenus en ligne provient du modèle commercial des plus grandes plateformes de médias sociaux, basé sur la collecte de grandes quantités de données sur les utilisateurs et leurs habitudes en ligne (données comportementales) et leur monétisation à travers la publicité (ciblée) en ligne.⁸¹ Les défenseurs du droit à la vie privée et des droits numériques, en particulier, ont longtemps alerté sur le fait que le « capitalisme de surveillance est à la fois manipulateur et nuisible aux droits humains des internautes.⁸² Il s'appuie sur des techniques très intrusives (par ex. suivi en ligne, enchères en temps réel, et ciblage) et incroyablement opaques.⁸³ Ces techniques cherchent à maintenir les internautes sur les plateformes et à alimenter la diffusion généralisée de contenus « appâts à clics », sensationnalistes et « extrémistes », sans que les utilisateurs comprennent réellement pourquoi ils voient un type particulier de contenu ou de publicité sur les médias sociaux.⁸⁴ Ces techniques publicitaires empiètent donc sur l'autodétermination informationnelle des utilisateurs. Elles tendent aussi à promouvoir moins de diversité et à réduire la qualité du contenu,⁸⁵ et présentent des risques importants de discrimination dans des domaines tels que les annonces immobilières et les offres d'emploi, entre autres.⁸⁶

Échec perçu de la régulation en solo et de l'autorégulation, et manque de responsabilisation démocratique

Autre moteur clé des propositions réglementaires actuelles : la perception que les grandes plateformes de médias sociaux n'ont pas réussi à remédier à un certain nombre de « préjudices » – de la « désinformation » et du « discours de haine »⁸⁷ au contenu terroriste et aux images pédopornographiques. En particulier, des critiques soulignent la disponibilité généralisée de contenus à la fois illégaux ou « préjudiciables » en ligne – y compris leur impact parfois tragique hors ligne⁸⁸ – et le manque de cohérence et de transparence dans l'approche des médias sociaux en matière de suppression de contenu.

Une grande partie du blâme a été attribuée au modèle dominant d'exonération conditionnelle de responsabilité, perçu comme donnant carte blanche aux plateformes. Ces dernières sont perçues comme profitant des maux sociaux – qu'elles ont elles-même rendus possibles – sans aucune redevabilité publique. De ce fait, les diverses initiatives adoptées de leur propre gré par les grandes plateformes de médias sociaux⁸⁹ ou en réponse à des codes de pratique « d'autorégulation »⁹⁰ sont considérées comme insuffisantes à la fois pour résoudre ces problèmes et pour combler le déficit de redevabilité.⁹¹

Nécessité de certitude juridique et de réelle transparence

Face aux défis exposés ci-dessus, et à une myriade de lois pour y faire face, certains gouvernements ont soutenu qu'un cadre réglementaire global apporterait plus de clarté et de sécurité juridique.⁹²

De plus, ces cadres sont jugés nécessaires pour forcer les entreprises de médias sociaux à être plus transparentes sur leurs opérations de modération de contenu et le fonctionnement de leurs algorithmes. Au-delà des statistiques, on ne sait pas quel contenu est supprimé, pour quelles raisons et pourquoi les utilisateurs peuvent voir des types de contenu particuliers. Dans ces nouveaux cadres, les entreprises seraient également tenues de mettre en place des mécanismes de recours en cas de suppression illicite de contenu. Ces nouvelles obligations seraient supervisées par un régulateur non impliqué dans la prise de décision sur le contenu.⁹³

Manque de diversité des médias

Enfin, les régulateurs pourraient intervenir pour protéger le pluralisme et la diversité des médias dans l'écosystème numérique. Certains législateurs ont déjà fait des propositions pour remédier à ce qu'ils perçoivent comme une défaillance des plateformes de médias sociaux.⁹⁴ Cela pose problème pour trois raisons principales :

- **Concentration excessive des médias sociaux et goulots d'étranglement** : Un nombre très limité de grandes plateformes décident de la manière dont les informations sont diffusées sur leurs services et des critères qui s'appliquent à cette diffusion. Les grandes plateformes sont capables d'agir ainsi parce qu'elles possèdent le plus grand nombre d'utilisateurs. Les entreprises de médias n'ont donc pas d'autre choix que de collaborer avec elles, en grande partie à leurs conditions, pour bénéficier de l'attention des utilisateurs.⁹⁵
- **Contenu personnalité** : Le modèle même des plateformes de médias sociaux implique que leurs utilisateurs ne sont exposés qu'à un contenu hautement personnalisé, plutôt qu'à une diversité de points de vue.⁹⁶ Cela a suscité des craintes de bulles de filtres⁹⁷ et engendré des accusations de parti pris de la part de certains partis politiques.⁹⁸
- **Domination de la publicité en ligne** : Avec la publicité en ligne, des médias traditionnels ont subi non seulement des pertes d'audience au profit des sources en ligne « gratuites » et des médias sociaux, mais également des pertes de revenus publicitaires. Le problème de la viabilité des médias constitue également une menace pour le pluralisme et la diversité des médias.

Position d'ARTICLE 19 sur les arguments en faveur d'une réglementation des plateformes

ARTICLE 19 reconnaît que les préoccupations suscitées par une plus grande réglementation des pratiques des plateformes sont tout à fait valables. En effet, nous les partageons. Cependant, nous pensons que bon nombre des solutions actuellement proposées sont susceptibles de rater la cible et de renforcer la « domination » des plus grands opérateurs, et de faire l'objet d'abus par les gouvernements en leur octroyant plus de contrôle sur les plateformes et les contenus. Au lieu de cela, nous présentons ci-dessous quelques solutions qui garantiraient mieux selon nous la protection de la liberté d'expression.

La concentration de pouvoir entre les mains de quelques grandes plateformes devrait être traitée par des outils proconcurrentiels

Selon ARTICLE 19, les évolutions les plus récentes de la « réglementation des plateformes » soulignent la complexité des problèmes posés par les grandes plateformes numériques et la mesure dans laquelle plusieurs disciplines – de la réglementation des contenus en ligne à la concurrence, en passant par la protection des consommateurs et la protection des données – sont interdépendantes. Cela démontre également que toute intervention réglementaire doit être claire sur ses objectifs et les « préjudices » auxquels elle cherche à remédier. En particulier, ARTICLE 19 estime que les problèmes relatifs à la taille, au pouvoir ou à la « domination » des entreprises de médias sociaux devraient être traités principalement par les outils de la concurrence et des cadres économiques proconcurrentiels plutôt que par une réglementation du contenu en ligne.

Dans le même temps, nous reconnaissons que certaines entreprises de médias sociaux sont devenues à ce point dominantes que des approches à plusieurs niveaux pourraient être justifiées. Mais si les plus grands acteurs seront généralement en mesure de respecter les nouvelles obligations réglementaires (par exemple, en matière de transparence, de procédure régulière ou même d'objectifs de suppression de contenu), les petits opérateurs pourraient simplement faire faillite en s'échinant à les respecter. Autrement dit, il est vital que les petites entreprises ne soient pas indûment accablées par de lourdes obligations réglementaires, ce qui les empêcherait de rivaliser avec les grandes entreprises. Dans la suite de cette note, nous formulons des recommandations sur la façon dont ce point devrait être abordé.

Nous notons toutefois que les approches multiniveaux devront être conçues de manière à ne pas renforcer la position des acteurs en place, qui seraient capables de faire face à de nouvelles exigences réglementaires susceptibles de rendre leurs services plus attrayants aux yeux des utilisateurs (par exemple, sécurité et fiabilité).⁹⁹ Nous notons également que des approches à plusieurs niveaux pourraient involontairement diriger les « mauvais acteurs » vers des plateformes moins surveillées ;¹⁰⁰ de ce fait, les « mauvais acteurs » continueraient d'opérer en marge, avec moins de visibilité pour le public, mais plus de visibilité entre eux.

L'efficacité d'approches réglementaires plus strictes n'a pas été confirmée

ARTICLE 19 estime également qu'il est difficile de savoir si de nouvelles réglementations étatiques sur la modération de contenu des plateformes sont strictement nécessaires pour lutter contre les expressions problématiques en ligne telles que les « discours de haine » ou le terrorisme.

Nous craignons que les nouvelles législations soient contraires au droit international des droits de l'homme.¹⁰¹ L'efficacité des mesures proposées n'est pas non plus confirmée, que ce soit en termes de lutte contre des problèmes spécifiques (comme « l'extrémisme » ou « le discours de haine »),¹⁰² ou de leurs dommages collatéraux sur la liberté d'expression.¹⁰³ Trop souvent, les législateurs cherchent à légiférer pour adresser un message politique au public et lui signifier que « quelque chose est fait » au lieu d'investir des moyens dans des solutions à long terme moins visibles, mais plus efficaces.

De ce fait, ARTICLE 19 reste profondément sceptique sur la capacité des nouvelles lois à proposer des solutions pour contrer les contenus problématiques ; elles sont plus susceptibles d'être contre-productives, de se tromper de cible et de nuire à la liberté d'expression.¹⁰⁴

Des modèles de gouvernance indépendante ou multipartite devraient être mis en place

ARTICLE 19 reconnaît également que les entreprises de médias sociaux ont tardé à réagir à un certain nombre de préoccupations légitimes – telles que les problèmes relatifs au « discours de haine », à la « désinformation » ou à la protection des enfants – sur leurs plateformes. Au fil du temps, cela a considérablement érodé la confiance du public et soulevé des questions légitimes sur leur manque de responsabilité publique.

Néanmoins, nous pensons qu'il est trop tôt pour qualifier d'inefficaces les mesures et initiatives qu'elles ont adoptées pour résoudre ces problèmes. Même s'il leur reste un long chemin à parcourir, les entreprises de médias sociaux ont adopté une série de mesures de plus en plus sophistiquées au fil du temps. Il s'agit notamment d'une mise à jour et une clarification régulières de leurs politiques,¹⁰⁵ un investissement dans la mise à niveau de leurs systèmes pour détecter les contenus problématiques dans un éventail de langues plus large,¹⁰⁶ un élargissement de leurs programmes de vérification des faits¹⁰⁷ ou du signalement de confiance à un plus grand nombre de pays afin de mieux comprendre les spécificités nationales,¹⁰⁸ et l'élaboration de réponses plus sophistiquées au contenu « problématique ». Elles cherchent aussi constamment à innover et à développer de nouvelles fonctionnalités qui renforcent le contrôle des internautes sur ce qu'ils visualisent,¹⁰⁹ et elles ont mis en place des mécanismes d'appel contre les suppressions de contenus jugés illicites selon leurs normes communautaires.¹¹⁰ Certaines entreprises ont commencé à entreprendre des évaluations des droits humains,¹¹¹ et continuent d'augmenter le nombre de modérateurs humains sur leurs plateformes,¹¹² souvent en réponse à des crises particulières ou des mesures réglementaires anticipées.

De plus, les nombreux aspects des contenus problématiques qui justifient une réglementation des plateformes sont complexes et difficiles à définir. C'est le cas par exemple de la « désinformation », notamment parce que, le plus souvent, la frontière entre les faits réels et les opinions est difficile à tracer. De ce fait, l'interdiction et d'autres restrictions légales du partage de fausses informations peuvent faire l'objet d'abus et avoir un impact dévastateur sur le discours politique. ARTICLE 19 note que l'impact des politiques des médias sociaux dans ce domaine reste largement méconnu.¹¹³ Dans l'ensemble, nous pensons que la réussite ou l'échec des initiatives actuelles contre la diffusion de la « désinformation » n'est pas démontré. De plus, certaines études académiques soulignent le caractère hautement spéculatif de la nocivité de la désinformation pour la démocratie.¹¹⁴ Par conséquent, les arguments en faveur d'une plus grande réglementation dans ce domaine ne nous semblent pas justifiés.¹¹⁵ Dans le même temps, nous reconnaissons qu'une surveillance publique indépendante, ou un audit des allégations des plateformes sur la manière dont elles luttent contre la « désinformation »,¹¹⁶ peut être justifiée.

De plus, de nombreux types de contenus problématiques soulèvent des questions difficiles et complexes à résoudre. Par exemple, un certain nombre de problèmes posés par le « discours de haine » ont sans doute été exacerbés en ligne, ce type de contenu étant susceptible de devenir viral. Il est à la fois souhaitable et nécessaire que les gouvernements et les entreprises prennent des initiatives pour y apporter des solutions. La question est de savoir si les mesures adoptées jusque-là par les entreprises de médias sociaux sont suffisantes – et dans le cas contraire, si une nouvelle législation contre le « discours de haine » en ligne est une solution appropriée. De même, les mesures prises par les médias sociaux pour lutter contre le « discours de haine » ne doivent pas être ignorées d'office.¹¹⁷

On peut donc difficilement conclure que les initiatives d'autorégulation ont totalement échoué. Cela ne signifie pas que les entreprises de médias sociaux peuvent se permettre de devenir complaisantes. Elles doivent continuer à faire plus pour lutter contre les « discours de haine » et être transparentes dans ce domaine. Il est particulièrement important que les différentes parties prenantes, notamment les gouvernements, les entreprises et la société civile, continuent de dialoguer pour préserver un juste équilibre entre la protection des droits à l'égalité et à la liberté d'expression.

À cet égard, ARTICLE 19 estime que de nouvelles formes d'autorégulation multipartite ou indépendante, tels que les Conseils de médias sociaux, pourraient faire partie de la solution.¹¹⁸ Cela est particulièrement vrai dans certains domaines, notamment lorsque le contenu incriminé n'est pas illégal. Par conséquent, certaines formes d'autorégulation indépendante restent nécessaires pour les raisons suivantes :

Les arguments en faveur d'un cadre réglementaire global sont problématiques

ARTICLE 19 note que les autres arguments avancés en faveur d'un cadre réglementaire global du contenu en ligne sont très discutables. En particulier :

- **La nécessité d'une certitude juridique** : C'est un argument de poids en faveur d'un cadre réglementaire plus unifié. C'est le cas, par exemple, dans l'UE.¹¹⁹ ARTICLE 19 note toutefois que, en ce qui concerne l'UE, cet argument semble difficile à concilier avec son approche jusqu'à présent, à savoir l'adoption de plusieurs instruments juridiques distincts pour protéger le droit d'auteur ou combattre le terrorisme en ligne, et la révision de la Directive sur les services de médias audiovisuels pour lutter contre les « discours de haine » et les contenus préjudiciables aux enfants en ligne.¹²⁰ Dans de nombreux cas, les mesures adoptées par l'UE ces dernières années n'ont pas encore été mises en œuvre. Il n'est donc

pas certain qu'elles puissent être efficaces. Dans l'ensemble, nous pensons que l'adoption concomitante d'une loi sur les services numériques n'apportera pas nécessairement une plus grande clarté dans ce domaine du droit. On peut également se demander si un régulateur unique, qu'il soit national ou européen, serait bien placé pour traiter des problèmes aussi divers que la « désinformation », le « discours de haine », les images pédopornographiques et les contenus terroristes.

- **Réglementation d'un large éventail de services** : En outre, comme nous l'expliquons plus en détail ci-dessous, nous craignons que la réforme dans ce domaine puisse être utilisée pour réglementer des services ou des acteurs tels que les applications de messagerie privée ou la presse. Ces acteurs devraient être exclus du champ d'application en raison de la menace qu'une telle réglementation représenterait pour la protection des droits à la liberté d'expression et à la vie privée.
- **Setting up a one-stop-shop regulatory body**: Enfin, si nous admettons qu'un organisme de réglementation peut contribuer au respect des obligations de transparence et de procédure régulière, nous craignons que les questions de procédure soient souvent mêlées à celles du contenu, telles que l'adoption de filtres pour détecter les contenus préjudiciables ou illégaux. Pour les gouvernements, les régulateurs restent également un puissant moyen d'exercer un contrôle sur les flux d'information, en particulier dans les pays où le régulateur n'est pas indépendant. L'attribution de pouvoirs aux régulateurs sur la fourniture de services de médias sociaux pourrait finalement avoir un effet dissuasif sur la liberté d'expression des internautes. Regulators also remain a powerful avenue for governments to exercise control over information flows, particularly in countries where the regulator is not independent. Giving regulators powers over the provision of social media services could ultimately have a chilling effect on users' freedom of expression.

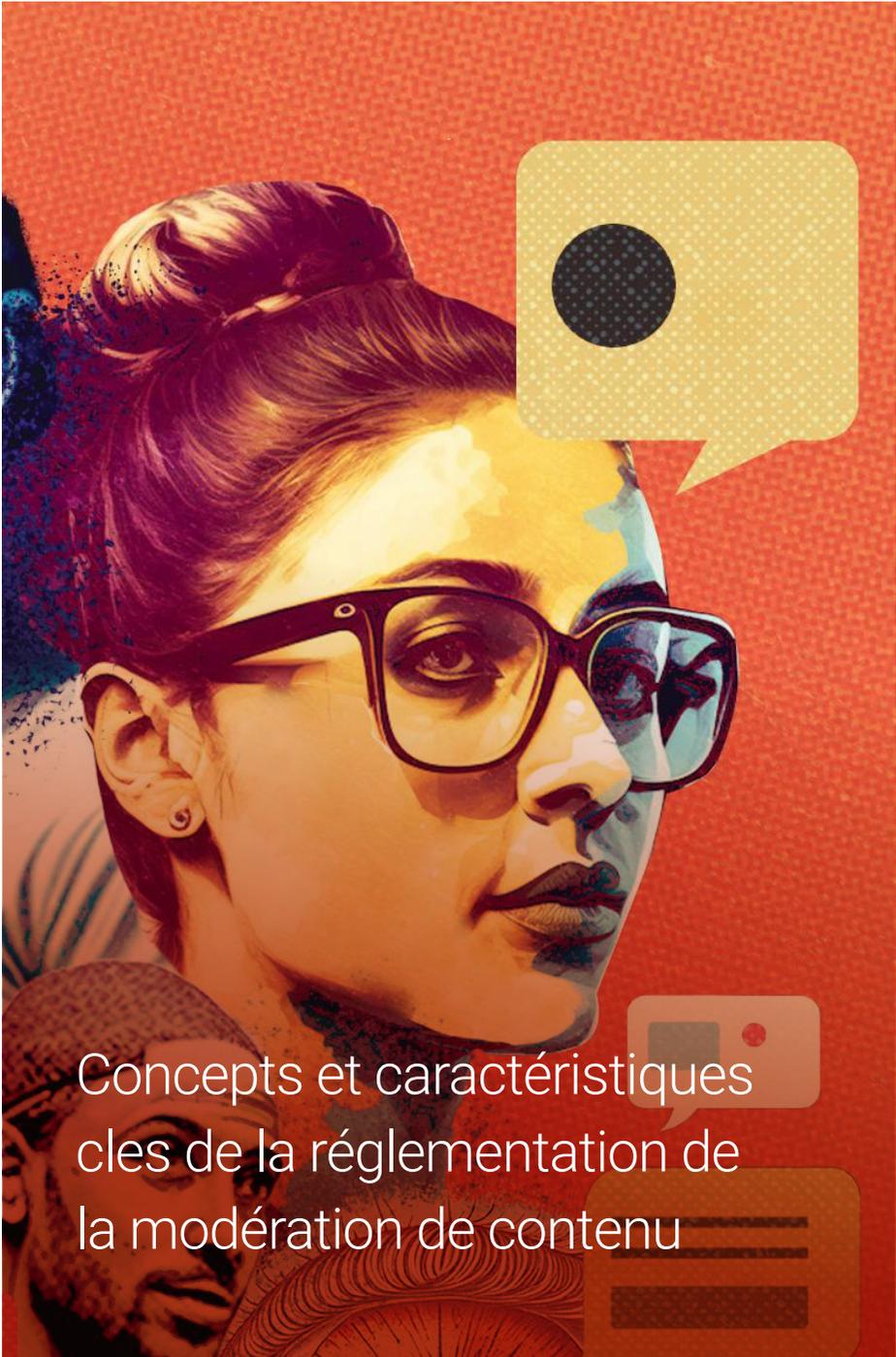
ARTICLE 19 n'est donc pas convaincu que les arguments en faveur de l'exonération de responsabilité soient justifiés, en particulier pour le contenu entrant dans la catégorie « préjudiciable, mais légal ». Nous pensons que d'autres mécanismes, tels que des organes indépendants multipartites (Conseils de médias sociaux), pourraient être plus aptes à veiller à la conformité des entreprises de médias sociaux à un ensemble de principes dérivés des normes relatives aux droits humains.

Il faut remédier au manque de diversité des médias

En tant qu'organisation de défense de la liberté d'expression, ARTICLE 19 prend très au sérieux les préoccupations relatives au pluralisme et à la diversité des médias. Si les inquiétudes sur les bulles de filtres ont pu être exagérées,¹²¹ il est tout aussi clair que les algorithmes utilisés pour la curation de contenu ne sont pas neutres en termes de diversité, et que leurs modifications sont susceptibles d'avoir un impact sur la consommation de l'information. Des faits probants démontrent également que les audiences en ligne sont plus polarisées, ce qui pourrait à son tour inciter à la production de contenus plus partisans.¹²² En fin de compte, la préoccupation primordiale reste que certaines plateformes ou services de recherche détiennent trop de pouvoir sur les conditions de diffusion des actualités – et au final sur la survie du journalisme – en raison de leur taille, de leurs parts de marché et de leur base d'utilisateurs. ARTICLE 19 estime qu'une partie de ces problèmes pourraient être résolus par un certain nombre de mesures, notamment les Conseils de médias sociaux (qui favoriseraient le dialogue entre les plateformes et les organes de presse) et le dégroupage des services d'hébergement et de curation de contenu (ce qui favoriserait la concurrence et donc la diversité et le choix des utilisateurs, dans la curation de contenu). Ci-dessous, nous expliquons plus en détail à quoi pourraient ressembler ces propositions.

Le modèle économique des plateformes doit être réformé pour se conformer à la législation sur la protection des données

ARTICLE 19 estime que le modèle commercial des grandes plateformes de médias sociaux et d'autres entreprises porte atteinte de manière significative au droit à la vie privée et peut avoir un effet dissuasif sur la liberté d'expression. Au minimum, il est clair que les entreprises de médias sociaux doivent se conformer à la législation sur la protection des données pour limiter la quantité de données personnelles collectées. ARTICLE 19 note également qu'à la lumière des évolutions récentes (en particulier autour des élections), les pratiques actuelles de publicité politique doivent être revues pour leur impact sur la liberté d'expression, et une réglementation pourrait être nécessaire dans ce domaine. ARTICLE 19 aborde cette question dans une autre note d'orientation.¹²³



Concepts et caractéristiques
cles de la réglementation de
la modération de contenu

Concepts et caractéristiques clés de la réglementation de la modération de contenu

En l'espace de quelques courtes années, le large consensus dans les pays occidentaux sur la nécessité d'une exonération de responsabilité conditionnelle en tant que préalable à l'innovation technique et à la liberté d'Internet a été progressivement remis en cause. De même, l'interdiction de surveillance générale du contenu sur les plateformes de médias sociaux a été constamment ébranlée. Comme nous l'avons précédemment noté, un certain nombre d'États ont progressivement adopté ou proposé une série de mesures qui érodent ces fondements importants de la protection de la liberté d'expression en ligne.¹²⁴ Dans ce chapitre, nous examinons quelques concepts clés promus dans les propositions législatives dans ce domaine.

Pour commencer, ARTICLE 19 relève un certain nombre d'idées fausses récurrentes dans les débats sur la réglementation des contenus en ligne :

- **Responsabilité vs réglementation** : Les règles de responsabilité permettent aux particuliers de poursuivre des personnes physiques ou morales pour leur avoir infligé un préjudice sur la base de causes d'action existantes. À l'inverse, l'exonération de responsabilité dans le secteur numérique signifie que les entreprises numériques sont protégées contre les poursuites, sauf si elles ne prennent aucune mesure malgré la reconnaissance de l'illégalité. Compte tenu des millions de contenus publiés chaque minute, et de la possibilité que ce contenu soit illégal, l'immunité de responsabilité est une protection très importante pour les entreprises – et pour la liberté

d'expression plus généralement. Si cette protection venait à disparaître, et vu le risque juridique encouru, les entreprises cesseraient probablement d'exister. Cela serait également néfaste pour la liberté d'expression. La réglementation, en revanche, signifie que les entreprises doivent se conformer à un ensemble défini d'obligations prévues par la loi et, en cas de non-respect de ces obligations, elles peuvent faire l'objet de sanctions telles que des amendes. L'exonération de responsabilité et la réglementation peuvent coexister.

- **Contenu vs réglementation des plateformes** : Le débat actuel sur la manière dont les entreprises numériques sont réglementées implique souvent des références à la « réglementation du contenu » ou à la « réglementation des plateformes ». Il convient de rappeler que le contenu est principalement réglementé par une loi qui s'applique directement aux individus, par ex. la loi sur la diffamation ou la pénalisation de l'incitation à la violence. La situation des entreprises est différente dans la mesure où elles autorisent des tierces parties (les utilisateurs) à publier du contenu sans vérification préalable (bien que cela change progressivement avec l'utilisation de la technologie du filtrage). Alors que les propositions législatives actuelles sont présentées comme des propositions de réglementation des plateformes, elles délèguent largement la responsabilité de la régulation de l'expression des utilisateurs aux plateformes, souvent au-delà des exigences de la loi (voir ci-dessous). Par ailleurs, il est important de garder à l'esprit que, dans la pratique, la réglementation vise les activités, le comportement et les services des entreprises et non les entreprises elles-mêmes.

- **Responsabilité de l'éditeur vs responsabilité de l'intermédiaire** : Les entreprises numériques sont de plus en plus souvent qualifiées d'« entreprises médiatiques ». Certains soutiennent que la responsabilité applicable aux éditeurs devrait également s'appliquer aux plateformes dominantes telles que Facebook, Twitter, et YouTube. Si l'on peut dire que certaines plateformes de médias sociaux se comportent de plus en plus comme des entreprises de médias traditionnelles, il subsiste néanmoins certaines différences essentielles entre les deux. En particulier, les plateformes ont des millions d'utilisateurs qui publient du contenu, et qui en sont donc les principaux éditeurs. Les plateformes de médias sociaux sélectionnent et organisent ce contenu. De ce fait, elles exercent une forme de responsabilité éditoriale – processus parfois qualifié de « curation de contenu ». Quoi qu'il en soit, elles ne sont pas responsables du contenu lui-même tant qu'elles ne le modifient pas au point qu'il devienne le leur. La question clé est de savoir si la simple utilisation d'algorithmes et de filtres permet véritablement aux entreprises technologiques d'avoir une connaissance réelle de l'illégalité de certains contenus sur leurs plateformes, perdant ainsi l'exonération de responsabilité. Nous répondons à cette question dans le prochain chapitre.

Caractéristiques positives des propositions législatives sur la réglementation de la modération de contenu

ARTICLE 19 constate que plusieurs propositions réglementaires dans ce domaine comprennent quelques aspects positifs en termes de protection du droit à la liberté d'expression, notamment :

- **Des exigences de transparence renforcées** sur la manière dont les médias sociaux et d'autres sociétés gèrent la modération de contenu, y compris la transparence sur la suppression de contenus des utilisateurs ou sur le traitement des plaintes des usagers.¹²⁵
- **Obligation de mettre en place des mécanismes de recours et diverses garanties procédurales**, notamment la notification de retrait aux utilisateurs dont le contenu a été supprimé, et la possibilité pour les utilisateurs de contester les décisions des sociétés devant les tribunaux.¹²⁶
- **Exonération de responsabilité** : La plupart des propositions¹²⁷ ne rejettent pas explicitement l'exonération de responsabilité. Cependant, elles la mettent en danger en imposant (plus ou moins explicitement) des « mesures proactives » pour détecter les contenus illégaux ou préjudiciables. La question est alors de savoir si les entreprises conservent suffisamment d'incitations à protéger la liberté d'expression quand elles bénéficient d'une exemption de responsabilité par le biais d'une « clause du bon samaritain »,¹²⁸ y compris lorsqu'elles mettent en place des filtres et d'autres mesures ? Dans l'ensemble, pour des raisons que nous expliquerons plus loin, nous pensons que la réponse à cette question est « oui ».

- **Approches systématiques** : Alors que les particuliers conservent la possibilité de poursuivre les entreprises qui ont omis de supprimer des contenus « illégaux », les propositions réglementaires actuelles ne recommandent pas de sanctionner les échecs uniques de suppression de contenu ; elles adoptent plutôt une approche « macro » de la modération du contenu en se concentrant sur les systèmes et les processus des entreprises.¹²⁹
- **Approches à plusieurs niveaux** : Dans la mesure où une telle réglementation est jugée nécessaire – ce qui n’est pas notre avis –, il y a un aspect positif dans nombre de propositions : elles tendent à prévoir des approches à plusieurs niveaux en fonction de la taille de l’entreprise, du moins par référence au nombre de ses utilisateurs.¹³⁰

Lacunes des propositions réglementaires sur la modération de contenu

Malgré plusieurs aspects positifs, la plupart des propositions actuelles sur la modération de contenu créent des problèmes importants du point de vue des normes juridiques internationales sur la protection de la liberté d’expression. Comme souligné précédemment, pour qu’une restriction du droit à la liberté d’expression soit justifiée au regard du droit international, elle doit être prévue par la loi, poursuivre un but légitime parmi les objectifs énumérés de manière exhaustive dans les traités relatifs aux droits humains, et être nécessaire et proportionnée à cet objectif. La grande majorité des propositions ne satisfont pas aux critères de légalité et de proportionnalité.

Le contenu « légal, mais préjudiciable »

La plupart des propositions actuelles vont au-delà de la portée du sujet, car elles couvrent un contenu qui n’est pas seulement illégal, mais aussi simplement préjudiciable.¹³¹

Cela est profondément préoccupant, car cela signifie que les entreprises seront probablement tenues légalement d'agir contre des contenus effectivement autorisés par la loi. Les individus seraient autorisés à dire des choses dans la rue qui ne sont pas permises dans les médias sociaux. De plus, la grande majorité des contenus « préjudiciables » sont intrinsèquement difficiles à définir, et donc susceptibles de conduire à une censure du contenu qui – bien qu'offensant pour certains – devrait rester en place. C'est le cas par exemple d'un contenu ayant pour effet « probable » d'humilier une personne,¹³² ou d'un discours « extrémiste » « avec une définition moins claire ».¹³³ Même si certaines des propositions offrent une plus grande certitude juridique (puisqu'elles concernent un contenu illégal), les lois auxquelles elles se réfèrent suscitent des inquiétudes du point de vue de la liberté d'expression.¹³⁴

Une portée trop large dans l'éventail des entreprises concernées

Il convient également de s'interroger sur la mesure dans laquelle un large éventail de services tels que la messagerie privée ou les sites d'actualités sont susceptibles d'être couverts¹³⁵ par les propositions existantes alors qu'ils ne le devraient pas. ARTICLE 19 est particulièrement préoccupé par le fait que la réglementation pourrait obliger les entreprises concernées à filtrer les contenus, ce qui affaiblirait considérablement le chiffre d'affaires, et donc la confidentialité des communications des utilisateurs. Pour les sites d'actualité, nous craignons qu'ils ne soient soumis à une réglementation susceptible de réduire au silence leurs contenus, en particulier lorsqu'ils critiquent la classe politique. De plus, dans les propositions existantes, il semble que l'on n'ait guère réfléchi à la mesure dans laquelle des prestataires de services différents peuvent être affectés différemment : par exemple, un fournisseur d'infrastructure (par ex. Cloudflare) ne devrait pas être soumis aux mêmes obligations qu'une plateforme de médias sociaux (par ex., Facebook).

Si certaines propositions tendent à adopter une approche multiniveaux des obligations imposées aux entreprises dans leur champ d'application, que ce soit par référence au nombre de leurs utilisateurs ou à leur chiffre d'affaires, cela n'est pas toujours le cas.¹³⁶ De plus, les critères utilisés ne sont pas toujours cohérents ; des exceptions ne sont pas toujours prévues pour les organisations à but non lucratif (par ex. Wikipedia), et de petits fournisseurs ayant une large base d'utilisateurs, mais un faible chiffre d'affaires, peuvent encore être soumis à des obligations onéreuses.¹³⁷

Obligations vagues et problématiques

Alors que certaines obligations proposées sont généralement positives (par ex. sur la transparence et l'obligation de procédure régulière), d'autres sont soit trop vagues, soit profondément hostiles aux droits à la liberté d'expression et à la vie privée :

- **Obligation de diligence** : Certaines propositions font référence à un « devoir de diligence » général sans définir clairement les types d'obligations qu'il implique¹³⁸ ou sa relation avec la responsabilité de l'intermédiaire. En général, l'approche du « devoir de diligence » est basée sur une évaluation du risque de « préjudice » pour les utilisateurs. Même si un « devoir de diligence » impliquait une évaluation des risques de « préjudice » découlant des systèmes des entreprises,¹³⁹ le terme de « préjudice » est difficile à définir, et la définition de ce qui constitue des risques « inacceptables » relèverait du pouvoir discrétionnaire du régulateur.¹⁴⁰ Sans plus de clarté, un « devoir de diligence » pourrait ressembler à l'une des obligations détaillées ci-dessous, c'est-à-dire la suppression de contenu dans des délais exagérément courts, des mesures proactives ou des obligations de suppression du chiffrage.¹⁴¹

ARTICLE 19 observe que les partisans d'un « devoir de diligence » obligatoire semblent suggérer que leurs propositions s'alignent sur le type d'évaluations d'impact sur les droits humains encouragées par les Principes directeurs.¹⁴² Dans cette analyse, la suggestion semble être que les évaluations des risques possibles pour les droits humains sont les mêmes que les évaluations des risques de « préjudice » potentiel. ARTICLE 19 estime qu'il s'agit là d'une erreur fondamentale. Les droits humains sont des concepts juridiques inscrits dans les lois nationales, les constitutions et les instruments internationaux. Ils sont généralement formulés dans des termes généraux, mais ils sont aussi définis par un vaste corpus de décisions rendues par des cours et des tribunaux. Par conséquent, malgré leur apparence très large, ils sont en fait définis minutieusement. La notion de « préjudice », en revanche, n'a pas de définition juridique générale. Dans la mesure où elle est utilisée dans des contextes juridiques, elle tend à être minutieusement définie et très spécifique au contexte, par ex. l'infraction pénale consistant à infliger des lésions corporelles réelles. Les Principes directeurs sont conçus pour prévenir des violations des droits humains et non pour prévenir des « préjudices ». En outre, contrairement au « devoir de prévenir un préjudice », les Principes directeurs n'obligent pas les entreprises à prévenir les violations des droits humains par d'autres en tant que tels.¹⁴³ Tout au plus, ils obligent les entreprises à adopter des procédures visant à éviter ou réduire leur implication dans ces violations.

- **Obligation de retrait de certains contenus** : La grande majorité des propositions délèguent effectivement des pouvoirs de censure à des entreprises privées pour supprimer des contenus illégaux ou préjudiciables. Si les entreprises ne le font pas, le régulateur peut ordonner le blocage de ces services.¹⁴⁴ Bien que certaines propositions incluent une obligation pour l'organisme de surveillance de s'assurer que les entreprises ne suppriment pas excessivement des contenus,¹⁴⁵ il n'y a généralement pas d'obligation de transparence équivalente pour les

contenus supprimés à tort. En tout état de cause, sauf s'il y a réclamation, le nombre de contenus supprimés à tort peut ne jamais être connu. De plus, cela dépend en partie de l'efficacité des mécanismes internes de plainte et d'autres formes de réparation judiciaire. Cependant, la plupart des propositions restent totalement muettes sur la manière dont ces dernières devraient être financées.

- **Obligation de supprimer des contenus dans de brefs délais :** En général, la plupart des propositions tendent à exiger la suppression du « discours de haine » ou du contenu « manifestement illicite » dans un délai de 24 heures et les contenus terroristes en une heure.¹⁴⁶ Ces délais sont manifestement trop courts pour procéder à une évaluation correcte – et encore moins juridique – des réclamations présentées.
- **Obligation de prendre des mesures proactives :** Certaines propositions contiennent une obligation pour les entreprises d'adopter des mesures « proactives » ou de faire « de leur mieux » pour lutter contre les contenus illégaux ou préjudiciables. Parfois, cette obligation est présentée comme une « recommandation » émise par le régulateur, bien que la loi indique clairement que le respect de ces recommandations est attendu.¹⁴⁷ Dans la pratique, les « mesures proactives » et les termes associés signifient l'adoption de filtres ou d'autres technologies visant à identifier ou empêcher le téléchargement de contenus « problématiques ». Bien sûr, les filtres ne peuvent pas évaluer la légalité d'un contenu et ils sont pratiquement toujours voués à l'erreur – en particulier lorsqu'une analyse du contexte est nécessaire, comme dans le cas du propos « terroriste », ou du « discours de haine ». Par conséquent, le contenu légitime peut être supprimé à tort.

L'organe de régulation

La plupart des propositions de réglementation des plateformes de médias sociaux viseraient à les placer sous la tutelle d'un organisme réglementaire de « l'audiovisuel », ce qui élargirait considérablement les attributions et les pouvoirs de ce régulateur. Pour ARTICLE 19, cela pose plusieurs problèmes :

- **Nécessité et légitimité** : Pour commencer, nous notons que ces propositions placeraient finalement le discours des utilisateurs sous le contrôle d'une autorité publique qui se verrait accorder de vastes pouvoirs discrétionnaires sur ce qui constitue un contenu « préjudiciable ». La nécessité d'une telle décision et son impact probable sur le droit des internautes à la liberté d'expression ne sont pas clairs. En tout état de cause, il est largement inapproprié pour une entreprise privée ou un régulateur de déterminer la légalité d'un contenu : c'est l'affaire des tribunaux.
- **Indépendance** : Une autre préoccupation majeure pour une organisation mondiale comme ARTICLE 19 consiste dans le fait que le régulateur peut ne pas être indépendant, que ce soit en droit ou en pratique. C'est loin d'être une préoccupation théorique, y compris dans les démocraties bien établies,¹⁴⁸ mais aussi dans de nombreux pays où l'état de droit est faible ou menacé.
- **Guichet unique** : Enfin, nous craignons que les propositions actuelles placent une vaste gamme de contenus sous la responsabilité d'un organisme réglementaire susceptible de ne pas être le mieux armé pour traiter ces problèmes, et encore moins pour fournir le type de solutions qui contribueraient à résoudre ces problèmes. Par exemple, le retrait de contenus pédopornographiques est un domaine spécialisé et pouvant impliquer des procédures et des considérations spéciales, qui seraient mieux gérées par une institution dédiée. À notre avis, un organisme de régulation de type « guichet unique » pour tous les contenus illégaux et préjudiciables a peu de chances d'être efficace.

Sanctions disproportionnées

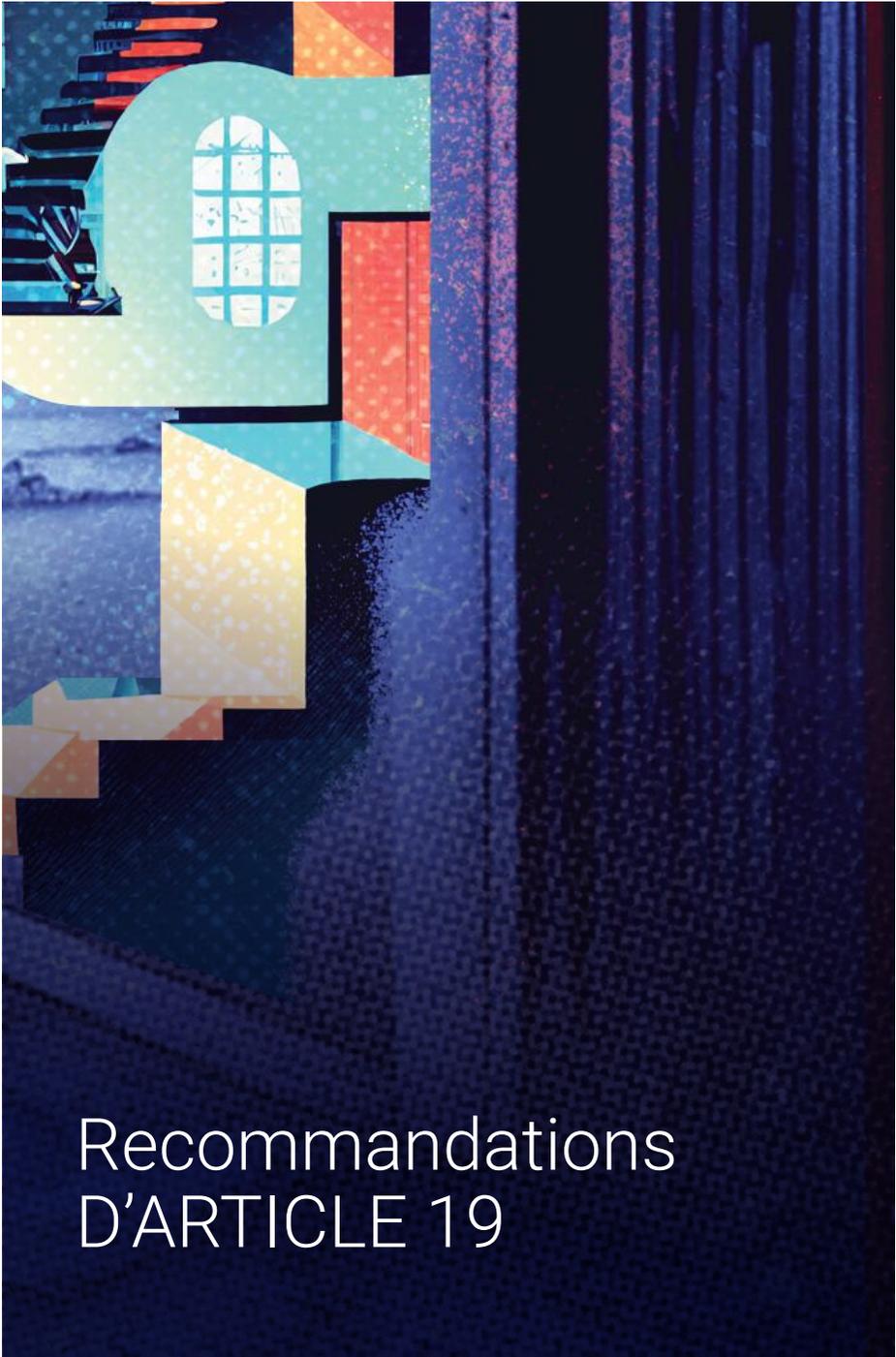
ARTICLE 19 note que l'éventail des sanctions actuellement envisagées dans divers pays, comme la France ou le Royaume-Uni, est particulièrement élevé, par ex. jusqu'à 4 % du chiffre d'affaires mondial des entreprises. Combinés à des obligations trop vagues, nous craignons que cela n'incite les entreprises à supprimer plus de contenu au détriment de la protection de la liberté d'expression en ligne. De plus, nous notons que le seuil auquel les sanctions deviennent probables n'est pas toujours clair¹⁴⁹ et que ces mesures sont disproportionnées.

Autres préoccupations

ARTICLE 19 craint que les lois ou projets de loi sur la réglementation des plateformes contiennent de plus en plus d'obligations pour les entreprises de désigner des points de contact ou des représentants « dans le pays ». En pratique, cela permet aux autorités d'exercer une pression significative sur les entreprises pour qu'elles se conforment à leurs exigences, même si celles-ci n'ont que peu ou pas de base légale et visent manifestement à faire taire les détracteurs des gouvernements.¹⁵⁰

Nous nous demandons également dans quelle mesure les entreprises ayant des activités mondiales peuvent être en mesure de se conformer à des régimes différents – et potentiellement contradictoires – à travers le monde. À notre avis, cela pourrait conduire à une fragmentation d'Internet, et potentiellement à une course vers le bas, les contenus controversés étant de plus en plus supprimés par défaut en vertu des conditions de service des entreprises.

Enfin, nous notons que les lois dans ce domaine sont fréquemment adoptées sans processus participatif approprié, dans lequel tous les secteurs, y compris les plateformes numériques petites et locales, seraient inclus.



Recommandations D'ARTICLE 19

Recommandations D'ARTICLE 19

Recommandation 1 :

Les États doivent s'abstenir de réglementer inutilement la modération de contenu en ligne

ARTICLE 19 affirme que nous ne soutenons pas les modèles de régulation qui ont émergé en Europe occidentale ou en Amérique latine. Nous craignons toujours que les règles proposées soient trop souvent vagues et fondées sur l'indépendance d'un organisme de réglementation de bonne foi. Ce n'est tout simplement pas le cas dans de nombreuses régions du monde. Le contexte plus large dans lequel les propositions réglementaires sont formulées est particulièrement important.¹⁵¹ En outre, en vertu du test en trois parties des restrictions à la liberté d'expression, les États doivent démontrer de manière exhaustive l'existence de preuves suffisantes pour justifier la nécessité d'une réglementation. Si la forte envie de réglementer les très grandes plateformes est compréhensible, ces dernières seront toujours capables de s'adapter et de répondre aux exigences réglementaires. Cela n'est probablement pas le cas pour les plateformes plus petites ou de moyenne taille. À tout le moins, une réglementation en ligne plus poussée risquerait de profiter aux très grandes entreprises et de renforcer leur position dominante.

Nous exhortons donc les législateurs à résister à la tentation de réglementer inutilement. Si les plus grandes préoccupations de nos sociétés résultent du pouvoir excessif et de la domination d'un petit nombre de plateformes, il nous semble que les législateurs devraient répondre à ces préoccupations en utilisant des recours disponibles dans la législation sur la concurrence et la réglementation économique proconcurrentielle. C'est également pour cette raison que nous pensons que nos positions et nos recommandations sur la responsabilité des intermédiaires et les règles de modération de contenu des plateformes de médias sociaux sont toujours valables.¹⁵²

De plus, ARTICLE 19 a plaidé pour une surveillance des entreprises de médias sociaux par une institution multipartite indépendante, telle que les Conseils de médias sociaux. Nous continuons de croire que la création d'une institution à partir de la base conduirait à de meilleurs résultats pour la liberté d'expression, en excluant des sanctions trop sévères. Cela favoriserait également une responsabilisation plus efficace puisque ses parties prenantes s'approprieraient le succès de l'organisation. Les Conseils de médias sociaux et une réglementation des très grandes plateformes par le biais de la législation sur la concurrence peuvent jouer des rôles complémentaires.

Recommandation 2 :

Les principes fondamentaux de tout cadre réglementaire doivent être la transparence, la responsabilisation et la protection des droits humains

Nonobstant ce qui précède, ARTICLE 19 reconnaît qu'une plus grande réglementation du secteur numérique, et des entreprises de médias sociaux en particulier, paraît quasiment inévitable.¹⁵³ De ce fait, nous pensons que les objectifs de tout nouveau cadre réglementaire ne peuvent se limiter à la seule « lutte contre l'illégalité ». De même, la « prévention des préjudices » est un concept beaucoup trop large pour être un objectif significatif, et encore moins légitime, d'un tel cadre.

Nous pensons plutôt que les principes fondamentaux de la réglementation doivent être la transparence, la responsabilisation et la protection des droits humains, cette dernière signifiant que les principes de légalité et de proportionnalité doivent être respectés de bout en bout. En outre, tout cadre de ce type doit être fondé sur des preuves solides pour adopter les solutions les plus appropriées.

Recommandation 3 :

L'exonération de responsabilité conditionnelle pour le contenu de tiers doit être maintenue – mais sa portée et ses procédures de notification et d'action doivent être clarifiées

ARTICLE 19 note que la suppression – ou la limitation abusive – de l'exonération de responsabilité pourrait inciter les entreprises numériques soit à filtrer et supprimer autant de contenu que possible, soit à rester totalement neutres et à n'en supprimer aucun. En d'autres termes, cela pourrait conduire à une censure accrue ou à la disparition de toute incitation par les entreprises à s'engager dans la modération de contenu. À notre avis, ces deux résultats seraient hautement indésirables pour la protection de la liberté d'expression en ligne.

De ce fait, nous estimons que les entreprises numériques et de médias sociaux doivent continuer à bénéficier d'une exonération de responsabilité étendue, ou du moins conditionnelle. À cet égard, nous notons que, par définition, la responsabilité ne peut être attachée qu'à un contenu illégal – plutôt que « préjudiciable ». En pratique, cela signifie que les entreprises ne devraient pas être tenues responsables de ne pas prendre des « mesures raisonnables » indéfinies pour traiter les contenus « préjudiciables ».

La responsabilité doit être différente selon les types d'activités et de services, comme suit :

Une large exonération de responsabilité pour les fournisseurs de services d'infrastructure, y compris le « simple conduit » et « l'hébergement » neutre, devrait être maintenue

ARTICLE 19 estime que les prestataires de services d'infrastructure essentiels au fonctionnement d'Internet, tels que les réseaux de diffusion de contenu, devraient bénéficier d'une exonération de responsabilité plus large que les services engagés dans la modération de contenu à l'échelle de l'application. Ils ne devraient être contraints de supprimer un contenu que sur injonction d'un tribunal. Dans la pratique, cela signifie que les fournisseurs d'infrastructure ne devraient pas être pénalisés pour l'hébergement de certains sites, à moins qu'ils n'aient pas respecté une décision de justice valide leur enjoignant de ne plus fournir leurs services à un tel site parce qu'il est illégal.¹⁵⁴ De même, les fournisseurs de services d'infrastructure ne devraient pas être tenus d'héberger un tel site s'ils ne le souhaitent pas. En d'autres termes, les fournisseurs de services d'infrastructure essentiels ne devraient pas être tenus de transporter du contenu.

Dans le même temps, les services fournissant un service essentiel devraient énoncer clairement les raisons pour lesquelles ils peuvent décider de cesser la prestation de leurs services à certains acteurs.

Procédures de notification et d'action pour les fournisseurs de services d'hébergement couplés à la modération de contenu

ARTICLE 19 estime en outre que le niveau actuel de connaissances requis pour bénéficier de l'exonération de responsabilité doit être maintenu, c'est-à-dire que la connaissance doit rester « réelle » plutôt que « constructive ». ARTICLE 19 continue de croire que la connaissance réelle de l'illégalité ne peut être obtenue que par une ordonnance du tribunal. Le contraire équivaldrait à accepter qu'un contenu est illégal simplement parce qu'un tiers, tel qu'un détenteur de droits d'auteur, l'a dit.

Dans le même temps, nous reconnaissons qu'un cadre réglementaire pourrait clarifier les différents types de procédures de notification et d'action applicables aux différents types de contenu – sans préjudice de la détermination de la légalité. En d'autres termes, la loi devrait clarifier la manière dont les entreprises obtiennent une « connaissance réelle » de l'illégalité présumée, et ce qu'elles doivent faire à ce sujet une fois qu'elles l'ont obtenue.

ARTICLE 19 a précédemment expliqué comment cela pourrait fonctionner dans la pratique.¹⁵⁵ Nous croyons que les processus que nous suggérons restent valables et offrent la meilleure voie pour protéger le droit à la liberté d'expression. En résumé, cela comprend :

« Notification et notification » pour les litiges privés, tels que les droits d'auteur ou la diffamation

Dans le cadre de cette procédure, le plaignant ou « signaleur de confiance » serait tenu de donner son nom et d'indiquer dans un avis pourquoi il pense que ses droits ont été enfreints, quel est le fondement juridique de sa requête, l'emplacement du matériel prétendument incriminé et l'heure et la date de l'infraction présumée. Le fournisseur de services d'hébergement sera tenu de transmettre la notification au contrevenant présumé (le fournisseur de contenu) dès que possible, mais dans un délai maximal (par ex. 72 heures). Le fournisseur de contenu aura le choix de supprimer le contenu ou de déposer une contre-notification dans un délai raisonnable (par ex. 14 jours).

Là encore, l'hébergeur serait tenu de transmettre la contre-notification dès que possible, mais dans un délai maximal (de 72 heures par exemple). Le plaignant disposerait alors d'un certain délai (14 jours par exemple) pour décider s'il souhaite porter l'affaire devant la justice. Le contenu serait supprimé suite à une décision du tribunal. Un hébergeur pourrait être tenu responsable de dommages-intérêts légaux s'il ne se conforme pas à ses obligations de « notification et notification » ou s'il ne supprime pas le contenu à la suite d'une ordonnance du tribunal.

En revanche, si le fournisseur de contenu ne répondait pas ou ne fournissait pas de contre-notification dans un délai donné, l'hébergeur perdrait l'exonération de responsabilité. Il pourrait soit supprimer le contenu prétendument illégal, soit être tenu responsable du contenu en cause si le plaignant décide de porter l'affaire devant un tribunal ou un autre organe juridictionnel indépendant. Pour protéger la liberté d'expression, tout nouveau cadre « de notification et notification » devrait également prévoir des sanctions pour notifications abusives.

« Notification et action » pour allégations de crime grave

Dans le cadre de cette procédure, un fournisseur d'hébergement serait tenu de retirer le contenu après réception d'une injonction du tribunal à cet effet. En d'autres termes, il serait tenu responsable s'il ne se conformait pas à cette injonction. En pratique, cela signifierait que si les forces de l'ordre estiment qu'un élément de contenu doit être supprimé et que l'affaire n'est pas urgente, elles devraient requérir une ordonnance auprès du tribunal, si nécessaire sur une base ex parte. Si la situation est urgente (par ex. la vie d'une personne est en danger), les forces de l'ordre devraient être habilitées par la loi à ordonner la suppression immédiate ou le blocage de l'accès au contenu en question. Toutefois, toute ordonnance de ce type doit être confirmée par un tribunal dans un délai déterminé (par ex. 48 heures). Le recours à des mécanismes informels – par ex. des appels téléphoniques ou des courriels demandant à l'hébergeur de supprimer le contenu – ne devrait pas être autorisé.

En revanche, si les hébergeurs reçoivent une notification d'un utilisateur ordinaire sur un contenu suspecté d'être criminel, l'hébergeur ou la plateforme devrait à son tour informer les autorités chargées de l'application de la loi s'ils ont des raisons de croire que la plainte est fondée et mérite une enquête plus approfondie. L'hébergeur ou la plateforme pourraient également décider de supprimer le contenu concerné, à titre provisoire, conformément à leurs conditions d'utilisation. Cependant, ils ne seraient pas obligés de le faire, et la non-suppression du contenu en cause n'engagerait pas leur responsabilité.

Le même processus devrait s'appliquer aux organismes privés qui travaillent avec les forces de l'ordre et exploitent des hotlines que des utilisateurs individuels d'Internet peuvent appeler s'ils soupçonnent qu'un contenu criminel a été publié en ligne.¹⁵⁶ En d'autres termes, la hotline signalerait le contenu en question à la fois à l'hébergeur et aux forces de l'ordre. L'hébergeur emploierait la même procédure que celle utilisée pour les plaintes des utilisateurs ordinaires, c'est-à-dire qu'il resterait libre de décider de supprimer ou non un contenu en se fondant sur ses conditions d'utilisation. Le même modèle pourrait s'appliquer à d'autres organismes, qu'ils soient publics ou privés, qui reçoivent des plaintes du public concernant des contenus potentiellement criminels en ligne, ou pour signaler des problèmes par des « signaleurs de confiance » (voir ci-dessous pour plus de détails sur les programmes de signaleurs de confiance). Quelle que soit l'option retenue, il est important que les autorités soient informées de toute allégation de comportement criminel grave afin qu'une enquête appropriée soit menée et traitée conformément à la procédure établie du système judiciaire pénal.

Les Principes de Manille sur la responsabilité des intermédiaires fournissent d'autres orientations utiles sur la manière dont les procédures de « notification et action » devraient fonctionner.¹⁵⁷ Nous pensons que c'est la manière la plus proportionnée, et la plus respectueuse des droits, d'appliquer les procédures de « notification et action », en particulier à l'encontre des petites entreprises.

Protection contre la responsabilité en cas de mesures de modération de contenu appliquées d'office par les entreprises

ARTICLE 19 estime que les plateformes de médias sociaux et autres entreprises numériques ne devraient pas être tenues responsables du fait d'avoir simplement adopté des normes communautaires et utilisé des modérateurs humains ou d'autres instruments pour les faire respecter.¹⁵⁸ En ce sens, nous soutenons l'adoption d'une « clause du bon samaritain » encourageant

les efforts de modération de contenu faits de bonne foi. À notre avis, le fait de ne pas le faire empêcherait l'adoption de solutions et d'outils techniques innovants, tels que la démonétisation ou la suppression de certaines fonctionnalités de la plateforme, qui permettraient de préserver un équilibre plus proportionné entre la protection de la liberté d'expression et la lutte contre les contenus illégaux – ou même « préjudiciables ». Dans le même temps, les entreprises qui utilisent ces outils devraient être soumises à des exigences strictes de transparence et de procédure régulière.

De même, les entreprises devraient bénéficier d'une large exonération de responsabilité pour les recommandations faites par leurs algorithmes, dans les cas où ces algorithmes recommandent un contenu illégal en réponse à un contenu consulté par les utilisateurs. Si les développeurs de systèmes et les codeurs définissent les paramètres dans lesquels les algorithmes fonctionnent, ils ne contrôlent ni ne déterminent le résultat de ces processus automatisés : les algorithmes produisent des résultats à partir d'ensembles de données de manière à la fois complexe et imprévisible. Ils sont également tous deux généralement enclins à faire des erreurs et incapables de distinguer un contenu licite d'un contenu illégal. Tenir les sociétés responsables de toutes les « erreurs » commises par leurs systèmes serait donc à la fois irréalisable et disproportionné. Dans la mesure où la responsabilité porte sur des cas spécifiques d'illégalité, c'est aussi un piètre instrument pour relever les défis systématiques soulevés par les algorithmes.

Au lieu de cela, les entreprises – en particulier celles qui sont dominantes sur le marché – devraient être soumises à des obligations de transparence accrues et tenues de réaliser des évaluations d'impact sur les droits de l'homme, comme indiqué ci-dessous. À notre avis, le même raisonnement devrait s'appliquer aux services de navigation ou de « trouvabilité », c'est-à-dire qu'ils ne devraient pas être pénalisés si l'algorithme

de leur moteur de recherche renvoie un contenu illégal, mais ils devraient être transparents et expliquer au public comment leur algorithme sélectionne des résultats de recherche.

En revanche, nous acceptons que les entreprises perdent l'exonération de responsabilité lorsqu'elles « promeuvent » – ou « optimisent » la présentation – des contenus illégaux dans la section publicité de leur plateforme à la suite d'accords commerciaux.¹⁵⁹

Recommandation 4 :

La surveillance générale du contenu doit continuer à être interdite

ARTICLE 19 estime que les gouvernements doivent s'abstenir d'imposer une obligation de surveillance générale des contenus par les entreprises.

Si l'on peut affirmer que la surveillance permet simplement aux entreprises de détecter des contenus potentiellement illégaux ou autres contenus problématiques, dans la pratique, la simple détection est presque toujours associée à la suppression ou à d'autres types d'actions qui réduisent la disponibilité de ces contenus. Cela est profondément problématique étant donné que la technologie de surveillance du contenu n'est pas aussi perfectionnée qu'on le suggère parfois. En particulier, les algorithmes de correspondance de hachage et les outils de traitement du langage naturel sont actuellement incapables de distinguer les contenus dont la légalité pourrait varier selon le contexte, comme les reportages d'actualité ou la parodie.¹⁶⁰ De grandes quantités de contenu légitime peuvent donc être supprimées. De plus, ces technologies portent atteinte au droit à la vie privée des utilisateurs, car elles nécessitent une analyse des communications des individus.

En outre, si une loi subordonnait l'exonération de responsabilité à une « surveillance générale » ou à l'adoption de « mesures proactives » ou d'« efforts améliorés » pour lutter contre les contenus illégaux,¹⁶¹ les entreprises pécheraient inévitablement par excès de prudence et supprimeraient des contenus par défaut pour éviter les risques juridiques et les frais d'exécution. Cela pourrait conduire les plateformes à n'autoriser que des intervenants préselectionnés, ou à utiliser leurs conditions de service pour interdire le contenu controversé,¹⁶² et cela pourrait également dissuader les nouveaux entrants sur le marché de défier les opérateurs en place.¹⁶³

Dans le même temps, ARTICLE 19 reconnaît que la surveillance « spécifique » et la suppression de vidéos ou autres images contenant des contenus incontestablement illicites d'abus sexuels d'enfants – c'est-à-dire la représentation d'une activité sexuelle (par ex., la pénétration) entre un adulte et un enfant – peuvent être compatibles avec les droits à la liberté d'expression et à la vie privée.¹⁶⁴ Nous le reconnaissons parce que nous tenons compte de la gravité du comportement en cause et du fait que ce type de contenu peut être reconnu de manière fiable comme illégal, quel que soit le contexte. Nous ne sommes toutefois pas d'accord que de telles obligations de surveillance spécifique s'appliquent à tout autre type de contenu.¹⁶⁵

Recommandation 5 :

Tout cadre réglementaire doit avoir une portée strictement limitée

Comme nous l'avons noté précédemment, ARTICLE 19 estime que tout cadre visant à réglementer les activités de modération de contenu des plateformes devrait avoir un champ d'application limité, y compris par référence à son objet, aux entités qu'il vise à couvrir et à son application géographique. En particulier, nous formulons les recommandations suivantes :

La réglementation devrait se concentrer sur les contenus illégaux plutôt que sur les contenus « légaux, mais préjudiciables »

ARTICLE 19 estime que tout cadre de ce type devrait être limité aux contenus « illégaux » plutôt que « préjudiciables », pour la simple raison que le contenu « préjudiciable » est un concept intrinsèquement vague. Cela le rend difficile à appliquer, sujet à des abus et susceptible d'être contesté pour des raisons de légalité. À notre avis, le contenu légal, mais néanmoins interdit par les normes communautaires des entreprises, devrait être soumis à une surveillance par des entités multipartites indépendantes, telles que les Conseils de médias sociaux proposés par ARTICLE 19.

Si le contenu « légal, mais préjudiciable » est inclus dans le champ d'application de la législation, contrairement à nos recommandations, la législation ne devrait alors imposer que des exigences de transparence et de procédure régulière pour l'application des normes communautaires de l'entreprise. Le rôle de l'organisme réglementaire se limiterait donc à s'assurer que les systèmes de modération de contenu des entreprises sont suffisamment transparents et que les utilisateurs disposent de mécanismes de recours clairs et efficaces.

Les services de messagerie privée et les organes de presse devraient être exclus du champ d'application

De même, nous pensons que le champ d'application de tout cadre réglementaire devrait être limité afin d'exclure les commentaires en dessous de la ligne sur les sites Web et les blogs de journaux. Idem pour les contenus journalistiques, y compris lorsque des plateformes de médias sociaux ou d'autres services (par e. des moteurs de recherche) l'hébergent. À notre avis, il n'est absolument pas souhaitable que des mesures visant à réglementer les pratiques des très grandes plateformes de médias sociaux servent de prétexte pour réglementer le contenu journalistique.

De même, les applications de messagerie et autres canaux de communication privés doivent être exclues du champ d'application. En particulier, les régulateurs ne devraient pas avoir le pouvoir d'imposer des obligations aux fournisseurs lorsque ces obligations pourraient entraîner une atteinte injustifiable au droit à la vie privée des internautes, notamment un affaiblissement du chiffrement de bout en bout ou des filtres obligatoires.

Les mesures ne doivent pas avoir une application extraterritoriale

Enfin, nous pensons que la mise en œuvre de mesures en vertu de ce nouveau cadre réglementaire devrait être géographiquement limitée au pays qui impose ces mesures, conformément aux principes internationaux de courtoisie et au principe de proportionnalité prévu par le droit international des droits humains. Autrement dit, aucun pays ne devrait être capable d'émettre des ordres de suppression ou de restriction d'un contenu qui pourrait être légal en dehors de ses frontières.

Recommandation 6 :

Les obligations prévues dans tout régime réglementaire doivent être clairement définies

ARTICLE 19 estime que toute obligation prévue par un nouveau régime réglementaire régissant les activités des plateformes et autres entreprises technologiques doit être clairement définie. Nous exposons ci-dessous les types de mesures qui pourraient être incluses dans un tel cadre et celles qui ne le devraient pas. En particulier, nous croyons qu'un nouveau cadre réglementaire pourrait imposer ce qui suit :

Obligations de transparence

À notre avis, la transparence devrait être une exigence fondamentale dans tout ce que font les entreprises. Elle doit notamment s'appliquer dans les domaines suivants :

- **La distribution de contenu** : Les plateformes de médias sociaux et les entreprises numériques devraient fournir des informations essentielles et expliquer au public comment leurs algorithmes sont utilisés pour présenter, classer, promouvoir ou rétrograder un contenu. Le contenu promu doit être clairement identifié comme tel, soit-il promu par l'entreprise ou par un tiers contre rémunération.¹⁶⁶ Les entreprises doivent également expliquer comment elles ciblent les utilisateurs avec du contenu promu (non sollicité), que ce soit de leur propre initiative ou pour le compte de tiers dans le cadre d'un service payant.¹⁶⁷
- **Conditions de service et normes communautaires des entreprises** : Les entreprises devraient publier des normes communautaires/conditions de services faciles à comprendre et donner des exemples de « jurisprudence » sur la manière dont elles sont appliquées. Elles devraient publier des informations sur les méthodes et les processus internes pour l'élaboration des règles communautaires, qui devraient continuer à inclure des consultations avec un large éventail d'acteurs, y compris la société civile.¹⁶⁸

- **Ressources humaines et technologiques utilisées pour assurer la conformité** : Les entreprises doivent inclure des informations détaillées sur les systèmes de signalement de confiance, notamment les personnes qui figurent sur la liste des signaleurs de confiance, comment elles ont été sélectionnées et tous les « privilèges » attachés à ce statut. Elles devraient également publier des informations sur la manière dont leurs algorithmes fonctionnent pour détecter des contenus illégaux ou prétendument « préjudiciables » selon leurs normes communautaires. En particulier, cela devrait inclure des informations sur les taux de faux négatifs/faux positifs et des indicateurs, le cas échéant, pour évaluer le contenu susceptible de devenir viral, par ex. par référence à l'exposition à un public plus large.¹⁶⁹
- **Prise de décision** : Les entreprises doivent informer les parties affectées de leurs décisions et donner des raisons suffisamment détaillées sur les mesures prises contre des contenus ou des comptes particuliers. Elles doivent également fournir des informations précises sur tout mécanisme de plainte interne.
- **Rapports de transparence** : Les entreprises devraient publier des informations détaillées conformes aux Principes de Santa Clara.¹⁷⁰ Nous notons qu'il est particulièrement important de ne pas limiter les informations statistiques à la suppression de contenu, et d'inclure également des données sur le nombre d'appels traités et leur résultat. Les rapports sur la transparence devraient également faire la distinction entre les contenus signalés par des tiers (y compris s'il s'agit d'organismes publics ou privés), les signaleurs de confiance (s'il s'agit d'organismes publics ou d'entités privées), et les algorithmes. Des informations supplémentaires devraient être fournies également sur les différents types de restrictions appliquées au contenu dans le cadre des processus de modération de contenu, telles que la démonétisation ou la rétrogradation ; pour chaque restriction, l'entreprise doit fournir des informations sur les règles sur lesquelles la décision a été fondée et, le cas échéant, sur le résultat des appels.

Plus généralement, nous notons que toute exigence de rapport sur la transparence devrait viser à fournir une analyse beaucoup plus qualitative des décisions de modération de contenu. La mesure du succès dans la lutte contre le contenu illégal ne doit pas être liée aux taux de suppression de contenu, car cela encourage la suppression excessive. De même, les rapports de transparence ne devraient pas se limiter aux informations fournies par les entreprises, et devraient également inclure des informations transmises par les agences gouvernementales compétentes. Ce qui précède est sans préjudice des mesures éventuellement applicables en vertu du droit des consommateurs.¹⁷¹

- **Audits de transparence** : Les entreprises devraient donner un meilleur accès aux ensembles de données aux organismes de régulation et aux chercheurs indépendants approuvés – qu'ils soient universitaires, journalistes ou autres – pour leur permettre de vérifier que les systèmes et les algorithmes de l'entreprise fonctionnent comme celle-ci le prétend. En particulier, les auditeurs devraient avoir accès à des données sur : (1) les programmes de modération de contenu des entreprises ; (2) la manière dont les entreprises classent, rangent, hiérarchisent, recommandent ou personnalisent le contenu ; et (3) la manière dont cela s'applique à la réclame politique.¹⁷² Alors que les organismes de régulation pourraient avoir accès à des données sensibles et commerciales, des tiers approuvés pourraient avoir accès à des ensembles de données anonymisés. Ces audits des opérations des plateformes devraient avoir lieu régulièrement.
- **Archives de la publicité numérique et politique** : De nouvelles règles devraient interdire « l'opacité des données » pour des publicités politiques et demander aux plateformes des mécanismes d'application clairs en cas de violation de leurs politiques. Les plateformes ne devraient pas être tenues de jouer un rôle d'arbitrage ou d'atténuation des discours politiques agressifs et de la « désinformation ».

De plus, nous appelons à une transparence renforcée sur toutes les dépenses publicitaires politiques des parties prenantes concernées, y compris les partis politiques, les entreprises technologiques et les annonceurs tiers. Les publicités politiques devraient être clairement distinctes du contenu éditorial, y compris les actualités, quelle que soit leur forme et y compris en ligne, et clairement cataloguées et accompagnées d'informations sur qui les a payées. En outre, nous soutenons l'utilisation de bases de données publicitaires numériques pour conserver et publier toutes les publicités réglementées, le montant pécuniaire dépensé en publicité, et le nom de la personne qui a autorisé la publicité. Ces bases de données devraient être accessibles dans un format permettant une récupération en masse par des chercheurs et des décideurs politiques.

Obligations internes de procédure régulière

Sans préjudice de la Recommandation 9, ARTICLE 19 estime que tout cadre réglementaire régissant les activités des plateformes dominantes devrait inclure une obligation de mettre en place :

- **Des règles de notification et d'action précises**, conformément aux Principes de Manille sur la responsabilité des intermédiaires.
- **Des mécanismes de recours internes** pour gérer les plaintes liées aux restrictions de l'exercice du droit à la liberté d'expression, telles que la suppression illicite de contenu ou l'application abusive d'étiquettes pouvant suggérer qu'une source d'information n'est pas digne de confiance. À l'inverse, les mécanismes d'appel devraient également être en mesure de répondre au refus d'une entreprise de supprimer un contenu qui est supposé enfreindre les normes communautaires de l'entreprise. Dans tous les cas, les mécanismes de plainte internes devraient respecter les garanties de procédure régulière.¹⁷³

- **Obligation de promouvoir la diversité des médias** : Compte tenu des risques de contenu trop personnalisé sur les plateformes de médias sociaux, les très grandes entreprises de médias sociaux devraient être tenues de prendre des mesures afin de garantir un degré suffisant de diversité des médias,¹⁷⁴ ainsi qu'une couverture équilibrée pendant les élections, sur leur service.¹⁷⁵ En particulier, elles devraient fournir suffisamment d'informations pour expliquer comment les fils d'actualité et le contenu qu'ils promeuvent sont sélectionnés. Elles devraient également fournir aux utilisateurs des alternatives viables parmi lesquelles choisir qui ne sont pas basées sur le profilage. Dans certaines circonstances limitées, comme pendant des élections, elles pourraient être contraintes de diffuser des messages provenant des médias du service public pour veiller à ce que le segment le plus large possible de la population dispose des informations de base nécessaires pour participer de manière significative aux élections.

S'abstenir d'imposer certaines obligations

En revanche, ARTICLE 19 estime qu'un tel système réglementaire ne devrait pas inclure les types non exhaustifs d'obligations suivants :

- **Un « devoir de diligence » étendu et non défini pour prévenir une notion tout aussi indéfinie de « préjudice »** : À notre avis, de telles notions auraient peu de chances de passer le test de légalité prévu par le droit international des droits humains. En réalité, il créerait à la fois une incertitude juridique et conférerait des pouvoirs largement illimités aux autorités de régulation, ce qui serait profondément problématique pour la liberté d'expression.

- **Une obligation générale de surveillance du contenu** : Comme indiqué ci-dessus, il ne devrait pas y avoir d'obligation de surveillance générale du contenu, ni de mesures substantiellement équivalentes, telles que l'obligation d'« efforts améliorés » ou de « mesures proactives » pour lutter contre le contenu illégal. De même, un tel cadre devrait s'abstenir de « pousser » les entreprises vers l'adoption de telles mesures en les présentant comme purement volontaires ou simplement « recommandées », alors qu'en réalité, le fait de ne pas les adopter pourrait entraîner de lourdes sanctions.
- **Délais trop courts** : Les entreprises ne devraient pas être tenues de supprimer des contenus dans des délais trop courts, en particulier lorsque le contenu en cause peut donner lieu à des questions d'interprétation difficiles telles que le « discours de haine » ou le contenu « terroriste ». Des délais trop courts n'incitent pas les entreprises à examiner les avis avec suffisamment de soin. En tant que tels, ils encouragent la suppression illicite de contenus et ne protègent pas la liberté d'expression. De plus, les suppressions dans un court laps de temps peuvent inciter les entreprises à allouer des ressources à la suppression des avis quelle que soit leur gravité ou à se concentrer sur un contenu simplement parce qu'il a été publié au cours des dernières 24 heures, au détriment d'un contenu plus ancien qui pourrait mériter plus d'attention.¹⁷⁶
- **Définition d'objectifs de conformité** : De même, la législation ou les organismes de régulation ne devraient pas imposer des objectifs de conformité chiffrés qui pourraient avoir pour effet d'inciter les entreprises à élargir la définition des contenus interdits sur leurs plateformes afin d'augmenter leur taux de conformité.¹⁷⁷ Autrement dit, des objectifs chiffrés pourraient encourager la suppression de quantités toujours plus importantes de contenus légitimes. Nous notons en outre que, dans la mesure où les législateurs pourraient envisager divers

paramètres et seuils pour assurer la conformité, ils devraient examiner dans quelle mesure la société est prête à tolérer un degré de risque de préjudice en ligne, comme elle le fait dans l'univers hors ligne.¹⁷⁸

- **Obligation de coopérer ou de signaler un contenu illicite :**
Les vagues obligations de coopérer sont problématiques, car elles pourraient impliquer de graves atteintes aux droits des utilisateurs, telles que l'accès aux données personnelles des utilisateurs par les forces de l'ordre, sans garanties suffisantes. Dans le même temps, leur manque de précision permet aux entreprises de soutenir qu'elles ont coopéré de manière moins intrusive. En bref, de telles obligations risquent d'être difficiles à appliquer ; par conséquent, on ne peut pas être sûrs qu'elles soient nécessaires. Les obligations de signaler des contenus illégaux pourraient inciter fortement les entreprises à se concentrer sur les avis qu'elles reçoivent sur des contenus prétendument illégaux, sans tenir compte de leur gravité, et à les signaler aux forces de l'ordre. Elles pourraient également dissuader les entreprises d'investir dans des outils automatisés pour détecter des contenus illégaux si elles pouvaient déterminer une reconnaissance d'illégalité ou être en infraction avec leurs obligations, pour avoir omis de signaler tous les contenus potentiellement illégaux identifiés automatiquement sur leurs réseaux. Ces deux résultats seraient indésirables et auraient probablement un impact négatif sur la liberté d'expression, dans la mesure où de grandes quantités de contenu légitime seraient signalées.
- **Obligations de diffuser (must-carry) :** Même si nous estimons que des « obligations de diffusion (« must-carry obligations») – c'est-à-dire l'obligation pour les plateformes de publier un contenu licite qui enfreint par ailleurs leurs conditions d'utilisation – peuvent être imposées dans des circonstances très limitées, nous pensons qu'il serait inapproprié d'imposer plus d'obligations de diffusion générales. En particulier, les

obligations de diffusion pourraient porter atteinte à la fois au droit à la liberté d'expression et au droit de propriété des plateformes de médias sociaux.¹⁷⁹ Le droit à la libre expression des plateformes doit impliquer qu'elles devraient être libres d'autoriser ou d'interdire par ex. des théories complotistes d'extrême droite sur leur site, même si ces théories peuvent être considérées légales dans certains pays.

En termes de proportionnalité, les utilisateurs qui souhaitent promouvoir de telles théories pourraient toujours avoir d'autres canaux à leur disposition pour exprimer ces opinions. Dans la pratique, des obligations générales de diffusion pourraient aussi compromettre quasi certainement la modération du contenu, ce qui serait indésirable.

Recommandation 7 :

Tout régulateur doit être indépendant et responsable à la fois devant la loi et dans la pratique

ARTICLE 19 estime que, pour qu'un système de régulation du contenu en ligne ait une quelconque légitimité, il doit être supervisé par un organisme de régulation indépendant ; c'est-à-dire un organisme libre de toute ingérence politique ou commerciale.

L'indépendance et l'autonomie institutionnelle d'un organisme de régulation doivent être garanties et protégées par la loi, notamment par :

- Un énoncé clair de la politique générale de la plateforme et de la réglementation du contenu en ligne ;
- Une définition précise des pouvoirs et des responsabilités de l'autorité de régulation ;

- Des règles d'adhésion ; et
- Des modalités de financement et de reddition de comptes au public à travers un organisme multipartite. Le gouvernement devrait être tenu à distance et ne pas être impliqué dans l'un ou l'autre de ces accords de financement et organes de responsabilisation.

Plus précisément, pour que la liberté d'expression soit protégée par un régulateur, la loi qui l'institue doit contenir :

- Des dispositions générales soulignant l'importance de la protection de la liberté d'expression, y compris l'expression qui peut choquer, offenser ou déranger ;
- Des dispositions précisant que la mission de toute autorité de régulation dans ce domaine consiste à protéger les droits de l'homme, y compris la liberté d'expression ;
- Des dispositions obligeant toute autorité de régulation à auditer les décisions de suppression de contenu et à déterminer la mesure dans laquelle les entreprises suppriment ou restreignent excessivement des contenus, que ce soit sur demande ou de leur propre chef ; et
- Des dispositions précisant que les entreprises ne devraient pas être pénalisées pour avoir omis de supprimer un contenu illicite.

Tout régulateur chargé de superviser les opérations d'un large éventail de fournisseurs de services numériques devrait garantir une coopération avec d'autres régulateurs concernés, tels que les autorités chargées de la protection des données, de la protection des consommateurs et de la concurrence.

Enfin, tout régulateur doit être en mesure de rendre des comptes au public, notamment par des obligations de transparence sur ses activités et des rapports annuels présentés au parlement.

Recommandation 8 :

Tout cadre réglementaire doit être proportionné

ARTICLE 19 estime que, pour être conforme aux normes internationales en matière de liberté d'expression, tout cadre réglementaire doit être strictement proportionné au but poursuivi :

- **Approche à plusieurs niveaux** : Les gouvernements doivent être extrêmement prudents sur l'adoption de mesures visant à obliger les grandes entreprises de médias sociaux à rendre des comptes, mais qui finiraient par imposer un fardeau excessif à d'autres services plus petits. À ce titre, nous estimons qu'une approche à plusieurs niveaux dans ce domaine serait nécessaire. En d'autres termes, les grandes plateformes de médias sociaux pourraient être soumises à des obligations plus strictes que les petits acteurs. Pour évaluer la dominance d'une plateforme, il faudrait tenir compte des facteurs suivants : (1) le nombre de ses utilisateurs ; (2) son chiffre d'affaires mondial annuel ; et (3) son pouvoir de marché. Les organisations à but non lucratif comme Wikipedia devraient être exemptées et continuer à fonctionner dans un large cadre d'exonération de responsabilité.

Dans le même temps, nous recommandons fortement que toute mesure proposée fasse l'objet d'évaluations d'impact rigoureuses, y compris d'éventuelles conséquences anticoncurrentielles. Les grandes entreprises de médias sociaux sont susceptibles de s'adapter à toutes les demandes qui leur sont imposées. De telles demandes pourraient finalement donner l'impression qu'elles sont « plus sûres ». Cela leur apporterait quasi certainement un avantage sur les petits entrants, qui ne seraient pas capables de s'engager dans le même type d'exercice de modération de contenu que les opérateurs en place.

- **Évaluation des défaillances systémiques** : Pour que la réglementation dans ce domaine soit viable et proportionnée, les entreprises ne doivent pas être évaluées parce qu'elles ont omis de supprimer un seul contenu ou parce qu'elles n'ont publié qu'un seul ensemble de données. Un organisme de régulation devrait évaluer si l'entreprise a manqué à ses obligations en vertu de la loi sur une base systémique. Le seuil des défaillances systémiques devrait être défini par la loi par référence à des critères précis permettant une évaluation globale plutôt que purement numérique. Par exemple, la loi ne devrait pas sanctionner des entreprises parce qu'elles ont omis de supprimer une quantité ou un pourcentage donné de contenu signalé soit comme illégal soit comme préjudiciable. Elle devrait plutôt contenir une évaluation globale des mesures que les entreprises ont adoptées pour atténuer les risques pour les droits humains.
- **Sanctions proportionnées** : Le non-respect des obligations décrites ci-dessus doit être passible de sanctions proportionnées. Cela peut inclure des amendes importantes, mais celles-ci ne doivent pas être trop élevées et dissuader de protéger la liberté d'expression. Nous estimons qu'un pourcentage de 4 % du chiffre d'affaires mondial est probablement trop élevé. De même, les sanctions pénales imposées aux chefs exécutifs pour non-respect de ces obligations pourraient avoir un effet dissuasif sur la liberté d'expression. Confrontés à la perspective de plusieurs années de prison, les dirigeants d'entreprise adopteraient quasi certainement des politiques favorisant une plus grande suppression de contenu ou d'autres types de restrictions de contenu. De ce fait, les gouvernements devraient s'abstenir d'adopter ce type de sanctions pénales.

Recommandation 9 :

Tout cadre réglementaire doit fournir un accès à des recours effectifs

Au-delà des mécanismes de plainte internes, ARTICLE 19 estime que les gouvernements devraient veiller à ce que les individus puissent accéder à des recours judiciaires pour contester le retrait illicite de leur contenu par les plateformes de médias sociaux sur la base de leurs conditions d'utilisation. Ces recours devraient inclure non seulement un accès aux tribunaux, mais également des mécanismes alternatifs de règlement des litiges, tels que des tribunaux en ligne ou un médiateur.¹⁸⁰ Dans la pratique, les gouvernements devraient élaborer des propositions pour financer de tels mécanismes, y compris, par ex. un prélèvement sur les plateformes de médias sociaux.

Cela devrait être sans préjudice des systèmes d'autorégulation, tels que les Conseils de médias sociaux, qui permettraient (entre autres) aux utilisateurs de contester les décisions de modération de contenu des plateformes de médias sociaux par référence à un ensemble de principes convenus tels qu'une « Charte des droits des utilisateurs ».

Recommandation 10 :

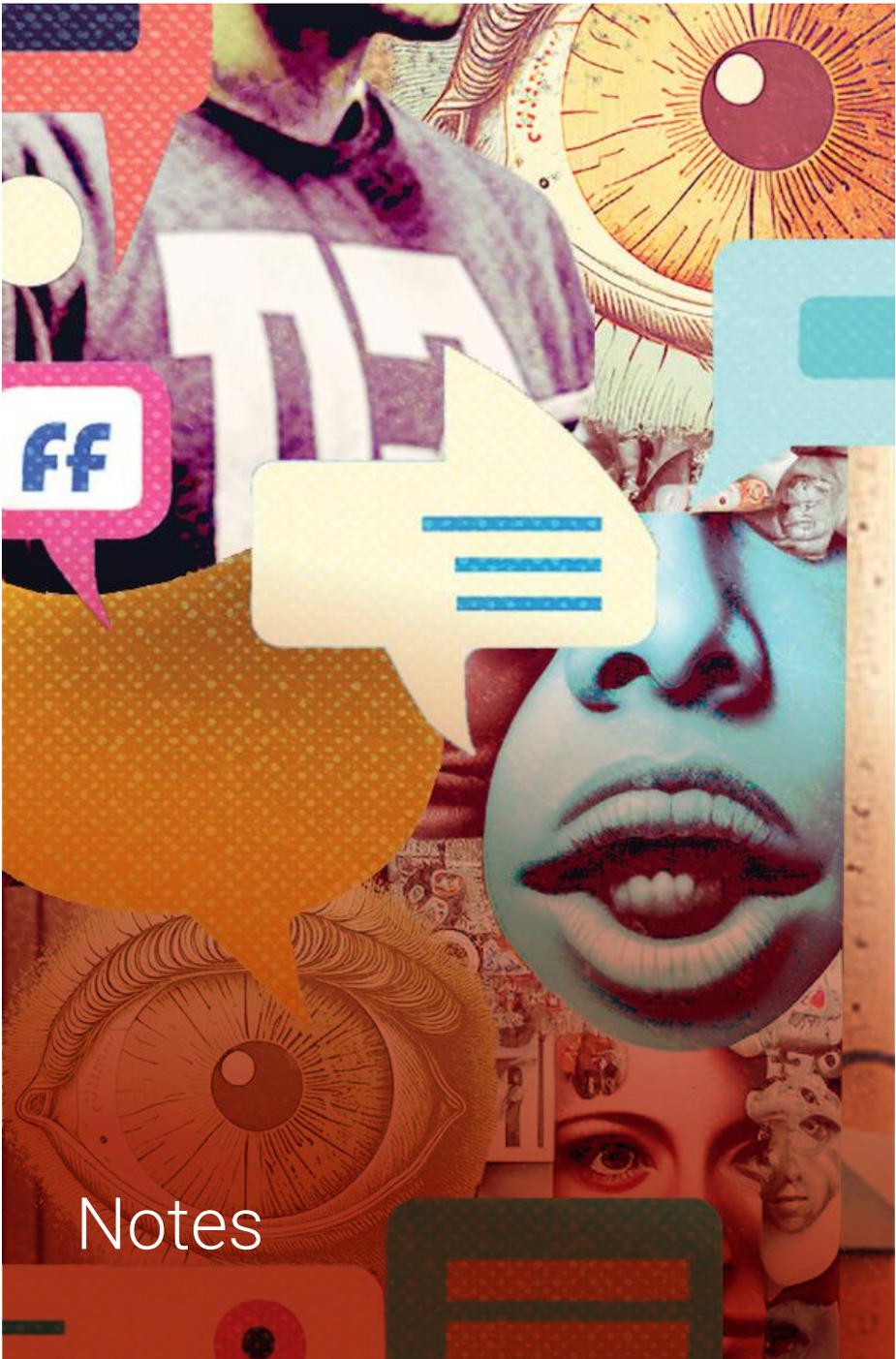
Les grandes plateformes devraient être tenues de dégroupier leurs fonctions d'hébergement et de curation de contenu et de garantir leur interopérabilité avec d'autres services

ARTICLE 19 estime que pour faire face au pouvoir de marché excessif d'une poignée de plateformes de médias sociaux, la curation de contenu devrait être décentralisée.

Les régulateurs pourraient contraindre les grandes plateformes à dissocier leurs fonctions d'hébergement et de curation de contenu afin de permettre à des tiers d'accéder à leur plateforme (en pratique, leur API) pour fournir des services de curation de contenu aux usagers.¹⁸¹

Cette forme de séparation fonctionnelle n'empêcherait pas les grandes plateformes d'offrir une curation de contenu à leurs utilisateurs. Cependant, les utilisateurs pourraient décider librement s'ils veulent y adhérer. En d'autres termes, au moment de créer son profil sur Facebook, par exemple, l'utilisateur devrait être invité à sélectionner un fournisseur de curation de contenu, et Facebook pourrait rester l'une des options à sélectionner. Idéalement, et pour éviter de rester bloqués sur cette option initiale, les utilisateurs devraient rester libres de changer d'option à tout moment via les paramètres de la plateforme. Nous pensons que ces types de solutions devraient être davantage explorés pour permettre aux utilisateurs de reprendre la main, de garantir une concurrence saine et l'innovation sur les marchés de médias sociaux et de renouer avec la promesse d'un Internet diversifié et décentralisé.¹⁸²

À tout le moins, nous pensons que les plateformes de médias sociaux devraient offrir aux usagers plus d'options de curation de contenu, à la fois en termes de type de contenu qu'ils souhaiteraient recevoir et selon quels critères, par ex. par ordre chronologique.



Notes

Notes

(Cliquez sur le numéro de la note de fin pour lire dans son contexte)

- 1  [John Perry Barlow, A Declaration of the Independence of Cyberspace](#), Electronic Frontier Foundation, 8 février 1996.Y
- 2  Les « hébergeurs » sont généralement ceux qui louent un espace de serveur web pour permettre à leurs clients de créer leurs propres sites Internet. Le terme d'hébergeur a pris également un sens plus général, et comprend toute personne ou société contrôlant un site ou une page Web qui permet à des tiers de télécharger ou publier du contenu. De ce fait, les plateformes de médias sociaux et les services de partage de vidéos et de photos sont généralement qualifiés d'« hébergeurs ». « L'hébergement » est donc un rôle caractérisé par l'absence d'intervention éditoriale dans le contenu » ; c.f. ARTICLE 19, [Les intermédiaires d'Internet : dilemme de la responsabilité](#), 2013, p.6.
- 3  Selon des estimations récentes, on compte aujourd'hui 2,5 milliards d'utilisateurs mensuels actifs sur Facebook et 330 millions d'utilisateurs de Twitter, et 1 milliard d'heures de vidéos sont visionnées chaque jour sur YouTube. Voir par ex. le Portail des statistiques, [Leading countries based on Facebook audience size](#); Portail des statistiques, [Number of monthly active Twitter users worldwide from 1st quarter 2010 to 1st quarter 2019](#); [YouTube for Press](#), [YouTube in numbers](#).
- 4  Voir par ex. Twitter, [Permanent suspension of @realDonaldTrump](#), 8 janvier 2021; Guy Rosen Monika Bickert, [Our response to the violence in Washington](#), Facebook, 6 janvier 2021.
- 5  K. Wagner et P. Martin, [Twitter locks out Chinese Embassy in U.S. over post on Uighurs](#), Bloomberg, 20 janvier 2021.
- 6  Voir par ex. Section 230 of the Communications Decency Act 1996 (47 USC. § 230) aux États-Unis, ou la [Directive 2000/31/CE du Parlement européen et du Conseil du 8 juin 2000 relative à certains aspects juridiques des services de la société de l'information, et notamment du commerce électronique, dans le marché intérieur](#) (« directive sur le commerce électronique »).
- 7  Voir par ex. [Online Harms White Paper](#) au Royaume-Uni, 2019, et [Draft Online Safety Bill](#), 2021; [Proposition de règlement du Parlement européen et du Conseil relatif à un marché intérieur des services numériques \(législation sur les services numériques\) et modifiant la Directive 2000/31/CE](#), Union européenne, 2020; [Online Safety and Media Regulation Bill](#), Irlande, 2020; [l'Australian Online Safety Bill](#), 2021; [Draft Intermediaries Guidelines \(Amendment\) Rules](#), 2018, Inde.
- 8  Voir ARTICLE 19, [#MissingVoices](#) campaign.
- 9  Voir par ex. [Online Harms White Paper: Full government response to the consultation](#), 15 décembre 2020; ou la proposition de loi française [visant à lutter contre les contenus haineux sur Internet](#) (dite Loi Avia, ou proposition de loi Avia), que le Conseil Constitutionnel a déclaré inconstitutionnelle par la suite.

- 10 ↩ Dilemme de la responsabilité, *op.cit.*
- 11 ↩ ARTICLE 19, Sidestepping rights: Regulating speech by contract, 2018.
- 12 ↩ *Ibid.*
- 13 ↩ Cette note d'orientation fait également référence à plusieurs propositions législatives de gouvernance des plateformes présentées récemment dans certains pays, qui peuvent servir d'inspiration pour l'adoption d'une législation similaire, en particulier la Draft Online Safety Bill, 2021, et la proposition de législation de l'Union européenne sur les services numériques (voir Commission européenne, Paquet législatif sur les services numériques ou Digital Services Act package). Nous comprenons que ces propositions pourraient avoir été modifiées ou changées entre-temps, et cette note d'orientation les cite sous leur forme au moment de la publication (Décembre 2021).
- 14 ↩ ARTICLE 19, Taming Big Tech: Protecting freedom of expression through the unbundling of services, open markets, competition, and users' empowerment.
- 15 ↩ ARTICLE 19, note d'orientation sur la diversité des médias en ligne, à paraître.
- 16 ↩ ARTICLE 19, note d'orientation sur les obligations de must-carry, à paraître.
- 17 ↩ Cela comprend, par exemple, les questions relatives à la publicité, aux publicités ciblées, et aux modèles commerciaux des entreprises et à la liberté d'expression et d'information.
- 18 ↩ Dilemme de la responsabilité, *op.cit.*
- 19 ↩ Voir par ex. L. Lovdahl Gormsen et J.T. Llanos, Facebook's anticompetitive lean in strategies, SSRN, 6 juin 2019.
- 20 ↩ C.f. par ex. Communication de la Commission au Parlement européen, au Conseil, au Comité économique et social européen et au Comité des régions, Les plateformes en ligne et le marché numérique – Perspectives et défis pour l'Europe, COM (2016) 288 final, 25 mai 2016.
- 21 ↩ Voir Sidestepping rights, *op.cit.*, p. 36.
- 22 ↩ Voir par ex. Dilemme de la responsabilité, *op.cit.*
- 23 ↩ En droit commun, voir par ex. D. Rolph, Liability in the Internet age, Blog du International Forum for Responsible Media (IFFRM), 17 février 2011.
- 24 ↩ À travers son adoption dans une résolution de l'Assemblée générale des Nations Unies, la DUDH n'est pas strictement contraignante pour les États. Cependant, un certain nombre de ses dispositions sont considérées comme ayant acquis une force légale en tant que droit international coutumier depuis son adoption en 1948; voir Filartiga c. Pena-Irala, 630 F. 2d 876 (1980) (US Circuit Court of Appeals, 2nd circuit).

- 25 ↩ Assemblée générale des Nations Unies, Pacte international relatif aux droits civils et politiques (PIDCP), 16 décembre 1966, Collection des traités des Nations Unies, vol. 999, p.171.
- 26 ↩ Article 10 de la Convention européenne pour la protection des droits humains et des libertés fondamentales, 4 septembre 1950; Article 9 de la Charte africaine des droits de l'homme et des peuples (Charte de Banjul), 27 juin 1981; Article 13 de la Convention américaine des droits de l'homme, 22 novembre 1969.
- 27 ↩ Comité des droits de l'homme (Comité DH), Observation générale n° 34 sur l'Article 19: Libertés d'opinion et liberté d'expression, CCPR/C/GC/34, 12 septembre 2011, par. 12, 17, et 39.
- 28 ↩ **La Déclaration conjointe sur la liberté d'expression et l'Internet de 2011, adoptée par le Rapporteur spécial des Nations Unies sur la promotion et la protection du droit à la liberté d'opinion et d'expression, le Représentant de l'Organisation pour la sécurité et la coopération en Europe (OSCE) pour la liberté des médias, le Rapporteur spécial de l'Organisation des États américains (OEA) pour la liberté d'expression et le Rapporteur spécial sur la liberté d'expression et l'accès à l'information de la Commission africaine des droits de l'homme et des peuples (CADHP, juin 2011.**
- 29 ↩ *Ibid.* Voir également le Rapport du Rapporteur spécial sur la liberté d'expression, A/66/290, 10 août 2011, par. 16.
- 30 ↩ Comité DH, *Belichkin c. Biélorussie*, Comm. No. 1022/2001, UN Doc. CCPR/C/85/D/1022/2001 (2005).
- 31 ↩ Observation générale n° 34, *op.cit.*, par. 43.
- 32 ↩ Déclaration conjointe 2011, *op. cit.*
- 33 ↩ Rapporteur spécial sur la liberté d'expression, Rapport du 16 mai 2011, A/HRC/17/27, par. 43.
- 34 ↩ Rapporteur spécial sur la liberté d'expression, Rapport du 6 avril 2018, A/HRC/38/35, par. 66.
- 35 ↩ *Ibid.*
- 36 ↩ *Ibid.*
- 37 ↩ Rapporteur spécial sur la liberté d'expression, Rapport du 13 avril 2021, A/HRC/47/25, par. 91.

- 38  Principes directeurs relatifs aux entreprises et aux droits de l'homme, mise en œuvre du cadre de référence « protéger, respecter et réparer » des Nations Unies (les Principes directeurs), élaborés par le Représentant spécial du Secrétaire général chargé de la question des droits de l'homme et des sociétés transnationales et autres entreprises, Rapport du Représentant spécial du Secrétaire général, John Ruggie, 7 avril 2008, A/HRC/8/5A/HRC/17/31. Le Conseil des droits de l'homme a approuvé les Principes directeurs dans sa Résolution 17/4 du 16 juin 2011.
- 39  *Ibid.*, Principe 15.
- 40  *Ibid.*, Rapport du Rapporteur spécial sur la liberté d'expression, 6 avril 2018, A/HRC/38/35, par. 48 et 7; Rapporteur spécial de l'OEA pour la liberté d'expression, Standards for a free, open and inclusive Internet, 15 mars 2017, par. 111-112.
- 41  *Ibid.*, Rapport 2018 du Rapporteur spécial sur la liberté d'expression, par. 45-46; Rapport du Rapporteur spécial de l'OEA, par.113. Voir également Rapporteur spécial de l'OEA pour la liberté d'expression, Standards for a free, open and inclusive Internet, 15 mars 2017, par. 99.
- 42  *Ibid.*, Rapporteur spécial sur la liberté d'expression, par. 47 et 76; Rapporteur spécial de l'OEA pour la liberté d'expression 2013, par.115.
- 43  Rapport du Rapporteur spécial sur la liberté d'expression avril 2018, *op. cit.*, par. 40-44.
- 44  *Ibid.*, par. 70-72.
- 45  Commissaire aux droits de l'homme du Conseil de l'Europe, La prééminence du droit sur l'Internet et dans le monde numérique en général, CommDH/IssuePaper (2014) 1, 8 décembre 2014.
- 46  *Ibid.*, p. 24.
- 47  Comité des ministres du Conseil de l'Europe, Recommandation CM/Rec (2012) du Comité des Ministres aux Etats membres sur la protection des droits de l'homme dans le cadre des services de médias sociaux, adoptée par le Comité des ministres le 4 avril 2012. Ces recommandations ont été reprises dans le Comité des ministres, Guide des droits de l'homme pour les utilisateurs d'Internet, Recommandation CM/Rec(2014)6 et exposé des motifs, p.4.
- 48  Recommandation CM/Rec (2018)2 du Comité des ministres aux Etats membres sur les rôles et les responsabilités des intermédiaires d'Internet, adoptée par le Comité des ministres le 7 mars 2018.
- 49  Voir Rapport OEA 2013, *op. cit.* et Rapport de l'OEA 2016, *op.cit.*
- 50  *Ibid.*

- 51  Voir Déclaration de principes sur la liberté d'expression et l'accès à l'information en Afrique, 2019, adoptée par la CADHP lors de sa 65e session ordinaire à Banjul, Gambie.
- 52  Déclaration de principes sur la liberté d'expression et l'accès à l'information en Afrique. Voir également Déclaration africaine des droits et libertés de l'Internet, 2014, élaborée en tant qu'initiative de la société civile.
- 53  Principes de Manille sur la responsabilité des intermédiaires (Principes de Manille), mars 2015.
- 54  *Ibid.*, Principe 4.
- 55  *Ibid.*, Principe 5(c).
- 56  Ranking Digital Rights, Corporate Accountability Index: 2015 Research Indicators, juin 2015.
- 57  Déclaration conjointe sur la liberté d'expression et la lutte contre l'extrémisme violent de 2016, adoptée par le Rapporteur spécial des Nations Unies sur la liberté d'expression, le Représentant de l'OSCE pour la liberté des médias, le Rapporteur spécial de l'OEA pour la liberté d'expression et le Rapporteur spécial sur la liberté d'expression et l'accès à l'information de la CADHP, 4 mai 2016.
- 58  Déclaration conjointe sur la liberté d'expression et les fausses nouvelles (« Fake News »), la désinformation et la propagande de 2017 adoptée par le Rapporteur spécial des Nations Unies sur la liberté d'expression, le Représentant de l'OSCE pour la liberté des médias, le Rapporteur spécial de l'OEA pour la liberté d'expression et le Rapporteur spécial sur la liberté d'expression et l'accès à l'information de la CADHP, 3 mars 2017.
- 59  Rapporteurs spéciaux des Nations Unies sur la liberté d'expression et la violence contre les femmes, UN experts urge States and companies to address online gender-based abuse but warn against censorship, communiqué de presse conjoint, 8 mars 2017.
- 60  Rapport du Rapporteur spécial sur la liberté d'expression, A/74/486, 9 octobre 2019, par. 57(d).
- 61  Déclaration conjointe 2016, *op.cit.*, par. 2 (e).
- 62  Déclaration conjointe 2017, *op.cit.*, par. 4 (b).
- 63  *Ibid.*, par. 58 (b).
- 64  Rapport du Rapporteur spécial sur la liberté d'expression, A/73/348, 29 août 2018.
- 65  *Ibid.*, par. 64.

- 66 ↩ *Ibid.*
- 67 ↩ *Ibid.*, par. 66.
- 68 ↩ *Ibid.*
- 69 ↩ *Ibid.*, par. 70.
- 70 ↩ Voir Article 17 du PIDCP, *op.cit.* Le Rapporteur spécial sur la promotion et la protection des droits de l'homme et des libertés fondamentales dans la lutte antiterroriste Martin Scheinin a fait valoir que, à l'instar des restrictions au droit à la liberté d'expression en vertu de l'Article 19, les restrictions du droit à la vie privée en vertu de l'Article 17 du PIDCP devraient être comprises comme étant soumises au test en trois parties ; voir le Rapport du Rapporteur spécial des Nations Unies sur la promotion et la protection des droits de l'homme et des libertés fondamentales dans la lutte antiterroriste Martin Scheinin, A/HRC/13/37, 28 décembre 2009.
- 71 ↩ Rapport du Rapporteur spécial sur la liberté d'expression mai 2011, *op.cit.*, par. 53.
- 72 ↩ *Ibid.*, par. 84.
- 73 ↩ Rapport du Rapporteur spécial au Conseil des droits de l'homme sur le recours au chiffrement et à l'anonymat dans l'exercice des droits à la liberté d'opinion et d'expression à l'ère du numérique, A/HRC/29/32, 22 mai 2015, par.60.
- 74 ↩ *Ibid.* Voir également le rapport du Rapporteur spécial sur la liberté d'expression, Encryption and anonymity follow-up report, Document de recherche 1/2018, juin 2018.
- 75 ↩ Par exemple, au troisième trimestre 2020, Facebook a enregistré 2,5 milliards d'utilisateurs mensuels actifs dans le monde – près de la totalité de la population de la Chine et de l'Inde combinées. En 2017, les recettes annuelles de Facebook ont dépassé le PIB de la Serbie; voir F. Belinchón et Q. Moynihan, 25 giant companies that are bigger than entire countries, *Business Insider España*, 25 juillet 2018. De même, en décembre 2019, Google détenait plus de 90% de parts de marché des services de recherche dans le monde ; voir Statcounter, Search engine market share worldwide, avril 2019-avril 2020. Bien que plus petit, Twitter compte toujours plus de 125 millions d'utilisateurs actifs quotidiens, soit environ 2,5 fois la population du Kenya.
- 76 ↩ Voir par ex. le projet de loi du Royaume-Uni sur les préjudices en ligne/la sécurité en ligne (Online Harms/Online Safety Bill), *op. cit.*, le projet de loi Avia désormais invalidé, *op. cit.*, ou le projet de loi autrichien sur les plateformes de communication, octobre 2020.

- 77 ↩ Par exemple, la proposition de loi Avia, déclarée inconstitutionnelle par le Conseil constitutionnel français en juin 2020, prévoyait que le gouvernement fixe des seuils au-delà desquels les entreprises sont censées remplir certaines obligations, telles que le retrait de discours de haine illégaux, dans un délai de 24 heures (voir note 7). De même, en Allemagne, un projet de loi sur la diversité des médias fixe un certain nombre d'obligations sur les « plateformes vidéo » ayant plus d'un million d'utilisateurs par mois en Allemagne ; la Proposition sur les questions d'interface radio, de contrôle des plateformes et des médias intermédiaires, réglementations spécifiques aux États sur l'approbation, la réglementation et les intermédiaires (appelée *Medienstaatsvertrag*), 2018.
- 78 ↩ Par exemple, la Directive européenne de 2019 sur le droit d'auteur prévoit que les fournisseurs de services de partage de contenu en ligne ne seront plus exonérés de responsabilité à moins qu'ils ne remplissent un certain nombre de conditions strictes. Toutefois, le respect de ces nouvelles obligations au titre de la Directive doit tenir compte du « type, de l'audience et de la taille du service », conformément au principe de proportionnalité. Les entreprises qui ont fourni des services pendant moins de trois ans dans l'Union européenne et dont le chiffre d'affaires annuel est inférieur à 10 millions d'euros sont soumises à des conditions moins strictes afin de bénéficier de l'exonération de responsabilité ; voir Directive (EU) 2019/790 du Parlement européen et du Conseil du 17 avril 2019 sur le droit d'auteur et les droits voisins dans le marché unique numérique et modifiant les Directives 96/9/EC et 2001/29/EC. En revanche, la Directive européenne sur les droits d'auteur de 2019 stipule que les fournisseurs de services de partage de contenu en ligne dont le nombre moyen de visiteurs uniques n'excède pas 5 millions sont soumis à des exigences supplémentaires.
- 79 ↩ Voir par ex. Ofcom, Online market failures and harms: An economic perspective on the challenges and opportunities in regulating online services, 28 octobre 2019; Commission australienne pour la concurrence et la consommation, Digital Platforms Inquiry: Final report, 26 juillet 2019; et les autorités de la concurrence française et allemande, Algorithms and competition, rapport conjoint, novembre 2019.
- 80 ↩ Par exemple, Cloudflare a décidé de ne plus fournir ses services à Daily Stormer – un site d'actualité américain d'extrême-droite, néonazi et suprémaciste blanc – en 2017 à la suite d'une attaque mortelle contre des manifestants anti-extrême droite à Charlottesville. En 2019, Cloudflare a de nouveau interrompu la fourniture de ses services à 8chan, un site d'actualité d'extrême droite, après que l'auteur de la fusillade d'El Paso y aurait publié son manifeste ; voir M. Prince, Why we terminated Daily Stormer, Cloudflare, 16 août 2017; M. Prince, Terminating service for 8Chan, Cloudflare, 5 août 2019.
- 81 ↩ Voir par ex. Amnesty International, Surveillance giants: How the business model of Google and Facebook threatens human rights, POL 30/1404/2019, novembre 2019; Privacy International, Privacy International's response to the open consultation on the Online Harms White Paper, juillet 2019.
- 82 ↩ *Ibid.*
- 83 ↩ Voir par ex. K. Iwańska, To track or not to track: Towards privacy-friendly and sustainable online advertising, Panoptykon Foundation, novembre 2020.

84 ↩

Ibid.

85 ↩

Ibid.

86 ↩

Voir par ex. Access Now, [Raising the alarm: Online tracking harms human rights](#), 14 décembre 2020

87 ↩

Alors que le « discours de haine » n'est pas défini dans le droit international des droits humains, l'expression de la haine à l'encontre d'un individu ou d'un groupe sur la base d'une caractéristique protégée peut être divisée en trois catégories, qui se distinguent par la réponse que le droit international des droits humains exige des États : a) les formes graves de « discours de haine » que le droit international contraint les États à interdire; b) d'autres formes de « discours de haine » que les États peuvent interdire ; et c) le « discours de haine » qui est légal mais soulève néanmoins des préoccupations en termes d'intolérance et de discrimination, méritant une réponse critique de l'État, mais qui doit être protégé. Compte tenu de la complexité qui entoure ce terme, ARTICLE 19 y fait référence en tant que « discours de haine » dans ce document et d'autres rapports.

88 ↩

Par exemple, au Royaume-Uni, le suicide de Molly Russell a suscité un tollé après la découverte sur son compte Instagram de messages graphiques sur le suicide et l'automutilation ; voir par ex. BBC News, [Molly Russell: Did her death change social media?](#), 27 octobre 2019.

89 ↩

Par exemple, dans le domaine de la désinformation, la politique de Facebook a changé radicalement au fil du temps, passant d'une focalisation sur les « comportements non authentiques » (c'est-à-dire l'utilisation de faux comptes) à la suppression ou la réduction explicite de la propagation de « fausses nouvelles » ou de « médias manipulés » (par ex. les deep fakes) ; voir par ex. ARTICLE 19, [Facebook Community Standards: Analysis against international standards on freedom of expression](#), 30 juillet 2018. Twitter, pour sa part, ne supprime qu'un nombre limité d'informations trompeuses dans le contexte des élections, mais continue de lutter contre la « manipulation des plateformes » et les spams ; voir [Twitter rules on election integrity](#), 2021, et [Twitter, Platform manipulation and spam policy](#), septembre 2020. Concernant le « discours de haine », les entreprises de médias sociaux ont dû prendre des mesures pour répondre aux graves critiques qui leur ont été adressées par divers groupes de défense des droits humains, journalistes, législateurs et institutions telles que l'ONU. Par exemple, Amnesty International a dénoncé Twitter pour ne pas avoir protégé les droits des femmes en ligne et pour avoir créé un environnement toxique qui a éloigné les femmes du discours public ; voir Amnesty International, [A Toxic place for Women](#), 15 mars 2018. YouTube a été critiqué pour avoir omis de prendre des mesures contre l'animateur conservateur de YouTube Steven Crowder après que celui-ci a harcelé de manière systématique le journaliste de Vox Carlos Maza, notamment en proférant contre lui des insultes racistes et homophobes sur des vidéos YouTube pendant une période de deux ans ; voir E. Stuart, ["We don't want to be knee-jerk": YouTube responds to Vox on its harassment policies](#), Vox, 10 juin 2019. Les enquêteurs de l'ONU ont cité le « rôle » de Facebook dans la propagation du « discours de haine » – facilitant ainsi l'éventuel génocide des Rohingyas au Myanmar ; voir Rapport de la mission internationale indépendante d'établissement des faits sur le Myanmar du Conseil des droits de l'homme, 12 septembre 2018, [A/HRC/39/64](#), par. 74.

- 90 ↩ Par exemple, dans l'UE, cela comprend le [Code de bonnes pratiques contre la désinformation](#), 17 juin 2019, et le [Code de conduite de l'Union européenne visant à combattre les discours de haine illégaux](#), 2018. Au niveau mondial, le [Forum mondial de l'Internet contre le terrorisme \(GIFCT\)](#), une initiative lancée par Facebook, Microsoft, Twitter et YouTube en 2017, permet à ses membres d'accéder à une base de données de hachage que les membres peuvent utiliser pour décider de supprimer ou non du contenu sur leur propre service et en fonction de leurs propres normes communautaires. Dans le cadre de l'[appel de Christchurch](#), une initiative mondiale lancée par le gouvernement néo-zélandais suite à la diffusion en direct d'attaques terroristes à Christchurch, les fournisseurs de services en ligne s'engagent, entre autres, à « prendre des mesures transparentes et spécifiques visant à empêcher le téléchargement de contenus terroristes et extrémistes violents et à empêcher leur diffusion sur les médias sociaux et des services similaires de partage de contenus, y compris leur suppression immédiate et permanente ».
- 91 ↩ Certains États ont déjà estimé que ces mesures étaient insuffisantes et ont adopté une législation interdisant la désinformation, comme la [France](#) et [Singapour](#).
- 92 ↩ Dans l'UE, la législation sur les services numériques est également jugée nécessaire pour assurer l'efficacité du marché unique ; *op.cit.* Au Royaume-Uni, la proposition de régulateur unique pour gérer les différents problèmes posés par les contenus en ligne a l'attrait d'une apparence simplicité ; voir la Proposition de loi sur les préjudices en ligne/la sécurité en ligne (Online Harms/Online Safety Bill proposal), *op.cit.*
- 93 ↩ Voir Secrétaire d'État français au numérique, [Créer un cadre français de responsabilisation des médias sociaux : agir en France avec une ambition européenne](#), Rapport de mission, version 1.1, mai 2019, p.11; ARTICLE 19, France: [ARTICLE 19 comments on French interim report for social media regulation](#), 19 juin 2019.
- 94 ↩ Par exemple, aux États-Unis, le sénateur Hawley a cherché à présenter une législation visant à modifier l'Article 230 du Communications Decency Act (CDA) afin que les plateformes perdent leur exonération de responsabilité si un audit externe ne démontre pas de manière convaincante que leurs algorithmes et leurs pratiques de suppression de contenu sont politiquement neutres ; voir [Senator Hawley introduces legislation to amend Section 230 immunity for Big Tech companies](#), 19 juin 2019. Bien que le projet de loi n'ait pas été adopté, le débat autour de l'Article 230 fait toujours rage aux États-Unis. Par exemple, en janvier 2020, Joe Biden a déclaré vouloir le révoquer ; voir par ex. Makena Kelly, [Joe Biden wants to revoke Section 230](#), *The Verge*, 17 janvier 2020. En Allemagne, l'Autorité de l'audiovisuel a présenté une proposition qui imposerait pour la première fois des obligations de diversité sur les plateformes vidéo et de médias sociaux ; voir N. Helberger, P. Leerssen, et M. Van Drunen, [Germany proposes Europe's first diversity rules for social media platforms](#), LSE Blog, 29 mai 2019.
- 95 ↩ Voir par ex. E. M. Mazzoli et D. Tambini, [Prioritisation uncovered: The discoverability of public interest content online](#), étude DGI du Conseil de l'Europe, 2020, 19, p.40ff.
- 96 ↩ Pour plus de détails sur les défis soulevés par la diffusion d'actualités par les grandes plateformes de médias sociaux, voir P.F. Docquir & M.L. Stasi, [The Decline of Media Diversity – and How we can Save it](#), 21 janvier 2020.

- 97 ↩ [The Guardian, How social media filter bubbles and elections and algorithms influence the election](#), 22 mai 2017.
- 98 ↩ [The New York Times, Conservatives accuse Facebook of Political Bias](#), 10 mai 2016.
- 99 ↩ Voir par exemple, [The Guardian, Watchdog cracks down on tech firms that fail to protect children](#), 22 janvier 2020.
- 100 ↩ Voir par ex. D. Keller & J. von Hoboken, [Design Principles for Intermediary Liability Laws](#), Groupe de travail transatlantique de haut niveau sur la modération de contenu, octobre 2019.
- 101 ↩ Voir ARTICLE 19, [Germany: The Act to Improve Enforcement of the Law in Social Networks](#), août 2017; ARTICLE 19, [France: Analysis of draft hate speech bill](#), 3 juillet 2019.
- 102 ↩ Par exemple, des initiatives telles que le Forum Internet mondial sur la lutte contre le terrorisme (GIFCT) sont restées incroyablement opaques. Peu de données sont disponibles sur le type de contenu supprimé en fonction de la base de données de hachage GIFCT; voir la page [Transparency sur le site Internet GIFCT](#). En Allemagne, les avis restent partagés sur l'efficacité de la loi NetzDG, qui vise en grande partie à lutter contre le « discours de haine » en ligne. Les experts ont noté les difficultés à fournir une évaluation objective de la loi étant donné le manque d'uniformité des rapports de transparence par les entreprises de médias sociaux. Cependant, il existe un certain consensus sur le fait que la nouvelle loi a contribué à la suppression plus rapide du contenu, bien que son impact réel sur le « discours de haine » reste flou et difficile à prouver empiriquement ; voir par ex. A. Held, [Germany is amending its online speech act NetzDG...but not only that](#), 6 avril 2020.
- 103 ↩ En France, par exemple, un groupe de militants LGBTI a récemment publié une lettre ouverte dans le quotidien *Libération* déclarant que la proposition de loi française sur le discours de haine en ligne (Avia) serait contre-productive et entraînerait une augmentation du nombre de suppressions de contenus produits par les communautés LGBT. Le groupe a réclamé plus de moyens pour que la justice puisse poursuivre les véritables auteurs de crimes haineux, notamment les délits d'insultes racistes, entre autres ; see [Libération, Féministes, LGBTI et antiracistes, nous ne voulons pas de la loi Cyberhaine](#), tribune du 21 janvier 2020.
- 104 ↩ Cela comprend par exemple la NetzDG en Allemagne et la proposition de loi Avia.
- 105 ↩ Voir p.ex. [Facebook's recent updates](#); YouTube, [Our ongoing work to tackle hate](#), 5 juin 2019.
- 106 ↩ *Ibid.* Voir également [Wired, Twitter and Instagram unveil new ways to combat hate – again](#), 7 novembre 2019.
- 107 ↩ Voir par ex. Google, autodéclaration sur [Google News Initiative, one year in](#), 20 mars 2019.

- 108  Voir par ex. Code de conduite de l'UE visant à combattre les discours de haine illégaux, [Factsheet](#), février 2019.
- 109  [The Verge, Twitter rolls out 'hide replies' to let you tame toxic discussions](#), 19 septembre 2019.
- 110  Pour les lacunes de ces mécanismes d'appel, voir ARTICLE 19, [Sidestepping Rights, op. cit.](#), et [#Missing Voices Campaign](#).
- 111  Voir par ex. Facebook, [An Independent Assessment of the Human Rights Impact of Facebook in Myanmar](#), 5 novembre 2018; et Facebook, [A Second Update on Our Civil Rights Audit](#), 30 juin 2019.
- 112  Voir, par ex. Forbes, [Facebook is Hiring 3,000 Moderators In Push To Curb Violent Videos](#), 03 mai 2017; Financial Times, [Facebook steps up efforts to combat hate speech in Myanmar](#), 16 août 2018; ou Telegraph, [Facebook to Hire 1000 new workers in the UK to help fight toxic content](#), 21 janvier 2020.
- 113  Par exemple, dans l'Union européenne, la Commission européenne a noté dans son premier rapport de suivi de janvier 2019 sur la mise en œuvre des engagements des entreprises de médias sociaux dans le cadre du Code de bonnes pratiques contre la désinformation, qu'elle était encouragée par les politiques élaborées par les entreprises pour lutter contre la désinformation mais qu'elle était profondément préoccupée par l'absence de critères de référence spécifiques pour mesurer les progrès ; voir Commission européenne, [First monthly intermediate results of the EU Code of Practice against disinformation](#), 28 février 2019. Dans son dernier rapport de suivi de mai 2019, la Commission européenne a pris note des progrès réalisés dans l'amélioration de la **transparence des publicités politiques** et la diffusion publique de ces publicités, mais elle a ajouté qu'il fallait en faire davantage pour lutter efficacement contre la désinformation ; voir Commission européenne, [Last intermediate results of the EU Code of Practice against disinformation](#), 14 juin 2019.
- 114  Voir, par exemple, le Centre d'études des médias, de la communication et du pouvoir du King's College de Londres, [Submission to the Culture Media and Sport Select Committee's Inquiry into Fake News](#), 2017.
- 115  Cela est sans préjudice de la position plus détaillée d'ARTICLE 19 sur la publicité politique et la désinformation, à paraître.
- 116  Voir par ex., B. Wagner et L. Kuklis, [Disinformation, data verification and social media](#), 16 janvier 2020.

- 117  Par exemple, la Commission européenne a noté dans sa quatrième évaluation du Code de conduite de l'UE sur la lutte contre les discours de haine illégaux que cette initiative avait donné de bons résultats. En particulier, elle a souligné que les entreprises informatiques étaient désormais capables d'évaluer 89 % des contenus signalés dans les 24 heures, 72 % des contenus considérés comme des « discours de haine » illégaux ont été supprimés, contre respectivement **40% et 28% lorsque le code a été lancé en 2016**. Voir Commission européenne, *op. cit.*; voir également Commission européenne, Countering illegal speech online: EU Code of Conduct ensures swift response, 4 février 2019. De même, les Français ont constaté que « la rapidité de déploiement et les progrès réalisés au cours des 12 derniers mois par un opérateur tel que Facebook montrent l'intérêt de capitaliser sur cette approche d'autorégulation déjà utilisée par les plateformes, en l'étendant et en la légitimant », et que « la capacité d'autorégulation observée chez ces opérateurs fournissant un réseau social permet de les positionner comme des éléments clés de la solution aux problèmes de cohésion sociale soulevés par la présence de certains contenus sur ces plateformes » ; voir Secrétaire d'État français au numérique, Créer un cadre français de responsabilisation des médias sociaux : agir en France avec une ambition européenne, mai 2019, p. 11.
- 118  Pour plus d'information sur le Conseil de médias sociaux d'ARTICLE 19, voir ARTICLE 19, Self-regulation and 'hate speech' on social media platforms, 2 mars 2018; ARTICLE 19, Social Media Councils: Consultation, 11 juin 2019.
- 119  Voir Note explicative aux propositions de l'UE pour une Législation sur les services numériques, *op. cit.*, p. 1.
- 120  Voir Commission européenne, Leaked note on a Digital Services Act, Netzpolitik, juin 2019.
- 121  Observation générale n° 34, *op. cit.*, par. 43.
- 122  Voir par ex. R. Fletcher, Polarisation in the news media, Digital News Report, Reuters Institute and University of Oxford, 2017.
- 123  Voir la note d'orientation d'ARTICLE 19 sur les publicités politiques, à paraître.
- 124  Pour un aperçu de certaines propositions législatives, voir ARTICLE 19, Kyrgyzstan: Platform regulation laws and freedom of expression, janvier 2021.
- 125  Voir, par exemple, le projet de propositions de l'UE pour une législation sur les services numériques ; *op. cit.*; le projet de loi autrichien sur les plateformes de communication 2020, *op. cit.*; ou le projet de Loi Avia désormais invalidé, *op. cit.* C'est également l'une des caractéristiques positives de la loi allemande NetzDG, voir par ex. B. Wagner, K. Rozgonyi, M. T. Sekwenz, J. Cobbe et J. Singh, Regulating Transparency?: Facebook, Twitter and the German Network Enforcement Act, janvier 2020.
- 126  Voir, par exemple, les propositions de l'UE pour une législation sur les services numériques, *op. cit.* ou le projet de loi autrichien sur les plateformes de communication 2020, *op. cit.*

- 127  *Ibid.*
- 128  La clause de « bon samaritain » fait référence à une disposition qui garantit que les plateformes ou autres fournisseurs d'hébergement de contenu ne sont pas pénalisés en perdant l'exonération de responsabilité s'ils prennent des mesures de modération de contenu de leur propre initiative. Voir par ex. J. Barata, Positive Intent Protections: Incorporating a Good Samaritan principle in the EU Digital Services Act, The Center for Democracy and Technology, juillet 2020.
- 129  Voir, par ex. la proposition de l'UE pour une législation sur les services numériques, *op. cit.* ou le projet de loi Avia désormais invalidé, *op. cit.* C'est également une caractéristique importante de la loi allemande NetzDG, voir par ex. ARTICLE 19's analysis of the Draft NetzDG law, 2017.
- 130  Voir, par exemple, la proposition de l'UE pour une législation sur les Services numériques, la proposition de loi du Royaume-Uni sur les préjudices en ligne/la sécurité en ligne (UK Online Harms/Online Safety Bill), *op. cit.*, la Directive (EU) 2019/790 du Parlement européen et du Conseil du 17 avril 2019 sur le droit d'auteur et les droits voisins dans le marché unique numérique et modifiant les Directives 96/9/EC et 2001/29/EC (Directive 2019 de l'UE sur le droit d'auteur), le projet de loi Avia désormais invalidé sur le discours de haine en ligne, *op. cit.*, ou le plus récent projet de loi autrichien sur les plateformes de communication de 2020, *op. cit.* C'est également un aspect important de la loi allemande NetzDG, *op. cit.*
- 131  Voir les propositions au Royaume-Uni, en Irlande, en Australie et dans le memorandum de l'UE sur la loi sur les services numériques, *op. cit.*
- 132  *Ibid.*, la proposition en Irlande.
- 133  Voir proposition de loi du Royaume-Uni sur les préjudices en ligne/la sécurité en ligne (UK Online Harms/Online Safety Bill), *op. cit.*
- 134  Voir par ex. la loi NetzDG en Allemagne ou la proposition de la Avia en France, *op. cit.*
- 135  Voir par ex. les propositions de loi sur les préjudices en ligne/la sécurité en ligne, *op. cit.*
- 136  Par exemple, dans la proposition irlandaise, le Commissaire à la sécurité en ligne aurait le pouvoir de désigner tout service en ligne qui permet aux utilisateurs de partager, diffuser ou recevoir des contenus que d'autres usagers ont rendus disponibles, *op. cit.*
- 137  Voir la proposition de rapport de mission français de Facebook, *op. cit.*
- 138  Voir par ex. les propositions de loi sur les préjudices en ligne/la sécurité en ligne, *op. cit.*
- 139  Voir gouvernement britannique, Online Harms White Paper - Initial consultation response, février 2020.

- 140  Voir ARTICLE 19, [Response to the Consultations on the White Paper on Online Harms](#), juin 2019.
- 141  C.f. proposition en Inde, *op.cit.*
- 142  Les Principes directeurs, *op.cit.*
- 143  *Ibid.*, Principe IV.
- 144  Voir par ex. les propositions en France, en Irlande, au Royaume-Uni et en Australie, *op.cit.*
- 145  Voir par ex. les propositions en France et en Australie, *op.cit.*
- 146  Voir par ex. la proposition de loi Avia désormais invalidée en France ou la loi NetzDG en Allemagne, *op.cit.*
- 147  Voir par ex. la proposition de loi Avia désormais invalidée en France, *op.cit.*
- 148  Par exemple, ARTICLE 19 a participé à une Mission internationale conjointe sur la liberté de la presse en Hongrie qui a conclu que le Conseil de médias n'est pas indépendant, tous ses membres étant nommés par le parti Fidesz; voir ARTICLE 19, [Hungary: Conclusions of the Joint International Press Freedom Mission](#), 3 décembre 2019. Voir également ARTICLE 19 et la Polish Helsinki Foundation for Human Rights, [Poland: independence of public service media](#), 31 janvier 2017.
- 149  Par exemple, la loi NetzDG ne définit pas ce qui équivaut à des manquements « systémiques » à se conformer aux obligations prévues par la loi. Nous sommes en outre préoccupés par la proposition de pouvoirs accordés aux régulateurs pour ordonner le blocage de sites Internet et la pénalisation de la haute direction pour non-respect d'un vague « devoir de diligence »; *op.cit.*
- 150  Par exemple, la police indienne a perquisitionné les bureaux de Twitter à Delhi, en Inde, après que la société a qualifié les tweets des membres du parti au pouvoir de médias « manipulés », voir BuzzFeed, [Police In Delhi Have Descended On Twitter Headquarters In The Country](#), 24 mai 2021.
- 151  Par exemple, les approches réglementaires proposées en Europe pourraient ne pas être facilement transposées ou appropriées ailleurs.
- 152  See: Dilemme de la responsabilité, *op. cit.* et Sidestepping Rights, *op.cit.*
- 153  C.f. les discussions sur la législation relative aux services numériques dans l'Union européenne, *op.cit.*
- 154  Cela signifie que des entreprises comme Cloudflare ne devraient pas être pénalisées pour l'hébergement de sites Internet tels que 8chan ou DailyStormer; *op.cit.*

- 155  Dilemme de la responsabilité, *op.cit.*
- 156  Voir par ex. la Internet Watch Foundation au Royaume-Uni ou SaferNet au Brésil.
- 157  Les Principes de Manille, *op.cit.*
- 158  Cela est également cohérent avec la jurisprudence de la majorité des États membres de l'Union européenne ; voir par ex. Commission européenne, Overview of the notice and action procedures in EU member states, SMART 2016/0039, juillet 2018.
- 159  Voir L'Oreal c. eBay, C-324/09, 12 juillet 2011.
- 160  Pour une discussion détaillée sur les technologies de modération de contenu, Voir Center for Democracy and Technology, Mixed Messages: the limits of automated social media content analysis, novembre 2017.
- 161  C.f. Directive du Parlement européen et du Conseil sur le droit d'auteur sur le marché unique numérique.
- 162  D. Keller & J. von Hoboken, *op.cit.*
- 163  *Ibid.*
- 164  Pour une analyse plus détaillée d'une approche respectueuse des droits de surveillance et de suppression de contenus pédopornographiques, voir Prostasia, Child protection and the limits of censorship, juillet 2018.
- 165  Pour une analyse des implications de Eva Glawischnig-Piesczek c. Facebook Ireland Limited, C-18/18, voir G. Smith, Notice and stay down order and impact on online platforms, *Bird & Bird*, octobre 2019. Voir également D. Keller, The CJEU's new filtering case, the Terrorist Content Regulation, and the Future of Filtering Mandates in the EU, *The Center for Internet and Society Blog*, 2 décembre 2019.
- 166  Pour un exemple de ce que pourraient être ces obligations de transparence, il est possible de s'inspirer de la proposition de loi allemande sur les règles de diversité des médias pour les plateformes de réseaux sociaux, voir Germany proposes Europe's first diversity rules for social media platforms. *op. cit.*
- 167  M. MacCarthy, Transparency requirements for digital social media platforms: Recommendations for policy makers and industry, Document de travail, Groupe de travail transatlantique de haut niveau sur la modération du contenu en ligne et la liberté d'expression, 12 février 2020.
- 168  C.f. Rapport de mission français de Facebook, *op.cit.*

- 169 ↩ Les entreprises elles-mêmes ont tendance à décrire cela comme la « prévalence » du contenu, c'est-à-dire le nombre de fois qu'il a été vu.
- 170 ↩ Les Principes de Santa Clara sur la transparence et la redevabilité dans la modération de contenu, 2017.
- 171 ↩ Voir par ex., The European Commission and Member States consumer authorities ask social media companies to comply with European consumer rules, 17 mars 2017.
- 172 ↩ M. MacCarthy, *op.cit.*
- 173 ↩ C.f. *Sidestepping Rights, op.cit.*
- 174 ↩ Voir note d'orientation d'ARTICLE 19 sur la diversité des medias dans l'écosystème numérique, à paraître.
- 175 ↩ Voir la note d'orientation actualisée d'ARTICLE 19 sur la liberté d'expression et les médias, à paraître.
- 176 ↩ Pour une discussion sur l'impact des objectifs de conformité, voir Facebook, Charting a way forward, Online content regulation, février 2020.
- 177 ↩ *Ibid.*
- 178 ↩ Voir par ex. V. Baines, On online harms and folk devils: Careful now, 24 juin 2019.
- 179 ↩ C.f. Cour européenne des droits de l'homme, *Appleby et autres c. Royaume-Uni*, App. No. 44306/98, 6 mai 2003.
- 180 ↩ Voir, en particulier, H. Tworek, R.Ó Fathaigh, L.Bruggeman & C.Tenove, Dispute resolution and content moderation: Fair, accountable, independent, transparent, and effective, Document de travail, Groupe de travail transatlantique sur la modération de contenu en ligne et la liberté d'expression, janvier 2020.
- 181 ↩ Voir par ex., H. Feld, The Case for the Digital Platform Act: Market structure and regulation of digital platforms, mai 2019.
- 182 ↩ Pour plus de détails, voir M. Masnick, Protocols, not platforms: A technological approach to Free speech, 21 août 2019; P.F. Docquir et M.L. Stasi, The decline of media diversity and how we can save it, 21 janvier 2020; M.L. Stasi, Ensuring pluralism in social media markets: Some suggestions, EUI Working Papers, RSCAS 2020/05, février 2020.

Séries de normes internationales



Défendre la liberté d'expression et de l'information

✉ info@article19.org

🌐 www.article19.org

🐦 [@article19org](https://twitter.com/article19org)

📧 [@article19](https://www.instagram.com/article19)

🌐 www.linkedin.com/company/article19/

📘 facebook.com/article19org

978-1-910793-48-0

© ARTICLE 19, 2023