
RÉGIMES DE RESPONSABILITÉ POUR LES RÉSEAUX SOCIAUX ET LEURS UTILISATEURS

CADRE DE RÉGULATION

SEPTEMBRE 2022



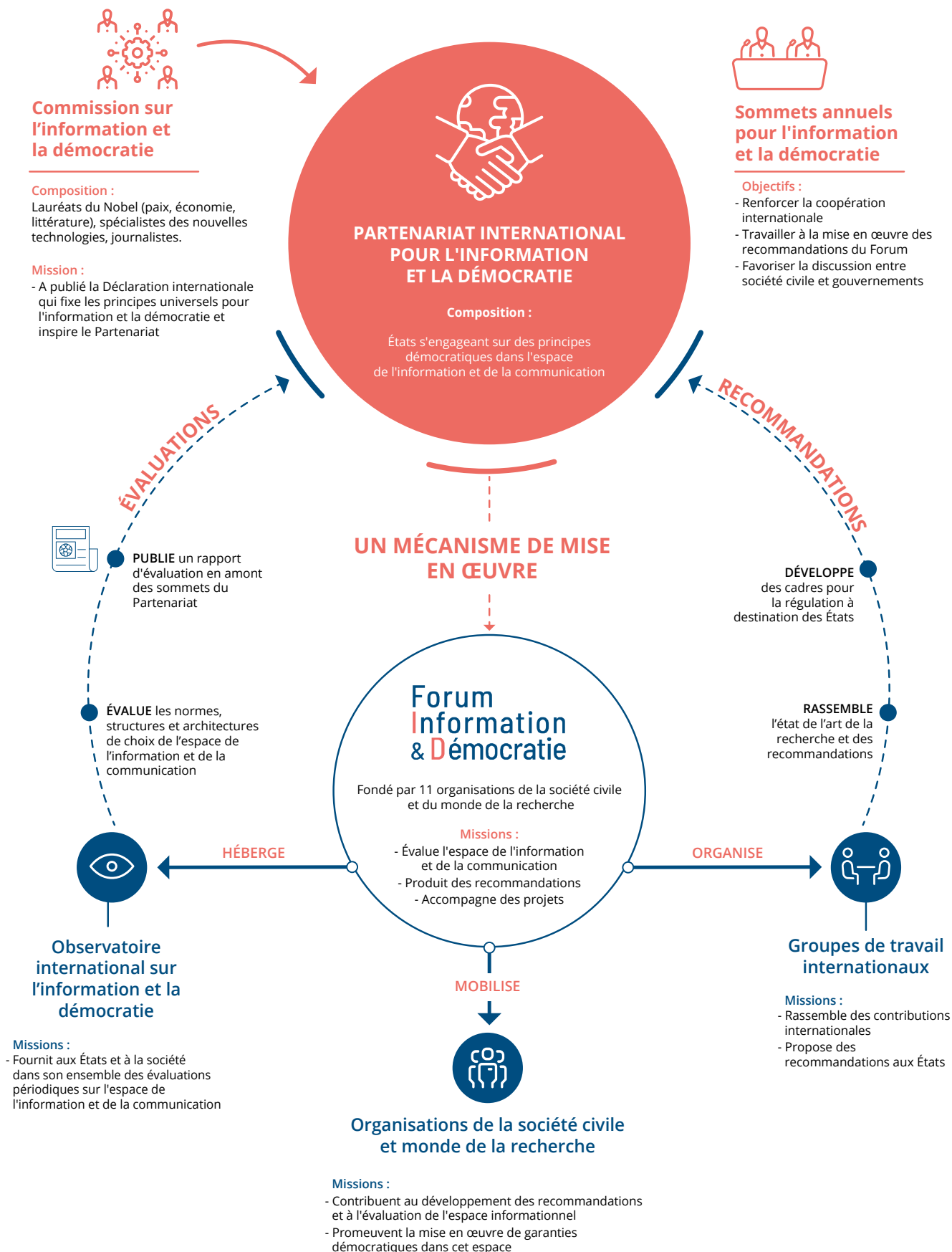
Forum
Information
& Démocratie

SOMMAIRE

A propos du Forum sur l'information et la démocratie	3
A propos du groupe de travail	4
Avant-propos de Kjersti Løken Stavrum et Damian Tambini	5
Résumé exécutif	7
Définitions	10
Introduction : D'un régime d'immunité à un régime de responsabilité	12
1/ Les différentes formes de régulation	12
2/ Les décisions des plateformes concernant les comptes	13
3/ Des lacunes à combler	14
CHAPITRE 1 : La responsabilité des réseaux sociaux	16
1.1 Les droits des réseaux sociaux	17
1.2 Les obligations des réseaux sociaux	17
1.3 Les sanctions	24
CHAPITRE 2 : Un régime général de responsabilité	25
2.1 Les droits des détenteurs de compte	26
2.2 Les obligations des détenteurs de compte	28
2.3 Les sanctions	29
CHAPITRE 3 : Des régimes spéciaux	32
3.1 Les médias d'information et les journalistes	33
3.2 Les mineurs	34
3.3 Les influenceurs	34
3.4 Les comptes très influents	36
3.5 Les administrateurs de groupe	37
CHAPITRE 4 : Gouvernance	38
4.1 Niveau national	39
4.2 Niveau international	44
Remerciements	49
Bibliographie	52

À PROPOS DU FORUM SUR L'INFORMATION ET LA DÉMOCRATIE

Pour des garanties démocratiques dans l'espace global de la communication et de l'information



À PROPOS DU GROUPE DE TRAVAIL

Le groupe de travail sur les régimes de responsabilité des réseaux sociaux et de leurs utilisateurs a été annoncé lors du Sommet pour l'information et la démocratie qui s'est tenu en marge de l'Assemblée générale des Nations unies le 21 septembre 2021.

Il est composé d'un comité de pilotage et d'une équipe de rapporteurs. Le comité de pilotage a donné des orientations de travail aux rapporteurs et au Forum.

Un appel mondial à contributions a permis au groupe de travail de recueillir des contributions venant de différentes régions du monde et de différentes disciplines scientifiques.

MEMBRES DU COMITÉ DE PILOTAGE :

- **Kjersti Løken Stavrum**, co-présidente, PDG de Tinius Trust et présidente de PEN Norvège
- **Damian Tambini**, co-président, chercheur émérite au département des médias et des communications de la *London School of Economics*
- **Damian Collins**, député à la Chambre des Communes, Royaume-Uni
- **Timothy Garton Ash**, professeur d'études européennes, Université d'Oxford
- **Bernd Holznagel**, directeur de l'Institut du droit de l'information, des télécommunications et des médias, Université de Münster
- **Anette Novak**, directrice de l'Institut suédois du Film, ancienne directrice du Conseil des médias suédois
- **Jacob Mchangama**, directeur du think tank Justitia
- **Olivier de Frouville**, professeur de droit international, Université Paris-Panthéon-Assas
- **Richard Malka**, avocat spécialisé dans la liberté d'expression
- **Karuna Nundy**, avocate, Cour suprême de l'Inde

RAPPORTEURES :

- **Sarah Jamal**, maître de conférences en droit public, Université Paris Panthéon Assas
- **Anuchika Stanislaus**, analyste des politiques technologiques

AVANT-PROPOS

LA DÉMOCRATIE EST EN CRISE ET LA LIBERTÉ D'EXPRESSION EST MISE À MAL DANS LE MONDE ENTIER.

Par Kjersti Løken Stavrum et Damian Tambini

Co-présidents du groupe de travail

L'idée démocratique - selon laquelle la légitimité de l'autorité doit être fondée sur la libre délibération entre des citoyens égaux - est confrontée aux défis existentiels de la guerre, de la crise climatique et de la pandémie, dans un contexte d'inégalités et d'injustices croissantes. Une simple révision de la réglementation des médias ne peut pas réparer la démocratie ou garantir la liberté d'expression face à ces défis et aux régimes autoritaires qui cherchent à les exploiter. Une réforme profonde des réseaux sociaux est nécessaire – et de plus en plus urgente – pour un renouveau démocratique.

La transformation radicale des systèmes d'information promise par Internet était censée promouvoir la liberté d'expression et une société transparente, permettant une organisation sociale ascendante et un nouveau débat public pluraliste et décentralisé. Cette promesse peut encore devenir réalité. Mais ces dernières années ont montré qu'elle ne se réalisera pas sans un changement politique ni une nouvelle réflexion sur la liberté d'expression. Sans règles et institutions nouvelles, le modèle économique actuel des réseaux sociaux et leur gouvernance - ou leur absence de gouvernance - continueront à favoriser l'autoritarisme et la démagogie. Fondé non pas sur la qualité de l'information et la vérification des faits mais sur le simple engagement des utilisateurs, le fonctionnement de ces médias sociaux, qui est orienté, personnalisé et adapté en fonction des données individuelles, doit être réformé. Mais notre volonté de réformes ne saurait servir de prétexte pour restreindre la liberté d'expression.

Cette volonté politique de changement vient des citoyens. Nous avons été témoins de la menace que représentent ces nouveaux « contrôleurs d'accès »¹ à Internet - portant atteinte aux médias libres dignes de confiance ; fragmentant et dégradant le débat public ; sapant le socle commun factuel indispensable à la confiance sociale et à une société civile forte. Dans la courte histoire d'Internet, des tentatives successives ont été entreprises pour qu'il contribue davantage à l'amélioration de la société : en développant des codes « volontaires » d'éthique, en encourageant la responsabilité sociale plutôt que le profit, et en imposant un ensemble de nouvelles règles sur la modération des contenus. Ces tentatives ont toutefois été largement inefficaces. Lorsqu'elles ont mené à des changements, ces remèdes étaient parfois pires que la maladie, car ces nouvelles formes de gouvernance ont servi à étouffer la liberté de parole et à s'emparer des médias libres.

Le premier paradigme en matière de responsabilité des plateformes est né de la concurrence entre les États, qui cherchent à maximiser les avantages économiques d'Internet. La prochaine génération sera fondée sur la nécessité de protéger la démocratie et les droits de l'homme tels que la liberté d'expression. Au cours de la dernière décennie, des géants du numérique immensément puissants sont apparus sur Internet et les sociétés civiles doivent convenir de nouvelles approches pour s'assurer que le pouvoir de

¹ Ce concept anglais de *gatekeeper* désigne la capacité pour certaines plateformes numériques d'imposer leurs règles et leurs normes dans l'espace numérique.

contrôle de ces nouvelles entités soit utilisé dans l'intérêt public, et non pour servir des intérêts privés ou opaques. Ces géants occupent une position critique dans la démocratie : leur modération fonctionne de plus en plus comme des adjudications, c'est-à-dire des arbitrages privés, sur les droits fondamentaux tels que la liberté d'expression, alors qu'ils sont de plus en plus sollicités par l'État pour modérer et réguler le débat démocratique. C'est pourquoi ces géants du numérique concluent des accords réglementaires avec les gouvernements du monde entier, ce qui constitue à la fois une opportunité et un danger pour la liberté d'expression et la démocratie.

Une véritable réforme prendra du temps. Le développement d'une nouvelle culture pour la responsabilité en ligne et d'une régulation équilibrée ne sera ni rapide, ni facile. Il s'agit d'un processus mondial, sur plusieurs niveaux, qui nécessite une coordination complexe entre de multiples parties prenantes. Et surtout : les réformes bénéfiques pour les démocraties pourraient avoir l'effet inverse dans les régimes totalitaires.

Ce dur travail de réforme des médias et des systèmes d'information a commencé. De nombreux pays ont présenté, et même adopté, de nouvelles lois, codes et réglementations visant à réformer radicalement les dispositions avantageuses des réseaux sociaux et autres intermédiaires numériques. Mais étant donné les menaces qui pèsent sur la liberté d'expression, ce processus d'expérimentation et de création d'institutions ne peut pas se faire uniquement dans un ou quelques pays. Il doit se faire au travers d'un véritable partenariat multilatéral et l'établissement de normes et d'institutions mondiales.

Avant tout, une réforme de gouvernance réussie exige un accord international clair sur des principes. Les recommandations de ce rapport constituent une avancée claire et décisive pour la révision des systèmes médiatiques libres et pour les réconcilier avec un nouvel écosystème de réseaux sociaux qui favorise la démocratie. Tous les États souhaitant garantir la liberté d'expression devraient de toute urgence leur apporter leur soutien.

RÉSUMÉ EXÉCUTIF

D'UN RÉGIME D'IMMUNITÉ À DES RÉGIMES DE RESPONSABILITÉ

- > Ces dernières années, l'espace de l'information et de la communication a été radicalement restructuré par les plateformes numériques et les comptes utilisés sur les réseaux sociaux.
- > Aucune disposition légale ne prévoit comment les plateformes doivent gérer les comptes d'utilisateur, ce qui donne aux entreprises une totale liberté pour établir les droits, les devoirs et les sanctions applicables à ces comptes.
- > Parce qu'elles définissent et appliquent leurs propres règles et décident unilatéralement comment sanctionner les comptes d'utilisateurs, les plateformes détiennent le pouvoir de décider qui peut ou ne peut pas participer au débat public qui prend place sur leurs services.
- > Compte tenu de l'importance que les plateformes ont prise dans la structuration du débat public, ce type de décisions devrait être encadré par des institutions démocratiques et suivre des principes démocratiques. En bref, il est nécessaire d'établir des régimes de responsabilité.
- > Cette responsabilité envers les autorités démocratiques doit s'inscrire dans un nouveau cadre de gouvernance national et international.

Ce rapport développe des régimes de responsabilité pour les plateformes numériques et leurs utilisateurs. Les recommandations présentées ci-dessous sont le résultat d'un appel à contributions international, d'une série d'entretiens et de recherches menés par le groupe de rapporteures travaillant sous la direction d'un comité de pilotage. Ces recommandations s'adressent aux États, aux plateformes de réseaux sociaux, aux organisations de la société civile et aux utilisateurs des réseaux sociaux.

Les régimes de responsabilité recommandés dans ce rapport doivent être envisagés dans le cadre du droit international encadrant les droits de l'homme, en particulier les articles 19 de la Déclaration universelle des droits de l'homme et du Pacte international relatif aux droits civils et politiques. Aucune recommandation de ce rapport ne saurait être interprétée ou utilisée comme prétexte par un État, un groupe ou une personne pour saper ou miner l'exercice des libertés et des droits de l'homme.

Un régime général de responsabilité

Les plateformes devraient :

- > Créer et appliquer un régime général de responsabilité pour les utilisateurs de réseaux sociaux qui définirait les droits, devoirs et sanctions applicables aux détenteurs de comptes ;
- > Détailler, dans le cadre du régime général de responsabilité, quels types de contenu (par exemple, la désinformation, la mésinformation, les discours de haine) et pratiques (comme le signalement massif, l'astroturfing, les comportements coordonnés inauthentiques) devraient engager la responsabilité des détenteurs de compte.

Les États devraient :

- > S'assurer que les sanctions appliquées sur les réseaux sociaux (limitation de la visibilité, suspension ou fermeture d'un compte, par exemple) soient prises dans le respect des droits de l'homme, c'est-à-dire que la sanction doit avoir une base légale, poursuivre un objectif légitime et être proportionnée ;
- > Mettre en place un mécanisme d'appel permettant aux détenteurs de comptes de contester les décisions des plateformes, en faisant appel à une autorité judiciaire.

Des régimes spéciaux

Les plateformes devraient :

- > Créer des régimes spéciaux pour les médias d'information et les journalistes, les influenceurs, les comptes très influents et les administrateurs de groupes, étant donné leur importance dans la formation du débat public et leur capacité à l'influencer ;
- > Créer un régime spécial pour les mineurs ;
- > Créer un code de conduite pour les influenceurs, les comptes très influents et les administrateurs de groupes afin d'encadrer leur activité ;
- > Imposer des sanctions qui reflètent le niveau d'influence de l'utilisateur et sa capacité à nuire.

La responsabilité des réseaux sociaux

Les plateformes devraient :

- > Créer des procédures internes d'appel permettant aux utilisateurs de contester un processus de modération ou n'importe quelle sanction prise par la plateforme ;
- > Avoir le droit de soumettre des cas spécifiques à des instances de régulation nationales ou transnationales pour demander des recommandations et fournir des rapports de transparence sur le fonctionnement de ces mécanismes.

Les États devraient :

- > Tenir les plateformes responsables en fonction de leur influence et de l'impact de leur action ;
- > Tenir les plateformes responsables quant au respect du droit national et international. Le droit international devrait être privilégié quand il est le plus favorable aux utilisateurs de compte ;
- > Imposer aux plateformes un devoir de transparence et une obligation d'information des détenteurs de compte lors de la mise en oeuvre de sanctions ;
- > Tenir les plateformes responsables quant à l'application des décisions des autorités et à l'application de sanctions ou de rétablissement de comptes conformément à leurs demandes.

La Gouvernance

Les États devraient :

Au niveau national, établir une nouvelle autorité administrative indépendante ou renforcer une autorité compétente déjà existante.

- > Cette autorité devrait avoir pour mandat de fournir de l'aide aux utilisateurs de services en ligne, de superviser les plateformes, de servir de pôle de recherche et d'émettre des conseils au niveau national en matière de politiques publiques et de réglementations.
- > Cette autorité devrait disposer des moyens et des ressources nécessaires pour atteindre ses objectifs. Une taxe spécifique sur les grandes plateformes numériques devrait être utilisée pour financer cette autorité.
- > Les décisions prises par l'autorité devraient être contraignantes pour les réseaux sociaux et entièrement indépendantes du pouvoir exécutif des États et des intérêts privés.

Créer une entité transnationale. Cette entité devrait :

- > Avoir le mandat de vérifier l'application des régimes de responsabilité par les États et les plateformes ;
- > Inclure un organe de règlement des différends fonctionnant sur les principes de l'arbitrage ;
- > Inclure un organe d'aide à la recherche chargé de collecter, consolider et conserver les informations et les preuves de violations du droit international humanitaire et de violations et d'abus des droits de l'homme.

DÉFINITIONS

1/ LES ENTITÉS

Les **réseaux sociaux** sont une catégorie de sites web avec des profils à URL unique, où chaque compte de profil peut écrire un commentaire public sur un autre profil (il peut disparaître au bout d'un certain temps ou lors de sa suppression), et où chaque détenteur de profil peut naviguer¹ sur le réseau par une série de connexions.

Les **fournisseurs de services en ligne**, comme l'énonce la Déclaration internationale sur l'information et la démocratie, sont des entités qui contribuent à structurer l'espace de l'information et de la communication en créant les moyens techniques, les architectures de choix et les normes de l'information et de la communication. Cela inclut les plateformes numériques (« plateformes »), qui sont définies comme : « *les sites et services en ligne qui (a) hébergent, organisent et font circuler les contenus partagés ou les interactions sociales des utilisateurs, (b) sans avoir produit ou commandé (la majeure partie) de ces contenus, et (c) s'appuient sur une infrastructure, sur laquelle circulent ces informations, pour traiter les données à des fins de service à la clientèle, de publicité et de profit.* »²

Dans ce rapport, les termes « plateformes » et « fournisseurs de services en ligne » sont utilisés indifféremment.

Ce rapport ne concerne pas toutes les plateformes, mais uniquement celles qui fournissent des services de réseaux sociaux dont l'activité contribue au débat public dans nos sociétés démocratiques.

Il convient de distinguer les fournisseurs de services/plateformes en ligne des :

- > **Services d'hébergement** : des services qui consistent à stocker à la demande du bénéficiaire du service des informations qu'il fournit (tels que l'hébergement web ou les services *cloud*) ;
- > **Services intermédiaires** : des services de « simple transport » qui consistent à transmettre dans un réseau de communication des informations fournies par un bénéficiaire, ou à fournir un accès à un réseau de communication (par exemple, les fournisseurs de services Internet).

Très grandes plateformes en ligne (TGPEL) : désignent les plateformes qui atteignent un seuil d'au moins 10% des consommateurs du marché régional.

2/ TYPES DE CONTENUS

Désinformation : Information fausse et délibérément créée pour nuire à une personne, un groupe social, une organisation ou un pays.³

Mésinformation : Information fausse mais qui n'a pas été créée dans l'intention de nuire.⁴

Légal mais néfaste : Contenu légal mais jugé néfaste par la plateforme.

Les **rumeurs** : Les affirmations qui ne reposent pas sur leur valeur de vérité intrinsèque mais dont l'influence découle de leur transmission sociale elle-même.⁵

Théories du complot : La croyance qu'un groupe caché d'individus puissants exerce un contrôle sur certains aspects de la société.⁶

1 Boyd, Danah M. & Nicole B. Ellison (2007).

2 Gillespie, Tarleton (2018).

3 UNESCO (2018).

4 Ibid.

5 Berinsky, A.J. (2015)

6 Vermeule, Adrian & Cass Sunstein (2009)

3/ LA MODÉRATION ET LA CURATION DE CONTENU

Modération de contenu : Lorsque les plateformes numériques analysent et filtrent le contenu généré par les utilisateurs sur la base de leurs propres règles et directives afin d'apprécier s'il faut héberger ou continuer à héberger certains éléments de contenu en vertu de leurs conditions de service. Ces décisions incluent le retrait de contenus, de manière permanente, temporaire ou par zone géographique.⁷

Curation de contenu : Décisions concernant la portée, la visibilité ou l'amplification de certains contenus. Il peut s'agir de booster, d'invisibiliser ou de rétrograder les contenus.⁸

4/ LA RÉGULATION

Régulation : Tout instrument juridique ou réglementaire contraignant que les institutions publiques locales, nationales ou régionales adoptent par le biais de leurs processus législatifs.⁹

Autorégulation : Les plateformes en ligne définissent le type de contenu acceptable dans le cadre de l'utilisation de leurs services, souvent par l'élaboration de leurs propres conditions d'utilisation. Les plateformes effectuent cette régulation principalement de deux manières : par la modération ou la curation du contenu.¹⁰

Co-régulation : Un système dans lequel les orientations générales et les résultats attendus des politiques des plateformes sont définis, avec la contribution de multiples secteurs, dans un instrument juridique qui doit être appliqué directement par les plateformes en prenant en compte les contextes locaux et régionaux, et conformément aux principes des droits de l'homme. Un organisme approprié, présentant des garanties d'indépendance et d'autonomie, doit superviser l'application de ces normes par les entreprises. La co-régulation devrait inclure la société civile et pourrait éventuellement exclure les gouvernements.

7 Access Now (2020).

8 Ibid.

9 Ibid.

10 Ibid.

INTRODUCTION

D'UN RÉGIME D'IMMUNITÉ À UN RÉGIME DE RESPONSABILITÉ

Comme l'énonce la Déclaration internationale sur l'information et la démocratie, l'espace mondial de l'information et de la communication est un bien commun de l'humanité qui revêt une grande valeur pour la démocratie. À ce titre, sa gestion devrait être fondée sur des institutions et des principes démocratiques.

Les plateformes numériques structurant cet espace détiennent un pouvoir considérable sur la capacité des citoyens à participer au processus démocratique, étant donné que les utilisateurs se servent de ces interfaces pour exercer leur droit à l'information et leur liberté d'expression et d'opinion. Ce pouvoir de façonner le débat public et la vie démocratique doit s'accompagner de garanties indissociables en matière de responsabilité.

Pourtant, la situation et la réglementation actuelles de l'espace numérique démontrent le contraire, particulièrement en ce qui concerne la gestion des comptes par les réseaux sociaux. Il y a trop peu de supervision et de transparence. Nous devons passer de l'immunité à la responsabilité.

1/ LES DIFFÉRENTES FORMES DE RÉGULATION

Le droit international traditionnel en matière de liberté d'expression a été élaboré pour garantir la liberté d'expression tout en déterminant des restrictions raisonnables à ce que les citoyens peuvent ou ne peuvent pas dire. Des normes internationales ont été formulées, notamment dans l'article 19 du Pacte international relatif aux droits civils et politiques (PIDCP), qui dispose que ces restrictions « *doivent toutefois être expressément fixées par la loi et qui sont nécessaires : a) au respect des droits ou de la réputation d'autrui ; b) à la sauvegarde de la sécurité nationale, de l'ordre public, de la santé ou de la moralité publiques* ». ¹¹ Les standards internationaux des droits de l'homme, et en particulier l'article 19 du PIDCP, devraient servir de premier cadre de référence concernant les décisions et les pratiques systémiques et individuelles de modération de contenu. En tant que tel, le Pacte international relatif aux droits civils et politiques doit primer en cas de contradiction avec des instruments internationaux ou régionaux relatifs aux droits de l'homme.

Cela implique que les pratiques de modération de contenu doivent être légitimes, nécessaires et proportionnelles dans le cadre de l'article 19(3) du PIDCP (restrictions à la liberté d'expression) qui définit les motifs de limitation de la liberté d'expression. Pour les discours de haine, les critères applicables devraient se baser sur l'article 20(2) du PIDCP (interdiction de l'appel à la haine) et prendre en considération la grille d'évaluation du seuil en six points du Plan d'action de Rabat concernant le contexte, le locuteur, l'intention, le contenu et la forme, l'ampleur de la diffusion et la probabilité d'un préjudice imminent avant d'appliquer des mesures. En ce qui concerne la désinformation, il convient de se conformer aux dispositions des articles 19(3) et 25 (droit de participer aux votes et aux élections) du PIDCP. En outre, les politiques de suppression vagues et générales sont incompatibles avec l'article 19 et seule la désinformation entraînant un préjudice réel et immédiat devrait faire l'objet des mesures restrictives les plus intrusives, comme la suppression du contenu. Pour déterminer les limites de la désinformation,

¹¹ Office of the United Nations High Commissioner for Human Rights (1966).

le contenu du message, son contexte, son impact, sa probabilité de causer un préjudice imminent et l'intention de l'auteur doivent être des paramètres décisifs.

Toutes les démocraties considèrent la liberté d'expression comme le fondement de leur régime politique, même s'il peut y avoir des interprétations culturelles différentes sur le caractère raisonnable de certaines restrictions à la liberté d'expression. Les lois et réglementations nationales peuvent prendre différentes formes dans le cadre d'une conception démocratique commune : celles-ci peuvent aller d'une absence d'intervention de l'État, la liberté d'expression est alors considérée comme la valeur la plus forte et la plus nécessaire, au détriment des autres, à un modèle plus interventionniste dans lequel l'intervention de l'État apparaît comme une garantie supplémentaire d'effectivité de cette liberté. L'effectivité de la liberté d'expression dans les pays démocratiques repose sur le caractère raisonnable des restrictions d'expression et sur l'encadrement de son application par l'État de droit.

Les lois actuelles sur la liberté d'expression permettent de statuer sur la légalité ou l'illégalité du contenu et, dans les cas pertinents, de définir le dédommagement que le contrevenant doit verser à la victime pour le préjudice subi. Ces lois s'appliquent à tous les participants au débat public en ligne et hors ligne.

De longue date, les médias ont joué un rôle central dans la formation du débat public et, à ce titre, ils ont été soumis à une réglementation spécifique (le droit des médias), tant pour garantir la liberté de la presse que pour fixer des exigences de base en matière de fiabilité et de pluralité de l'information.

Le droit des médias définit les droits et, plus souvent, les obligations des médias d'information et du public. Les médias d'information et leurs directeurs éditoriaux sont responsables de leurs publications et peuvent être poursuivis pour des délits de presse tels que la diffamation ou la violation de la vie privée. Traditionnellement, ils suivent des normes éthiques et professionnelles auto-réglementées¹² pour respecter la vérité et accorder des droits au public, tels que le droit de réponse ou de correction. Des obligations de fournir à leur public des informations pluralistes leur sont souvent imposées, notamment dans le domaine audiovisuel. Parallèlement à ces devoirs, les médias d'information ont reçu des droits spécifiques pour servir l'intérêt public et les droits fondamentaux du public, sans aucune interférence des autorités publiques ou des intérêts privés.

Ces dernières années, l'espace de l'information et de la communication a été radicalement restructuré par les plateformes numériques, qui ne sont pas tenues responsables du contenu qu'elles hébergent sur leurs services. Elles ont donné naissance à un nouvel acteur de l'espace de l'information et de la communication : les comptes des réseaux sociaux. Aucune disposition légale ne précise comment les plateformes doivent gérer les comptes, ce qui laisse aux entreprises technologiques une complète latitude pour définir les droits, les devoirs et les sanctions applicables à ces comptes.

2/ LES DÉCISIONS DES PLATEFORMES SUR LES COMPTES

Les droits, devoirs et sanctions relatifs aux comptes ne sont pas encadrés par la loi, mais uniquement par les conditions générales d'utilisation établies par les plateformes, le plus souvent à des fins commerciales et sans diligence raisonnable quant à leur impact sur les droits de l'homme et la démocratie. En raison de leur application internationale, indépendamment des frontières et des cultures, ces conditions générales ne sont pas toujours conformes au droit international ou national.

En cas de violation de leurs conditions d'utilisation, les plateformes ont pris des sanctions à l'encontre des comptes, allant de la réduction de la portée d'un compte spécifique au bannissement définitif de l'utilisateur. Ces décisions ont un impact majeur sur le débat public, en particulier lorsqu'il s'agit de personnalités publiques, de responsables politiques ou de leaders d'opinion, et sont prises sans qu'il soit nécessaire de rendre des comptes au public ou à une autorité indépendante.

¹² Par exemple, la Déclaration des devoirs et des droits des journalistes, également connue sous le nom de Charte de Munich, définit dix devoirs et cinq droits pour la profession de journaliste. Pour plus d'information, voir : European Federation of Journalists (1971).

Les plateformes ont également pris des mesures supplémentaires pour fournir des informations ou du contexte sur les comptes, en vérifiant ou catégorisant certains d'entre eux. La politique de Twitter concernant la certification des comptes de médias gouvernementaux et affiliés à l'État ou encore la politique de vérification des profils de Facebook, illustrent ces efforts d'autorégulation visant à protéger l'intégrité du débat public, en accordant parfois aux comptes de personnalités publiques des exemptions de modération spécifiques.¹³

Mais cette autorégulation des comptes reste préoccupante du fait de son opacité et son caractère arbitraire et unilatéral. D'une part, les plateformes numériques ne sont pas tenues responsables de l'application de leurs propres conditions générales et accordent parfois des privilèges à certains détenteurs de comptes, ce qui leur permet d'enfreindre les politiques de la plateforme.¹⁴ D'autre part, d'autres détenteurs de compte sont confrontés à des décisions de modération inexpliquées et discrétionnaires, alors même qu'ils sont publiquement reconnus comme des sources d'information fiables et professionnelles.¹⁵ En l'absence de diligence raisonnable en matière d'impact de leurs règles et politiques internes, les plateformes ont adopté des pratiques préjudiciables aux droits de leurs utilisateurs.

L'impossibilité pour les détenteurs de comptes de contester convenablement les sanctions appliquées par les plateformes est également problématique. Premièrement, le *shadow banning* ou d'autres sanctions non signalées empêchent les détenteurs de comptes d'exercer leur droit de recours, par manque de connaissance de la sanction. Deuxièmement, la procédure d'appel est souvent effectuée dans le cadre de mécanismes internes obscurs. Chaque entreprise a développé ses propres modalités et mécanismes d'appel, consistant en général à un nouveau passage en revue du contenu par l'équipe d'assistance.¹⁶ L'exception la plus célèbre est le Conseil de surveillance de Meta, qui est composé d'experts indépendants. Depuis 2020, le Conseil a sélectionné les cas de suppression ou de suspension les plus complexes et a décidé, en seconde instance, si la modération était appropriée ou devait être annulée. Meta peut également demander au Conseil de surveillance des recommandations sur des questions de modération délicates, mais celles-ci ne sont pas contraignantes.

Avec leurs propres normes et leurs décisions unilatérales de sanctionner des comptes, les plateformes détiennent le pouvoir de décider qui peut ou ne peut pas participer au débat public ayant lieu sur leurs services. Compte tenu du rôle que ces plateformes ont pris dans la structuration du débat public, ces types de décisions devraient être encadrés et contrôlés par des institutions démocratiques, et selon des principes démocratiques.

3/ DES LACUNES À COMBLER

Des droits et obligations spécifiques doivent être définis pour les détenteurs de comptes et les plateformes numériques afin d'établir un cadre clair respectant les droits de l'homme et les principes de l'État de droit. Ce cadre démocratique doit établir des sanctions nécessaires, proportionnées et équitables en cas de violation de ces droits et obligations, y compris en cas de violation des conditions générales d'utilisation par les détenteurs de comptes.

Une procédure de sanction équitable et transparente doit être mise en place pour empêcher les plateformes de modérer de manière opaque et arbitraire. Elle doit offrir aux détenteurs de compte un droit de recours auprès d'un mécanisme interne de règlement des différends établi par la plateforme, et

13 Twitter (2019).

14 Plusieurs exemples pourraient être mentionnés, tels que le programme interne XCheck de Facebook, révélé par le Wall Street Journal en septembre 2021, et la modification par Meta de sa politique de glorification de la violence dans certains pays d'Europe et d'Asie centrale. Pour plus d'informations, voir respectivement : Horwitz, Jeff (2021) et Vengattil, Munsif & Elizabeth Culliford (2022).

15 Pour ne citer que quelques cas : le cas du journaliste français David Thomson, le récit du magazine The Kashmir Walla en Inde, ou encore le récit de l'organe d'information d'investigation équatorien Periodismo de Investigación. Pour plus d'informations, voir respectivement : Reporters sans frontières (2016), Reporters sans frontières (2021) et Observacom (2022).

16 Pour plus d'informations sur les politiques de chaque plateforme, voir l'analyse comparative "Comment faire appel" de Electronic Frontier Foundation (EFF), consulté le 18 août 2022 via : <https://www.onlinecensorship.org/en/pages/1214ccb8-e1d5-472a-91e8-745d895b4dff>.

les informer sur les moyens de saisine d'une autorité nationale ou d'un tribunal au cas où ils souhaiteraient faire appel de la décision interne.

Ce passage d'un régime d'immunité à des régimes de responsabilité pour les plateformes et les détenteurs de compte est essentiel pour garantir le légitime contrôle des institutions démocratiques sur le débat public. Cette redevabilité envers les autorités démocratiques doit s'inscrire dans un nouveau cadre de gouvernance. Une autorité administrative nationale devrait être mandatée pour s'assurer que les plateformes numériques respectent toutes leurs obligations et les droits de leurs détenteurs de compte. Les utilisateurs devraient avoir facilement accès à des institutions judiciaires indépendantes dédiées aux litiges impliquant les plateformes.

Le pouvoir que les plateformes numériques détiennent sur le débat public dépasse largement l'échelle nationale. Pour contrebalancer leur influence considérable sur l'espace mondial de l'information et de la communication, les institutions nationales ne suffiront pas : des organismes internationaux doivent être créés ou mandatés pour garantir le respect des droits humains universels par les plateformes et les tenir responsables en vertu du droit international.

Alors que les dernières législations sur les plateformes numériques étaient principalement axées sur la réduction des risques systémiques qu'elles représentent pour la société,¹⁷ aucune initiative législative ne vise à combler les vides problématiques qui entourent la responsabilité des plateformes et des utilisateurs concernant les comptes et leur contenu.

Ce rapport développe des régimes de responsabilité pour les plateformes numériques et leurs utilisateurs. Conformément aux principes de la Déclaration internationale et du Partenariat pour l'information et la démocratie, il entend créer des garanties démocratiques protégeant l'exercice de la liberté d'opinion et d'expression et le droit à l'information dans l'espace numérique de l'information et de la communication.

17 Nous pouvons voir cette approche dans la législation sur les services numériques (DSA sous son acronyme en anglais) adoptée dans l'UE (voir annexe 1) ou dans le projet de loi sur la sécurité en ligne (Online Safety Bill) au Royaume-Uni.

Chapitre 1 :

La responsabilité des réseaux sociaux

Il est essentiel que les réseaux sociaux respectent les principes et les garanties qui protègent la nature démocratique de l'espace de l'information et de la communication. Conformément à la Déclaration internationale sur l'information et la démocratie, les réseaux sociaux doivent être responsables en proportion de leur influence et de l'impact de leur contribution.

1.1 LES DROITS DES RÉSEAUX SOCIAUX

Les réseaux sociaux devraient avoir le droit de :

Demander conseil à un organisme de régulation transnational

Les plateformes peuvent faire appel à l'expertise de l'organisme de réglementation transnational lorsqu'elles sont confrontées à une situation difficile.

Établir une définition du contenu jugé légal mais néfaste

Les réseaux sociaux ont le droit d'inclure dans leurs conditions générales une définition des contenus considérés comme légaux mais néfastes, et de prendre des décisions de modération appropriées¹⁸ pour atténuer leur diffusion, si les conditions suivantes sont remplies :

- leur définition devrait être élaborée dans le cadre d'un processus de délibération inclusif, avec la participation des personnes, des communautés et des organisations de la société civile concernées;
- tant la définition que les décisions prises sur sa base doivent être conformes aux standards internationaux des droits de l'homme, notamment à l'article 19 du PIDCP.

1.2 LES OBLIGATIONS DES RÉSEAUX SOCIAUX

Nous recommandons aux réseaux sociaux de s'engager à respecter les obligations suivantes en ce qui concerne leur relation avec les utilisateurs, leur conformité avec les demandes des autorités et leurs prises de décisions.

Droit national et international

Les réseaux sociaux devraient respecter les obligations qui leur incombent en vertu du droit international¹⁹ et du droit national conforme aux obligations internationales des États dans le domaine des droits de l'homme et des libertés fondamentales. Dans l'esprit de l'alinéa 2 de l'article 5 du PIDCP, les réseaux sociaux devraient appliquer la loi la plus favorable aux droits de l'homme et spécifiquement à la liberté d'expression de l'utilisateur qui a posté le contenu.

Processus de décision équitable et transparent

Les réseaux sociaux devraient respecter cinq principes dans leurs solutions de modération : les principes de personnalisation, de compatibilité, de communauté, de précaution et de décentralisation.²⁰ De plus, la procédure de modération doit être nécessaire, proportionnée et conforme aux principes de leurs conditions de service.

Les plateformes devraient analyser une série de paramètres avant de prendre toute décision de blocage, de démonétisation ou de promotion de contenus ou de comptes, que cette décision soit prise par des pratiques de modération automatisées ou par des modérateurs humains, en prenant en compte :

- la qualité de l'expéditeur (il convient d'être plus prudent pour les journalistes ou les personnes ayant un haut niveau d'influence) ;
- l'intention de l'auteur de la publication (l'image est-elle postée dans un but militant/informationnel, etc.) ;

18 Sander, Barrie (2020).

19 United Nations (2011).

20 Ermoshina, Ksenia (2022).

- le format du message et le contexte entourant son contenu ;
- la probabilité d'un préjudice imminent pour les utilisateurs et le public ;
- l'étendue de la diffusion du contenu ;
- la participation (ou non) de l'information à un débat d'intérêt général dans la société.

Modération automatisée du contenu

Les réseaux sociaux devraient expliquer la logique de l'algorithme et les critères auxquels il répond. Les algorithmes devraient suivre un **principe d'explicabilité** envers les utilisateurs et expliquer²¹ :

- le nombre de fois où le contenu a été sélectionné, modéré et ordonné ;
- les chiffres détaillés pour chaque action de modération effectuée ;
- les critères et les données utilisés pour entraîner les algorithmes à identifier, modérer, hiérarchiser et personnaliser le contenu ;
- la collecte de données, y compris des données personnelles ;
- les biais des données et les biais reproduits par les algorithmes ;
- les mauvais usages et abus potentiels des algorithmes, les taux de faux positifs / faux négatifs ;
- les procédures utilisées pour corriger les erreurs.

Les plateformes devraient mettre en place des équipes dédiées à la qualité du dialogue avec les utilisateurs à travers un service de médiation pour s'assurer que les utilisateurs comprennent correctement les informations. Les explications de l'algorithme devraient être fournies dans toutes les langues dans lesquelles le réseau social opère.

Les plateformes doivent organiser une forme de « joignabilité » des systèmes algorithmiques, notamment en identifiant systématiquement au sein de chaque entreprise ou administration une équipe responsable du fonctionnement d'un algorithme dès lors qu'il traite des données personnelles. Cette équipe doit être facilement joignable et avoir les moyens de répondre rapidement aux demandes reçues. Cela pourrait se faire grâce à :

- la mise en place d'équipes dédiées à la qualité du dialogue avec les utilisateurs ;
- mettre en place un service de médiation pour s'assurer que les utilisateurs comprennent les informations dont ils ont besoin.

Efficacité des procédures de signalement

Les réseaux sociaux devraient mettre en œuvre pour chaque détenteur de compte une procédure de signalement efficace²² qui réponde aux principes d'accessibilité, de clarté et de lisibilité. Les utilisateurs doivent être informés immédiatement si un contenu est retiré, bloqué, rétrogradé ou démonétisé, ou si un avertissement est associé au contenu.

Les réseaux sociaux devraient mettre en place une procédure de signalement encadrée pour signaler les contenus répréhensibles sur les plateformes, comme ceux qui constituent un appel au terrorisme. De plus, l'utilisateur qui a signalé le contenu devrait être informé de l'évolution de sa demande. La **procédure de signalement devrait être simplifiée** : il devrait être possible de faire le signalement en trois clics maximum et elle devrait être standardisée sur toutes les plateformes.

Procédure d'appel

Afin de garantir un procès équitable aux utilisateurs, les réseaux sociaux devraient s'engager à mettre en place des processus d'appel.²³ Cela inclut une instance d'appel où contester des décisions de *shadow-banning*, et l'établissement d'un dialogue entre les utilisateurs et les équipes responsables du fonctionnement des algorithmes utilisant les données personnelles. Ces équipes devraient être représentatives et formées sur l'impact de leur activité.

21 Ermoshina, Ksenia (2022).

22 Council of Europe (2021).

23 Common, MacKenzie, (2019).

Régimes spéciaux

Les réseaux sociaux devraient reconnaître, au minimum, les régimes spéciaux identifiés au chapitre 3 de ce rapport et établir des procédures pour s'adapter en conséquence à ces droits et devoirs supplémentaires.

Coopération avec d'autres entités

Les réseaux sociaux devraient soumettre à un tribunal arbitral les litiges concernant la modération du contenu quand la situation présente un certain niveau de complexité (un cas qui concernerait une personne publique internationale devrait être considéré comme une affaire complexe).

Ils devraient, de leur propre initiative, transmettre à un mécanisme d'enquête international toute information contribuant à démontrer la violation des droits de l'homme, ou tout acte illégal qualifié de crime grave (génocide, crime contre l'humanité, crime de guerre), ainsi que toute information pertinente demandée par ce mécanisme d'enquête.

La plateforme devrait transmettre l'élément identifiant l'utilisateur à une autorité judiciaire indépendante en cas de procédure judiciaire sous certaines conditions :

- l'acte doit être considéré comme illégal en vertu d'une loi nationale conforme au droit international des droits de l'homme.
- l'État doit avoir un rang élevé au regard des principes de justice et de liberté d'opinion et d'expression fixés par l'organisme de régulation transnational. Si l'organisme de régulation transnational est persuadé que les droits de l'homme des utilisateurs seront respectés, l'identité des utilisateurs peut être révélée à l'autorité.
- l'identification de l'utilisateur doit être demandée par une autorité judiciaire indépendante et impartiale dans le cadre d'une procédure respectant les principes d'un procès équitable.

Les **autorités compétentes** devraient²⁴ :

- avoir accès aux données permettant d'identifier les personnes qui ont posté le contenu problématique ;
- avoir accès au contenu concerné ;
- mettre en place une transmission automatique d'alerte vers une plateforme publique dédiée, voire une plateforme accessible par les enquêteurs internationaux.

Respect des décisions des autorités

Les réseaux sociaux devraient s'engager à respecter les décisions des autorités nationales qui sont conformes aux standards internationaux des droits de l'homme.

Lorsqu'une autorité judiciaire indépendante prive le détenteur du compte de l'accès au réseau social, **la plateforme a l'obligation d'appliquer la sanction.**

Coopération avec l'organisme de réglementation transnational

Les réseaux sociaux devraient s'engager à collaborer et coopérer avec l'organisme de régulation transnational (voir chapitre 4).

Nomination d'un directeur juridique

Il est essentiel que les réseaux sociaux s'engagent à nommer un directeur légal pour assurer leur représentation dans les États dans lesquels ils opèrent. Le directeur juridique sera le principal interlocuteur des autorités nationales compétentes et représentera les plateformes dans les procédures judiciaires.

Responsabilité algorithmique

Les réseaux sociaux devraient s'engager à être responsables des dommages causés par leurs algorithmes. Tout réseau social utilisant des algorithmes qui créent une expérience utilisateur toxique, par exemple en

24 Heldt, Amélie & Stephan Dreyer (2021).

amplifiant les contenus haineux, devrait être désigné comme producteur de contenu et tenu responsable du préjudice causé.

Indemnisation des victimes

Les réseaux sociaux devraient s'engager à mettre en place une procédure respectant les principes du procès équitable afin de déterminer l'existence d'un préjudice, ainsi que la forme et le montant de l'indemnisation requise pour la victime.

Ils devraient avoir la capacité de **procéder à une indemnisation** si et quand une erreur de modération a imposé une perte financière à l'utilisateur.

Conditions générales d'utilisation

Les réseaux sociaux devraient rendre publiques leurs politiques concernant les contenus et les comportements générés par les utilisateurs qui sont autorisés ou non : pour les règles relatives aux contenus et au ciblage publicitaire ; pour la modération, le classement et le ciblage des contenus ainsi que pour les recommandations de socialisation (suggestions de nouveaux "amis", pages et comptes à suivre, etc.) ; et pour le traitement et la divulgation des données des utilisateurs. Les conditions générales d'utilisation doivent être claires pour que les utilisateurs puissent y accéder et les comprendre, et doivent préciser les sanctions applicables aux détenteurs de comptes en cas de non-respect.

Les conditions générales devraient aligner leurs définitions des contenus problématiques, tels que le « discours de haine » et l'« apologie du terrorisme », sur celles élaborées par les institutions internationales publiques, telles que les organes des traités des Nations unies concernés, les entités mandatées dans le cadre de procédures spéciales et d'autres experts, comme l'OMS, et les tribunaux internationaux.²⁵

Modification des conditions générales d'utilisation

Les réseaux sociaux devraient mettre en place **une procédure publique et transparente de modification** de leurs conditions générales d'utilisation et prévoir des mesures transitoires appropriées.

- En cas de mise à jour, les utilisateurs devraient être informés et avoir accès aux versions précédentes de la charte d'utilisation, ainsi qu'aux raisons pour lesquelles elle a évolué.
- Les conditions générales doivent être faciles à trouver sur le site lors de l'inscription et lors de l'utilisation générale des plateformes, et être disponibles dans différentes langues en fonction de leur diffusion à l'international.

Les conditions générales doivent inclure une politique relative à la collecte, à la conservation et au partage de contenus susceptibles d'être utilisés comme preuves en cas de violation du droit international, et une procédure permettant aux utilisateurs de signaler des violations des droits de l'homme lors de troubles sociaux, et en cas de restriction de l'accès à l'internet dans un pays.

Participation des utilisateurs

Les réseaux sociaux devraient mettre en place un mécanisme permettant aux utilisateurs de modifier les conditions générales d'utilisation. Ce mécanisme devrait inclure une représentation des utilisateurs dans la gouvernance de la modération, les organes de décision, les conseils consultatifs de la plateforme (conseils consultatifs du contenu, conseil confiance et de sécurité, etc.), comme c'est déjà le cas dans certaines entreprises pour les salariés.

Inclusion de la société civile

Les réseaux sociaux devraient inclure la société civile dans leur autorégulation²⁶ à travers une plateforme numérique qui pourrait agréger les retours des utilisateurs et les cas litigieux en construisant des outils

²⁵ Kaye, David, (2019)..

²⁶ Isaac, Henri & Louis-Victor de Franssu (2022).

de régulation appropriés (indicateurs, algorithmes, systèmes de signalement, etc.). La société civile devrait être impliquée lorsque la plateforme réalise ses évaluations de risques, conçoit ses mesures d'atténuation des risques et élabore des codes de bonne conduite.

Les modalités de participation de la société civile doivent respecter une procédure transparente, assurer la représentation de différentes catégories (journalistes, ONG, etc.), et avoir une représentation égale au niveau des zones géographiques et des genres.

Curation et modération de contenu

Les réseaux sociaux devraient rendre publiques leurs politiques concernant la modération, le classement et le ciblage du contenu ainsi que les recommandations de socialisation (suggestions de nouveaux "amis", pages et comptes à suivre, etc.). Ils devraient maintenir à jour les documents de référence sur chaque fonction essentielle des algorithmes, notamment le classement, le ciblage, la modération et les recommandations sociales, ainsi que la détection de contenu.²⁷

Détection de nouvelles formes de discours illégaux ou néfastes

Les réseaux sociaux devraient investir dans l'identification des communications néfastes complexes en ligne, en particulier lorsque les utilisateurs extrémistes emploient des euphémismes génériques pour échapper à une détection facile.

Modérateurs humains et automatisés

Les réseaux sociaux devraient mettre en place de manière transparente un ratio entre modérateurs humains et modérateurs automatisés. La suppression de contenu devrait être effectuée par des modérateurs humains et non par des algorithmes, à partir d'un certain seuil de complexité du contenu. Ce seuil devrait être déterminé par une coalition d'experts en droit international relatif aux droits de l'homme et d'experts issus de la société civile. Ils devraient être impliqués dans la conception et l'évaluation du degré de complexité du contenu à modérer.

Modération humaine

Les réseaux sociaux devraient être obligés d'**assurer un certain niveau de modération humaine pour chaque pays, qui devrait être représentatif de la population**, afin que les plateformes puissent prendre des décisions en fonction du contexte.

- Les modérateurs doivent représenter les minorités.
- Les modérateurs doivent être régulièrement formés au contexte géopolitique local, au droit local applicable et au droit international.
- Les modérateurs humains doivent maîtriser la langue vernaculaire utilisée dans les messages qu'ils traitent.

Les modérateurs doivent être soutenus psychologiquement et bénéficier de conditions de travail convenables.

Accès des chercheurs et des organisations de la société civile aux informations

Les réseaux sociaux devraient permettre aux chercheurs et aux organisations de la société civile d'accéder aux **informations contextualisées** qui entourent les données agrégées actuellement fournies pour tous les contenus modérés.

Cette obligation de transparence permettra de mettre en lumière tout type de partialité ou de schéma de discrimination, et contribuera à protéger les minorités ou les régions du monde où la sur-modération serait disproportionnée.²⁸ Les chercheurs et les organisations de la société civile peuvent aider les

²⁷ Forum on Information & Democracy (2020).

²⁸ Pen America (2021).

plateformes à déterminer si un contenu est illégal ou doit être supprimé. Ils peuvent également soutenir les utilisateurs victimes au cours de la procédure. Cela permettrait aux législateurs d'adopter des lois et des réglementations adaptées aux besoins grâce à des données exhaustives issues du terrain.

Ces fiches de données doivent être anonymisées.

Les réseaux sociaux devraient ouvrir les **systèmes algorithmiques (codes sources)** et les rendre accessibles aux chercheurs indépendants.

Les réseaux sociaux devraient également mettre en place des **procédures de test en mettant à disposition les données d'activité des utilisateurs** via une interface de programmation d'application (*application programming interface*, API sous son acronyme anglais), ce qui pourrait être la base d'actions collectives s'appuyant sur leurs données (*unionization* des données).

Obligations de divulgation

Compte tenu du fait que les plateformes de données partagent déjà les informations²⁹ concernant les contenus terroristes et extrémistes violents, elles devraient élargir leurs rapports de transparence à toutes les autres catégories de contenus. Toutes les obligations de rendre publiques des informations supplémentaires devraient être faites sous le contrôle d'un audit indépendant, comme c'est le cas pour les rapports financiers des entreprises.

La publication de ces informations devraient s'inspirer des principes de Santa Clara sur la transparence et la responsabilité dans la modération des contenus,³⁰ établis par des organisations de la société civile, qui définissent le niveau minimum de détails que les plateformes devraient divulguer.

Évaluation d'impact et audits

Les réseaux sociaux devraient s'engager à publier des évaluations d'impact annuelles sur leurs algorithmes, conformément aux responsabilités qui leur incombent en la matière, telles que définies par différents États.

Les plateformes devraient mettre en œuvre une évaluation régulière et publique des risques systémiques afin de s'assurer de la conformité des algorithmes avec leurs objectifs initiaux, et d'identifier les potentiels biais dommageables qu'ils présentent. Les algorithmes devraient être testés régulièrement, en faisant appel à des modérateurs humains.

Recrutement inclusif

Les réseaux sociaux devraient promouvoir un recrutement favorisant la diversité culturelle, sociale et de genre pour chaque profession impliquée dans la conception d'algorithmes, et être transparents quant à la diversité sociale parmi les concepteurs d'algorithmes.

Formation

Les réseaux sociaux devraient fournir une formation continue aux concepteurs d'algorithmes (développeurs, programmeurs, codeurs, scientifiques des données, ingénieurs) pour encourager la prise de conscience des implications de leur activité et de leur devoir de vigilance.

Paramètres de priorisation des contenus

Les réseaux sociaux devraient laisser les paramètres de hiérarchisation des contenus accessibles et modifiables par les utilisateurs. Cela devrait inclure :

- de la transparence sur la politique de *shadow-banning* – les raisons, la quantité de contenu qui a fait l'objet de cette mesure, etc. – ainsi que de la transparence en ce qui concerne des biais de « bulle de filtres » (*filter bubble*) et l'accessibilité des données utilisées par l'algorithme pour sélectionner le

29 OCDE (2021).

30 The Santa Clara Principles (2018).

prochain contenu que l'utilisateur verra ;

- des informations à l'utilisateur quand son contenu est en train d'être dépriorisé.

Signalement de masse

Les réseaux sociaux doivent s'engager à interdire les signalements massifs dans leurs règles de modération et à limiter le nombre de signalements qu'un même compte peut effectuer en une journée.

Contenu ou comportement illégal

Les réseaux sociaux devraient créer un portail accessible à tous les utilisateurs afin de signaler les contenus ou comportements illicites. Ce portail aiderait les utilisateurs à identifier les responsables de ces contenus et pourrait aboutir à la saisine des autorités judiciaires.

Comptes automatisés

Les réseaux sociaux devraient obliger les utilisateurs à se soumettre régulièrement à un test Captcha (par exemple une fois par mois), et pas seulement lors de la première connexion, afin de limiter l'action des comptes automatisés. Les comptes exploités par des *bots*, ou robots, seront facilement repérables et pourront faire l'objet d'une déconnexion ou d'autres sanctions.

Détection de la viralité artificielle

Les réseaux sociaux devraient mettre en œuvre les moyens nécessaires pour mieux détecter la viralité artificielle de contenus. Les plateformes devraient utiliser leur capacité pour développer un mécanisme de friction limitant la viralité des contenus.

Réapparition de contenu illégal signalé

Les réseaux sociaux devraient s'assurer techniquement que les contenus qui ont été signalés et supprimés ne puissent pas réapparaître.

Des informations fiables et complètes

Les réseaux sociaux devraient promouvoir, à travers leurs algorithmes, la diffusion d'informations conformes à une méthodologie garantissant leur fiabilité et leur indépendance, telle qu'une norme ISO (par exemple, la *Journalism Trust Initiative* [JTI]). Les grandes plateformes qui participent à la structuration de l'espace de l'information et de la communication doivent respecter le principe de neutralité politique, idéologique et religieuse. Les algorithmes qui diffusent et traitent les informations et les idées doivent être neutres par rapport aux intérêts de ceux qui les contrôlent, à l'exception de la publicité politique et sociale qui devrait être explicitement identifiée. En effet, la publicité politique et sociale doit être présentée aux utilisateurs dans un contexte présentant différents points de vue.

Information des utilisateurs sur les modifications apportées à l'algorithme

Les réseaux sociaux devraient informer les utilisateurs des changements apportés à leurs algorithmes, en particulier lorsque ces changements ont un impact sur la viabilité financière des détenteurs de comptes (comme pour les médias d'information et les influenceurs).

1.3 LES SANCTIONS

Il est nécessaire que les sanctions adoptées à l'encontre des réseaux sociaux soient adaptées à la gravité de l'infraction commise et à la capacité financière du réseau. La sanction doit être prononcée soit par l'autorité nationale indépendante compétente, soit par une autorité judiciaire indépendante.

Les sanctions doivent être à la mesure du niveau d'influence et des ressources de la plateforme.

Le choix de la sanction doit tenir compte de la gravité du non-respect de l'obligation, ainsi que des capacités financières de la plateforme.

L'échelle des sanctions

Les sanctions pourraient aller **d'une alerte publique à une amende civile pouvant s'élever jusqu'à 6 % du chiffre d'affaires annuel mondial de la plateforme.**



RECOMMANDATIONS AUX ÉTATS

- > Garantir que les réseaux sociaux respectent leurs obligations.
- > Imposer des sanctions en cas de non-respect.

Chapitre 2 :

Un régime général de responsabilité

Il est important de reconnaître que les détenteurs de compte, étant de nouveaux participants à l'espace mondial de l'information et de la communication, ont des droits à faire valoir et des devoirs à respecter. Les détenteurs de compte agissent trop souvent en toute impunité, cachés derrière des pseudonymes. Il est indispensable de reconnaître qu'ils sont responsables des contenus qu'ils diffusent ou contribuent à diffuser, et qu'ils peuvent être tenus pour responsables des dommages occasionnés. Les sections suivantes présentent les droits et obligations attribués aux détenteurs de comptes en général, ainsi que les sanctions applicables et voies de recours disponibles.

2.1 LES DROITS DES DÉTENTEURS DE COMPTE

Les paragraphes suivants décrivent les droits qui devraient être reconnus aux détenteurs de comptes quand ils utilisent les réseaux sociaux.

Règles régissant la liberté d'expression

Les utilisateurs ont le droit d'être informés des règles régissant la prise de parole qu'ils sont censés respecter en ligne. Les détenteurs de compte ont le droit de disposer d'informations claires et exhaustives sur la politique de modération de la plateforme.

- Les informations doivent inclure une liste de contenus clairement répréhensibles qui feront l'objet de sanctions. Cette liste devrait inclure tous les modèles actuels, mais aussi ceux qui passent à travers les mailles du filet, comme les images d'abus d'enfants et le programme « XCheck » de Meta, entre autres.
- Tous les utilisateurs devraient connaître les critères sur lesquels se fondent les décisions de modération d'une plateforme, savoir à qui la politique s'applique, quels critères seront utilisés pour définir la notion d'intérêt public, quelle décision est prise et dans quelles circonstances, et quelles sont les sanctions possibles.

Rétrogradation et suppression de contenu

Les utilisateurs ont le droit d'être informés si et quand leur contenu fait l'objet d'une modération ou d'une suppression (« invisibilisation »)³¹. Les raisons de la décision ainsi que les voies de recours disponibles pour la contester doivent être précisées.

Les notifications devraient être suffisamment détaillées pour permettre à l'utilisateur d'identifier spécifiquement le contenu soumis à une restriction, et devraient inclure des informations sur la manière dont le contenu ou le compte a été détecté, comment il a été évalué, et pourquoi il a été supprimé ou rétrogradé. Néanmoins, dans certains cas, un juge indépendant et impartial peut ordonner qu'un utilisateur qui, par exemple, a fait preuve à plusieurs reprises d'activités de harcèlement à l'égard d'un autre utilisateur ne soit pas informé de l'invisibilisation du contenu qu'il a publié, afin de préserver la sécurité de la ou des victimes pendant la procédure contre l'auteur.

Appel

Les détenteurs de compte ont le droit de contester des décisions de modération d'un réseau social, telles que la suppression d'un message ou la fermeture d'un compte. Si le mécanisme d'appel du réseau social n'est pas satisfaisant, les utilisateurs ont le droit de faire appel à une autorité administrative indépendante ou aux autorités judiciaires.

Procédure équitable

Les détenteurs de compte ont le droit de bénéficier d'une procédure de procès équitable et le droit de saisir une autorité administrative indépendante ou des autorités judiciaires pour contester les décisions de modération et les sanctions qui leur sont imposées par les réseaux sociaux.

³¹ Nicholas, Gabriel, (2022).

Accès aux réparations

Les victimes ont droit à la réparation d'un préjudice subi en ligne et doivent pouvoir participer au processus de réparation. Les réparations doivent compenser toutes les formes de dommages, matériels ou moraux. Le préjudice doit être réparé dans son intégralité, sans que le montant de la réparation ne dépasse le montant du dommage subi.

La réparation doit privilégier la compensation financière, sauf lorsqu'une mesure symbolique peut contribuer à la réparation du préjudice subi. De la sorte, la réparation peut également prendre la forme d'excuses publiques publiées sur le réseau social.

Accès aux données générales pertinentes

Les utilisateurs devraient avoir le droit d'accéder aux données générales pertinentes concernant la modération du contenu sur les réseaux sociaux, telles que :

- le contenu qui a échappé à l'IA et aux modérateurs humains, qui n'a pas été signalé ou qui n'a pas été supprimé mais qui aurait dû être modéré ;
- les procédures mises en place pour minimiser les erreurs de modération ;
- le nombre de contenus restaurés pendant la période donnée ;
- les procédures automatisées de modération du contenu, les informations sur leurs objectifs, les indicateurs d'exactitude et les mesures de protection appliquées ;
- le nombre total de décisions de modération appliquées par les réseaux sociaux.

Autonomie et proactivité

Les utilisateurs devraient avoir le droit d'être autonomes et de participer de manière proactive aux processus de modération.

- Les utilisateurs devraient disposer d'outils adéquats pour modérer le contenu publié sur leur propre compte, page ou groupe.
- Les utilisateurs devraient être actifs dans la désactivation, la suppression ou le filtrage des commentaires sous leurs publications, et être alertés lorsque des contenus malveillants sont signalés sur leur espace, afin qu'ils puissent les modérer eux-mêmes en appliquant leurs propres règles spécifiques et des *plugins* de blocage automatique.

Les utilisateurs devraient avoir la possibilité d'intervenir par l'intermédiaire de leurs représentants dans le processus de modération.

- Avant que les modérateurs des réseaux sociaux ne soient informés du contenu, celui-ci devrait être envoyé à un groupe d'utilisateurs représentatifs qui pourraient prendre des décisions de manière collaborative, ou individuellement par l'auto-modération de leurs propres pages.

Participation à la modification des conditions générales

Les utilisateurs devraient pouvoir participer et intervenir par l'intermédiaire de leurs représentants dans la procédure de modification des conditions générales d'utilisation d'un réseau social, ainsi que dans les chartes applicables aux régimes spéciaux.

Sélection d'un représentant des utilisateurs

Les détenteurs de compte ont le droit de choisir régulièrement des représentants des utilisateurs par le biais d'une procédure publique et transparente. Il pourrait s'agir d'une élection régionale par les utilisateurs, ou d'un tirage au sort tenant compte des différentes catégories d'utilisateurs.

Interopérabilité des données

Les utilisateurs devraient avoir la possibilité de choisir le degré d'interopérabilité (la capacité des systèmes

informatiques ou des logiciels à échanger et à utiliser des informations) relatif à leurs données, sur la base de normes ouvertes.

Droits sur les données

Les données des utilisateurs sont la matière première du modèle économique des réseaux sociaux.

« Les données détenues par les utilisateurs répondent à toutes les caractéristiques d'un « actif » dans les lois sur la propriété, indépendamment du fait que les données pourraient être considérées comme intangibles. »³²

Ainsi, les utilisateurs devraient avoir le contrôle de leurs données. Cela signifie qu'ils soient en mesure de³³ :

- accéder aux données pour examiner les informations qui ont été collectées ou en ont été déduites ;
- effacer les données d'un appareil et d'un serveur ;
- transférer des données personnelles vers une autre plateforme ;
- corriger des données pour s'assurer que leur identité est correctement représentée dans un système ;
- être protégés en cas de délit lié à leur identité (vol d'identité et fraude à l'identité) ;
- être libres de toute prise de décision automatisée, afin de s'exclure des profilages automatisés ;
- s'opposer, retirer leur consentement et restreindre leur utilisation au cas où l'utilisateur reviendrait sur sa décision.

2.2 LES OBLIGATIONS DES DÉTENTEURS DE COMPTE

Les utilisateurs devraient également s'engager à respecter les obligations ci-dessous lorsqu'ils interagissent sur les réseaux sociaux.

Personnes responsables du contenu

Les participants au débat public sont responsables du contenu qu'ils diffusent ou contribuent à diffuser. En outre, compte tenu de l'effet de viralité sur les médias sociaux, nous recommandons de tenir les détenteurs de compte responsables de l'amplification d'un contenu néfaste, même s'ils n'en sont pas les auteurs, sauf si le contenu expose des informations d'intérêt public. Le caractère d'intérêt public du contenu devrait être déterminé par une autorité judiciaire indépendante.

Identification des détenteurs de comptes

Comme l'énonce la Déclaration internationale sur l'information et la démocratie, la responsabilité de tous les participants au débat public est un principe essentiel qui implique la transparence sur leur identité. En vertu du principe de responsabilité individuelle, chaque participant est responsable du contenu qu'il publie sur les plateformes. Il est donc fondamental de pouvoir identifier la personne qui se cache derrière la publication.

Néanmoins, les exceptions au principe de transparence peuvent être légitimes si elles facilitent l'obtention d'informations critiques relevant de l'intérêt public ou si elles contribuent à la sécurité des participants. En effet, l'anonymat peut représenter une garantie permettant aux participants de s'exprimer.

³² Jurcys, Paulius, et al. (2020).

³³ The World Wide Web Consortium (2022).

Nous recommandons de tempérer le principe de garantie de l'anonymat en cas de nécessité en permettant la levée de l'anonymat dans des conditions spécifiques et strictes.

Les utilisateurs ont le droit d'être informés si un réseau social a engagé une procédure visant à révéler leur identité, et peuvent demander une évaluation par des organismes internationaux indépendants tels que le Comité des droits de l'homme.

Signalement de masse

Les utilisateurs ne doivent pas effectuer de signalements de masse. Ils doivent respecter les limites du nombre de signalements pouvant être faits au cours d'une période donnée.

Usurpation d'identité de personnes, de groupes ou d'organisations

Les détenteurs de compte ne peuvent pas se faire passer pour d'autres personnes, groupes ou organisations dans le but de tromper les autres. Toutefois, les parodies ou les pages de fans peuvent bénéficier d'exemptions. Dans ce cas, ces types de comptes doivent être explicitement présentés comme tels.

Manipulation de la plateforme

Les utilisateurs ne peuvent pas manipuler la plateforme pour modifier l'accès aux informations ou amplifier artificiellement des contenus. Cela inclut le *spam* ou d'autres activités qui manipulent l'expérience des utilisateurs.³⁴

Intégrité électorale et civique

Les utilisateurs ne doivent pas utiliser les réseaux sociaux dans l'intention de manipuler les processus électoraux ou démocratiques.

Mésinformation

Les utilisateurs ne doivent pas partager de fausses informations pouvant entraîner un préjudice physique imminent, ni partager des informations fausses et trompeuses en temps de crise, comme les conflits sociaux armés, les urgences de santé publique et les catastrophes naturelles.

2.3 LES SANCTIONS

Les détenteurs de compte doivent s'assurer qu'ils respectent les conditions d'utilisation d'un réseau social. En cas de non-respect, les utilisateurs doivent être tenus pour responsables. Les sanctions imposées par les réseaux sociaux doivent être prises dans le respect des droits de l'homme, c'est-à-dire que la sanction doit avoir une base juridique, poursuivre un objectif légitime et être proportionnée.

Le critère de proportionnalité - également appelé test nécessaire et proportionné - exige que le choix de la sanction soit fondé en particulier sur les critères suivants : nature et gravité de l'acte, préjudice subi par la victime, intention de l'auteur de l'acte, risque de récurrence et excuses à la victime, entre autres.

De nombreuses dispositions légales couvrent déjà certaines sanctions liées à la violation des conditions

³⁴ Tony Blair Institute (2021).

générales d'utilisation, mais ne portent pas sur les moyens qui ont été utilisés pour perpétrer l'acte lui-même sur le réseau social. Le détenteur du compte qui a commis un acte illégal ou enfreint les conditions générales d'utilisation, quel qu'en soit le degré de gravité, conserve les moyens de commettre un acte illégal sur le réseau social après avoir été reconnu responsable. Afin de tenir responsables les détenteurs de compte et de prévenir la répétition des infractions, nous recommandons l'adoption d'un cadre de sanctions plus large visant à empêcher les détenteurs de compte à continuer à commettre des activités illégales en leur enlevant la possibilité d'utiliser leur compte.

Les détenteurs de compte ont le droit de faire appel des sanctions devant un régulateur indépendant ou devant l'autorité judiciaire nationale.

ÉCHELLE DES SANCTIONS POUR LES UTILISATEURS

Échelle de sanctions en fonction des comportements spécifiques (cumulatif et complémentaire aux sanctions pénales)

Une alerte et un avertissement	Réduction de la visibilité du contenu	Retrait du contenu	Limitation des fonctionnalités du réseau social	Suspension temporaire du compte				Suspension de l'accès à la plateforme pour plus de 10 ans
				< 1 an	> 1 an	> 5 ans	> 10 ans	
Une violation isolée des conditions d'utilisation		Violations répétées des conditions d'utilisation						Haute fréquence de violations des conditions d'utilisation
								Menace de mort, discours de haine, incitation à la violence, cyberharcèlement
Violation des conditions d'utilisation de la part de comptes à forte audience			Violation des obligations en matière d'identification et de transparence des sponsors					Violations répétées des obligations en matière d'identification et de transparence des sponsors
Non-respect des obligations en matière de publicité			Non-respect répété des obligations en matière de publicité					



RECOMMANDATIONS AUX **ÉTATS**

S'assurer que les réseaux sociaux reconnaissent et protègent les droits des détenteurs de comptes.

- > **En cas de non-conformité, les États peuvent imposer des sanctions, comme indiqué au chapitre 1, section 1.3.**
- > **Fournir un mécanisme permettant aux détenteurs de compte de contester les pratiques de modération et les sanctions des réseaux sociaux.**
- > **Tenir les détenteurs de compte responsables des délits commis en ligne.**
- > **Prévoir une procédure judiciaire pour déterminer les réparations à accorder aux victimes. En outre, les États devraient mettre en place un fonds de solidarité nationale alimenté par les réseaux sociaux pour indemniser les victimes au cas où l'auteur de l'acte ne pourrait pas le faire pour des raisons financières.**

Chapitre 3 :

Des régimes spéciaux

Profondément convaincus de l'égalité de tous devant la loi, nous réaffirmons reconnaître le principe d'égalité devant la loi garanti par l'article 26 du Pacte international relatif aux droits civils et politiques (PIDCP).

Toutefois, une différence de traitement n'est pas discriminatoire si la mesure en question est fondée sur un objectif légitime. Non seulement la différence de traitement est permise, mais elle est moralement nécessaire. En effet, certaines catégories de détenteurs de compte se trouvent dans des situations qui exigent qu'ils reçoivent un traitement différencié. L'importance du rôle de certaines catégories de détenteurs de compte (par exemple, ceux qui œuvrent en faveur de la protection des droits de l'homme ou du droit à l'information) nécessite qu'ils bénéficient de garanties procédurales supplémentaires. Certains détenteurs de compte doivent se soumettre à des obligations plus importantes, en raison de leur influence dans le débat démocratique et donc de leur capacité de nuisance.

Par conséquent, **nous recommandons l'élaboration de régimes de responsabilité spéciaux pour certaines catégories d'utilisateurs** en raison de l'importance de leur rôle dans le débat public, des différentes obligations qui leur incombent ou de leur forte capacité d'influence et de nuisance.

3.1 LES MÉDIAS D'INFORMATION ET LES JOURNALISTES

Le journalisme est confronté à une menace existentielle. La publication des médias d'information et des personnes exerçant la fonction de journaliste sur les plateformes peut être modérée ou invisibilisée, ce qui mène à la censure. Ces décisions de modération se font sans consultation et donc sans débat contradictoire.

Il est nécessaire de prévoir un régime de responsabilité avec des garanties supplémentaires pour permettre aux médias d'information et aux journalistes de remplir leur rôle dans nos démocraties.

DROITS SUPPLÉMENTAIRES

Statut spécifique

Les médias d'information devraient bénéficier d'un statut spécifique. Les plateformes des réseaux sociaux devraient reconnaître que les médias d'information journalistique ont une fonction sociale spécifique, et que leurs comptes devraient bénéficier de certaines garanties et protections, à condition que les médias appliquent les principes du journalisme. Les médias d'information devraient être identifiés en tant que tels par des entités indépendantes des gouvernements, selon des processus non discrétionnaires, sans aucune interférence politique, en utilisant autant que possible des normes d'autorégulation et des mécanismes de conformité transparents.

Préserver la visibilité

Compte tenu de la responsabilité éditoriale des médias d'information, le contenu journalistique ne doit pas être déclassé, rétrogradé ou supprimé. Les décisions de modération prises à l'encontre des médias d'information et des contenus journalistiques ne doivent pas avoir d'incidence négative sur la visibilité de leurs publications ou leur monétisation.

Examen par un tiers

La modération des contenus journalistiques ne devrait se faire que dans le cadre d'une procédure menée par une autorité indépendante, à l'initiative du média ou à la demande de toute personne intéressée, y compris par le biais de la plateforme. La procédure devrait prendre la forme d'un appel accéléré.³⁵

Informations sur les changements d'algorithmes affectant la stabilité financière

Les médias et les journalistes ont le droit d'être informés des modifications apportées aux algorithmes qui mettraient en péril leur stabilité financière.

Rémunération équitable

Étant donné que les algorithmes d'une plateforme peuvent avoir un impact sur la stabilité financière des médias en réduisant la visibilité de leur contenu, les médias d'information ont le droit de demander une compensation pour les dommages subis par les modifications de l'algorithme, afin d'assurer leur viabilité financière.

³⁵ Aspen Institute (2021).

3.2 LES MINEURS

Les plateformes prennent une place de plus en plus importante dans la plupart des aspects de la vie des enfants. Il est donc nécessaire d'établir un régime de responsabilité spécifique qui tienne compte de l'âge et des capacités des utilisateurs.

Nous devons également reconnaître le droit des mineurs à la liberté d'expression sur les médias sociaux, conformément à l'article 13³⁶ de la Convention internationale des droits de l'enfant (1989). Ce droit comprend la liberté de rechercher, de recevoir et de répandre des informations et des idées de toute espèce, sans considération de frontières, sous une forme orale, écrite, imprimée ou artistique, ou par tout autre moyen du choix de l'enfant.

DROITS SUPPLÉMENTAIRES

Accès à un régime juridique

Les mineurs ont le droit d'accéder à un régime juridique fondé sur l'intérêt supérieur de l'enfant.

Accès à des contenus adaptés aux enfants

Les mineurs ont le droit d'accéder à des chaînes et à des contenus sûrs et adaptés aux enfants. Ils ont également le droit de demander en toute sécurité des conseils et de l'aide concernant les contenus sexuellement explicites qu'ils ont eux-mêmes générés ou qu'ils ont reçus.

Accès à des procédures adaptées selon l'âge

Les mineurs ont le droit d'accéder à des procédures adaptées à leur âge en fonction de leur intérêt supérieur : par exemple, les plateformes peuvent avoir recours à des tiers de confiance à l'inscription³⁷ d'un mineur afin qu'ils puissent confirmer l'identité ou l'âge de l'utilisateur sans révéler d'information permettant son identification. De plus, les plateformes peuvent être encouragées à étudier la faisabilité d'un système de vérification de l'âge³⁸ basé sur le système d'exploitation d'un smartphone, d'une tablette ou d'un ordinateur, qui ne peut être modifié que par l'opérateur ou par le parent sur présentation d'une preuve de filiation.

3.3 LES INFLUENCEURS

Les plateformes ont permis l'émergence d'un nouveau métier : les influenceurs. Il s'agit d'un individu qui, grâce à sa visibilité sur Internet, a une influence sur les utilisateurs qui le suivent et sur leurs décisions d'achat. Les influenceurs se distinguent des simples détenteurs de compte par leur grande visibilité et

36 "L'enfant a droit à la liberté d'expression ; ce droit comprend la liberté de rechercher, de recevoir et de répandre des informations et des idées de toute espèce, sans considération de frontières, sous une forme orale, écrite, imprimée ou artistique, ou par tout autre moyen du choix de l'enfant."

37 CNIL (2022).

38 Institut Montaigne (2021).

le fait qu'ils exercent une activité rémunérée. Les avantages économiques qu'ils tirent de leur visibilité ainsi que leur forte capacité à influencer les choix des utilisateurs justifient l'établissement d'un régime de responsabilité spécifique.

Nous recommandons aux réseaux sociaux de reconnaître les influenceurs comme des utilisateurs qui reçoivent de l'argent ou toute autre forme de récompense en échange de la promotion d'un produit ou d'un service via leur influence sur une plateforme.

DROITS SUPPLÉMENTAIRES

Rétrogradation et suppression de contenu

Étant donné que la décision de modération d'une plateforme peut avoir un impact sur la situation financière des influenceurs en réduisant la visibilité de leur contenu, les influenceurs ont le droit d'être informés lorsqu'une décision de modération ou une modification des algorithmes peut avoir un impact financier sur eux.

OBLIGATIONS SUPPLÉMENTAIRES

Code de conduite

Les influenceurs doivent adhérer à un code de conduite imposé par la plateforme, encadrant leur activité selon des principes clairs.

Contenu sponsorisé

Tous les contenus qui ont été sponsorisés et postés par un compte d'influenceur doivent être présentés en tant que tel à tous les utilisateurs du réseau social.

Informations sur les contrats de publicité

Les influenceurs doivent rendre publics tous les accords publicitaires, ainsi que le montant des sommes perçues et la stratégie de ciblage de la publicité politique. Ces informations publiques devraient inclure :

- le lieu et la source des paiements ;
- la taille du public cible, ainsi que le nombre de vues de la publicité concernée et le taux d'engagement des utilisateurs après la visualisation de la publicité ;
- les critères de sélection pour le ciblage des destinataires (tout en protégeant la vie privée) tels qu'ils sont communiqués aux annonceurs, y compris la source des données, le profil inféré, les audiences similaires, les audiences personnalisées et le recours au test A/B ;
- les revenus issus de la publicité ciblée.

SANCTIONS

Echelle des sanctions

Les sanctions doivent être à la mesure du niveau d'influence de l'influenceur. La détermination de la sanction applicable doit tenir compte du fait que ces détenteurs de compte agissent en tant que professionnels et doivent donc se soumettre à des obligations renforcées par leur influence.

3.4 LES COMPTES TRÈS INFLUENTS

Les comptes à forte visibilité, même s'ils n'opèrent pas nécessairement sur les plateformes pour faire du profit, ont une forte influence sur les utilisateurs en raison de la taille de leur communauté, ce qui entraîne une forte visibilité de leur contenu. Cette visibilité augmente nécessairement le risque de causer un préjudice plus important, ce qui nécessite une surveillance plus stricte pour prévenir ou arrêter les abus plus rapidement.

Un compte est considéré comme très influent lorsqu'il remplit une ou plusieurs des conditions suivantes : il compte **50 000 followers ou plus** ; il s'agit d'un compte appartenant à une personne exerçant une fonction représentative (responsables politiques, représentants de l'autorité publique) ; ou à d'autres personnalités publiques certifiées par les plateformes.

DROITS SUPPLÉMENTAIRES

Procédure d'appel accélérée

Les comptes très influents devraient avoir le droit d'accéder à une procédure d'appel accélérée pour décider des décisions de modération et des sanctions.

OBLIGATIONS SUPPLÉMENTAIRES

Code de conduite

Les personnalités publiques et les hommes politiques devraient adhérer à un code de conduite imposé par la plateforme, encadrant leur activité selon des principes clairs, tels que l'identification et la transparence de leur affiliation politique ou d'entreprise.

Atteinte aux droits d'autrui

Il faut reconnaître que les personnalités publiques ont une obligation renforcée de ne pas porter atteinte aux droits d'autrui, y compris le droit à l'information.³⁹

Informations sur les contrats de publicité

Les comptes très influents doivent rendre publics tous leurs accords publicitaires, ainsi que le montant des sommes perçues et la stratégie de ciblage de la publicité politique.⁴⁰ Ces informations publiques devraient inclure :

- le lieu et la source des paiements ;
- la taille du public cible, ainsi que le nombre de vues de la publicité concernée et le taux d'engagement et le taux d'engagement des utilisateurs après la visualisation de la publicité ;
- les critères de sélection pour le ciblage des destinataires (tout en protégeant la vie privée) tels qu'ils sont communiqués aux annonceurs, y compris la source des données, le profil inféré, les audiences similaires, les audiences personnalisées et le recours au test A/B ;
- les revenus issus de la publicité ciblée.

39 Brennan Center for Justice (2021).

40 Reisach, Ulrike (2021).

SANCTIONS

Échelle des sanctions

Les sanctions doivent être à la mesure du niveau d'influence et de la capacité de nuisance du compte très influent. La détermination de la sanction applicable doit tenir compte de leur forte capacité d'influence et donc de nuisance. Elle doit tenir compte du contexte et de la visibilité du contenu.

3.5 LES ADMINISTRATEURS DE GROUPE

La gestion de certains contenus sur les plateformes est assurée par un administrateur de groupe. Son rôle est de superviser, gérer et administrer un groupe ou un compte sur un réseau social, ce qui lui confère un pouvoir de modération important. Ainsi, les administrateurs de groupe peuvent avoir la capacité d'agir pour faire cesser la diffusion de contenus illégaux ou néfastes dans leur groupe. C'est pourquoi les administrateurs de groupe devraient se soumettre à des obligations spécifiques et bénéficier de droits supplémentaires.

DROITS SUPPLÉMENTAIRES

Une charte pour les membres du groupe

Les administrateurs de groupe ont le droit d'adopter et d'imposer à leurs membres une charte qui, sous réserve qu'elle soit conforme aux droits de l'homme et à la politique de la plateforme, devrait pouvoir inclure des spécificités justifiées par l'objet du groupe.

Modération active du contenu

Les administrateurs de groupe ont le droit de participer activement à la modération du contenu de leur groupe.

OBLIGATIONS SUPPLÉMENTAIRES

Code de conduite

Les administrateurs de groupe doivent suivre un code de conduite imposé par la plateforme, encadrant leur activité selon des principes clairs.

Coopération avec la plateforme et les autorités

Les administrateurs de groupes devraient être obligés, lorsqu'ils sont informés d'un acte illégal publié sur leur groupe, de coopérer avec la plateforme et les autorités compétentes pour faire en sorte que l'individu mette fin à son activité et soit poursuivi en justice.



RECOMMANDATIONS AUX ÉTATS

- > S'assurer que les réseaux sociaux reconnaissent des régimes spéciaux et établissent des procédures en conséquence.
- > En cas de non-respect, les États peuvent imposer des sanctions, comme indiqué au chapitre 3, section 1.3.
- > Adopter une définition commune des influenceurs en tenant compte de leurs activités professionnelles sur les réseaux sociaux.

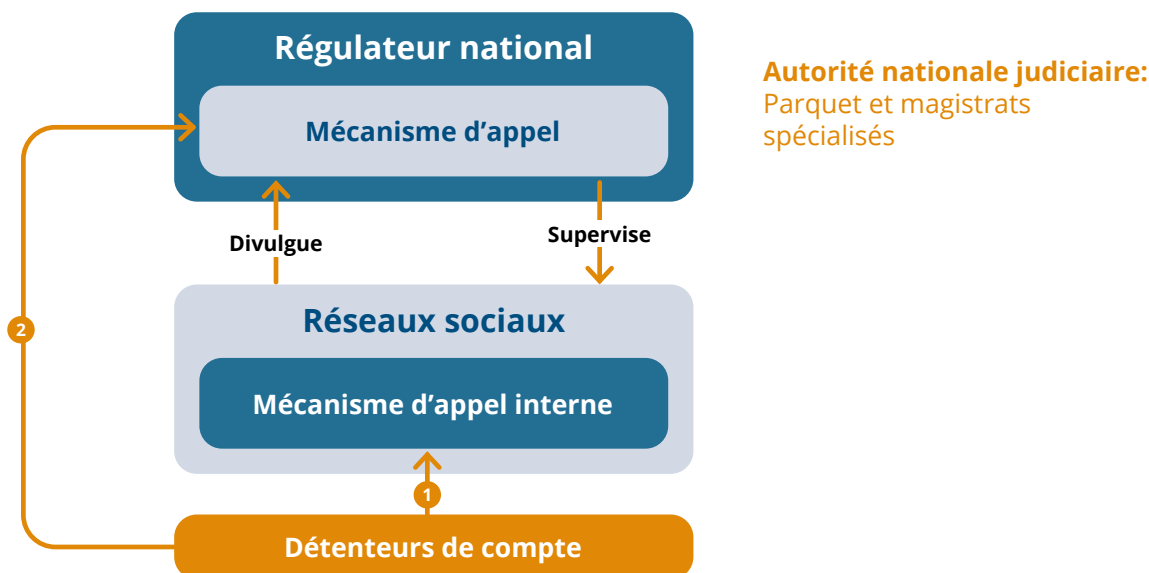
Chapitre 4 :

Gouvernance

La Déclaration internationale sur l'information et la démocratie affirme que la responsabilité de tous les participants au débat public est un principe essentiel. Afin de mettre en œuvre les régimes de responsabilité exposés dans le chapitre précédent, des organes et des procédures de gouvernance doivent être mis en place.

Les réseaux sociaux sont par nature une activité transnationale qui nécessite une coopération entre les différents États et participants pour limiter efficacement les comportements illégaux en ligne et préserver la démocratie. Il est donc essentiel de renforcer la gouvernance nationale par la création d'organismes internationaux.

4. 1 NIVEAU NATIONAL



Le premier niveau d'action pour la mise en œuvre des régimes de responsabilité devrait être le niveau national. Il serait composé d'un régulateur indépendant, d'un mécanisme de signalement des comportements ou des contenus illégaux, de l'accès à un système judiciaire indépendant et à un parquet et des magistrats spécialisés.

4.1.1 UNE AUTORITÉ DE RÉGULATION INDÉPENDANTE

Nous recommandons la mise en place d'un régulateur indépendant, ou le renforcement d'un organe administratif indépendant déjà existant, pour superviser, promouvoir et protéger les droits des utilisateurs sur les plateformes et pour résoudre les différends entre les utilisateurs et la plateforme.

MANDAT :

Nous recommandons que cette autorité ait quatre compétences :

1. Supervision des plateformes

Nous recommandons que cette autorité veille à l'application et au respect des règles par les plateformes, à moins que des missions spécifiques n'aient été confiées à d'autres autorités compétentes.

Nous recommandons que ce contrôle soit exercé par le biais des différentes fonctions suivantes. L'autorité compétente devrait :

- recevoir les rapports d'information des plateformes sur leurs pratiques et les évaluer ;
- disposer de certains pouvoirs d'investigation, tels que le pouvoir d'exiger des informations aux fournisseurs, de procéder à des inspections sur site et de demander aux employés d'expliquer leurs décisions ;
- être compétente pour contrôler les conditions générales d'utilisation (vérification *ex ante*) afin de s'assurer que les plateformes respectent toute la législation applicable ;

- pouvoir imposer des sanctions aux plateformes ;
- être compétente pour clarifier toute question relative à l'identification de comptes spécifiques, y compris les comptes très influents, les personnalités publiques ou les influenceurs (voir chapitre 3) ;
- être en mesure d'émettre des recommandations, afin d'aider les plateformes à adapter leurs pratiques ;
- examiner les évaluations d'impact des algorithmes et superviser la tenue régulière de tests de résistance.

Les décisions de cette autorité devraient être contraignantes pour toutes les questions concernant l'accès à l'information, y compris la résolution des différends liés à ces questions.

Ces pouvoirs sont de nature administrative et ne cherchent pas à remplacer ceux des tribunaux.

L'autorité devrait avoir le pouvoir de sanctionner les plateformes qui ne font pas tout leur possible, conformément aux normes industrielles élevées de diligence professionnelle, pour **éviter la diffusion de contenus illégaux, de désinformation néfaste et d'œuvres non autorisées sur leurs services**.

L'autorité doit tenir les plateformes responsables en cas de :

- absence d'action rapide, dès qu'elles en ont connaissance, pour supprimer ou désactiver l'accès à un contenu en ligne illégal, qu'il s'agisse d'un discours de haine ou d'un produit illégal ;
- non-coopération pas avec l'autorité nationale ;
- suppressions abusives et répétitives de contenus/erreurs multiples de modération s'appuyant sur l'IA ;
- absence de mécanismes d'appel permettant aux utilisateurs de contester une décision ;
- violation des obligations de transparence à l'égard des autorités nationales ou internationales, des utilisateurs et des chercheurs agréés ;
- refus de donner aux utilisateurs le pouvoir de choisir un algorithme qui ne soit pas basé sur un profil comportemental ;
- développement, utilisation et application de l'intelligence artificielle, des algorithmes et d'autres technologies similaires manifestement incompatibles avec le droit et les standards internationaux en matière de droits de l'homme ;
- utilisation de *dark patterns* ou interfaces truquées, qui conçoivent une architecture de choix qui induit en erreur ou oriente les utilisateurs vers des décisions qu'ils n'auraient peut-être pas prises autrement ;
- non-réalisation des audits annuels sur la manière dont leurs algorithmes ont une incidence sur la démocratie, les droits de l'homme et la santé physique et mentale des mineurs et des autres utilisateurs ;
- pratiques commerciales trompeuses si leur communication publique n'est pas conforme à leurs pratiques de modération, y compris en cas de changement de règles sans avertissement ;
- profilage de personnes mineures et utilisation de données sensibles relatives à l'orientation sexuelle, à la religion et à l'origine ethnique à des fins publicitaires.

2. Rôle consultatif

Nous recommandons que cette autorité fasse des recommandations et participe à l'élaboration des politiques publiques et des réglementations au niveau national. Nous encourageons la création d'un organe multipartite au sein de l'autorité administrative visant à élaborer des principes communs, des règles, des procédures basées sur des recommandations et des normes réglementaires internationales, ainsi que des programmes décisionnels pour définir l'utilisation et l'évolution d'Internet.

- Cet organe national devrait avoir une compétence consultative.
- **Il devrait être consulté pour chaque projet législatif ayant un impact sur les plateformes.**
- Il devrait être composé d'experts ayant des compétences reconnues dans le domaine de la modération et des plateformes.

- Ses membres devraient représenter les différentes parties prenantes concernées par ces questions (société civile, médias, ONG, etc.).

3. Soutien aux utilisateurs de services numériques

Nous recommandons que cette autorité assure le rôle d'intermédiaire avec les plateformes en :

- mettant en relation des utilisateurs et des plateformes ;
- aidant et guidant les victimes ;
- traitant les recours des détenteurs de comptes. Nous recommandons que ces décisions soient contraignantes pour la plateforme.

4. Pôle de recherche

Les plateformes détiennent un contrôle monopolistique de l'accès de la recherche à leurs services. À ce titre, **nous recommandons la création d'un pôle de recherche (cluster)** au sein de l'autorité administrative. Il réunirait des chercheurs et des spécialistes du numérique pour faire progresser la connaissance des systèmes de modération algorithmique des contenus, en développer de nouveaux, faire des audits de ceux des principales plateformes, et mettre en relation les plateformes et les chercheurs pour améliorer l'accès aux données des plateformes dans un format anonymisé.



MISE EN ŒUVRE PAR LES ÉTATS

- > **Les membres de cet organisme public de régulation doivent être indépendants.**
- > **Leur processus de nomination doit être ouvert, transparent et associé des parties prenantes pertinentes.**
- > **Les membres sélectionnés doivent être des personnes de haute moralité reconnues pour leur compétence dans les domaines de la technologie, et comprendre des personnes ayant une expérience juridique.**
- > **Toutes les parties prenantes intéressées, y compris la société civile, doivent être représentées équitablement.**
- > **Une taxe spécifique sur les grandes plateformes numériques devrait servir à financer cette autorité.**

4.1.2 MÉCANISME DE SIGNALEMENT DES COMPORTEMENTS OU CONTENUS ILLÉGAUX

Nous recommandons de mettre en place un mécanisme permettant de signaler les comportements ou contenus illégaux en ligne.

Un grand nombre de contenus illégaux publiés sur les plateformes sont modérés sans que les autorités nationales n'en soient informées, ou seulement si une victime engage une procédure. La suppression du contenu ou son invisibilisation peut en effet être nécessaire pour mettre fin au préjudice subi par la victime. Cependant, il peut alors être difficile pour les autorités de lancer une enquête en l'absence de connaissance ou de preuve sur l'objet du litige. De même, les victimes sont confrontées à la difficulté de rapporter le contenu, en particulier s'il a été supprimé.



MISE EN ŒUVRE PAR LES ÉTATS

- > Nous recommandons aux États d'identifier une autorité centralisée compétente pour recevoir et évaluer les contenus jugés illégaux par une plateforme lorsque cela est nécessaire, par exemple en cas de recours par un utilisateur. Cette autorité peut prendre la forme d'une autorité existante telle qu'un parquet spécialisé, d'une agence indépendante ou d'une nouvelle plateforme de signalement des comportements ou contenus illégaux sur Internet.
- > Si une agence indépendante se voit confier ce rôle, nous recommandons qu'une coopération entre cette agence et le procureur de la République soit établie, afin de faciliter les lancements d'enquête.
- > Nous recommandons de simplifier le degré de preuve exigée de la part des victimes qui signalent un contenu illégal qui a déjà été retiré. Cela pourrait être fait par une autorité judiciaire mettant en place une procédure de signalement gratuite.

4.1.3 ACCÈS À UN SYSTÈME JUDICIAIRE INDÉPENDANT

Nous recommandons que l'accès à une juridiction indépendante soit prévu pour les procédures civiles.

La juridiction *ratione loci* (c'est-à-dire dont la compétence est soumise aux limites territoriales) impose que le procès ait lieu là où la plateforme a son siège social, ce qui soulève des difficultés pour déterminer la compétence des tribunaux nationaux. Nous recommandons de ne pas tenir compte de ce critère, qui n'est pas adapté à l'activité internationale des plateformes. Cette approche est conforme à la jurisprudence des juridictions nationales.

En raison de la particularité de ce contentieux, qui nécessite des connaissances spécifiques et des ressources importantes, il est nécessaire que les juges impliqués dans ces contentieux soient spécialisés.

En raison de l'importance des dommages irréparables qui peuvent être causés par la diffusion d'un contenu, il est essentiel de permettre aux juges d'ordonner rapidement aux hébergeurs de retirer des contenus, ou au fournisseur d'accès à Internet (FAI) de bloquer l'accès à des contenus.

Cette procédure accélérée de retrait de contenu pourrait être initiée soit par le procureur, soit plus largement par toute personne y ayant un intérêt.

L'anonymat est un obstacle considérable à la mise en œuvre des procédures. Dans certaines circonstances, des procédures spécifiques doivent permettre d'obtenir l'identité de la personne dans un délai court.



MISE EN ŒUVRE PAR LES ÉTATS

- > Nous recommandons aux États de reconnaître le droit des utilisateurs de faire appel d'une décision prise par une plateforme à leur encontre devant un tribunal civil national.
- > Nous recommandons aux États d'exercer leur autorité sur la base d'un critère de *ratione personae*, en retenant la compétence de la juridiction nationale du lieu de résidence de la victime, ou de l'utilisateur qui a posté le contenu illégal.
- > Nous recommandons que, dans la mesure du possible, les États fassent usage de leur compétence universelle pour participer à la lutte contre l'impunité des crimes les plus graves.
- > Les juges nationaux chargés des litiges impliquant des réseaux sociaux devraient avoir des qualités particulières comme l'impartialité et l'indépendance.
- > Les États doivent accorder des ressources humaines et financières suffisantes aux institutions judiciaires.
- > Nous recommandons la création d'un tribunal civil indépendant spécialisé dans les litiges concernant les plateformes. Sa composition devrait inclure une représentation de la société civile.
- > Nous recommandons l'établissement de procédures sommaires (c'est-à-dire une procédure civile ou pénale menée sans formalités telles que des plaidoiries) pour le règlement rapide d'une affaire.
- > Les États devraient mettre en œuvre une procédure de demande d'identification d'un auteur de contenu litigieux.

4.1.4 PARQUET ET MAGISTRATS SPÉCIALISÉS

Compte tenu de la spécificité des litiges en question, nous recommandons la mise en place de magistrats spécialisés et de procureurs spécialisés.

4.2 NIVEAU INTERNATIONAL

UNE ENTITÉ TRANSNATIONALE

RÉGULATION	MÉCANISME DE RÈGLEMENT DES DIFFÉRENDS	ASSISTANCE AUX ENQUÊTES
1. Définir des normes pour les plateformes. 2. Identifier les circonstances exceptionnelles. 3. Analyse systémique des pratiques des plateformes afin de contribuer à l'amélioration de la réglementation. 4. Participer à la supervision des plateformes. 5. Superviser l'obligation des États à respecter la liberté d'expression. 6. Organe d'experts chargé de fournir des conseils à la demande des plateformes. 7. Aider à la formation des modérateurs.	Fournir aux plateformes un système de règlement des différends plus souple que les procédures judiciaires nationales pour déterminer l'application de la loi.	1. Collecter, consolider et préserver les informations et preuves de violations des droits de l'homme et d'abus. 2. Analyser les preuves collectées afin de faciliter la tenue de procédures pénales équitables et indépendantes. 3. Partager les informations et les éléments de preuve recueillis, ainsi que le travail d'analyse produit avec les organisations nationales, régionales et internationales.

L'espace mondial de l'information et de la communication est un bien commun de l'humanité qui doit être protégé et réglementé au niveau international, en tenant compte de la diversité culturelle mais aussi des principes communs acceptés et reconnus par ses différents participants.

Nous recommandons la création d'une entité transnationale, étant donné que les mécanismes nationaux pourraient ne pas répondre aux exigences imposées par les standards internationaux et qu'ils doivent être contrôlés pour s'assurer du respect des droits des détenteurs de compte. Cette entité transnationale aurait trois mandats distincts : la régulation internationale, un mécanisme de règlement des différends et une assistance en matière d'enquête.

Par ailleurs, le droit applicable aux plateformes, issu des standards internationaux des droits de l'homme, n'est pas suffisamment précis en l'état et devrait être précisé. Enfin, les plateformes doivent connaître clairement le droit applicable afin de procéder à une juste modération des contenus illégaux.

4.2.1 LA RÉGULATION INTERNATIONALE

L'organe transnational devrait être composé de représentants de tous les participants, dont des représentants des plateformes, de la société civile, des États et des organisations internationales intéressées, afin d'encourager la régulation tout en assurant la coordination entre les différentes parties prenantes.

En effet, l'élaboration de régimes de responsabilité nécessite un travail considérable que notre groupe de travail n'a pas les capacités de réaliser. Le présent rapport recommande donc que cette mission soit confiée à un organe permanent. Ce dernier pourra déterminer plus précisément les critères à établir pour appliquer des sanctions selon les circonstances ou lors de cas particuliers, afin de guider les plateformes et les autorités nationales dans la régulation des réseaux sociaux. Son caractère permanent lui permettra d'adapter la réglementation en fonction de l'évolution des situations.

Cette entité transnationale apportera son soutien aux plateformes pour la mise en œuvre des réglementations nationales et internationales.

Cette entité pourrait être créée au sein d'une organisation existante. Le Forum sur l'information et la démocratie dispose déjà de la structure nécessaire pour exercer les compétences requises pour promouvoir et contrôler la réglementation des plateformes conformément aux normes internationales.

Mise en place d'une entité transnationale

MANDAT

1. Conseiller et guider les plateformes

- Préciser les grands principes que les plateformes doivent respecter.
- Établir les moyens de mise en œuvre de ces principes par l'élaboration de normes et de lignes directrices.
- Guider les plateformes dans leur réponse à des situations exceptionnelles telles que la guerre.
- Fournir des conseils à la demande des plateformes.

2. Supervision des pratiques des plateformes

- Évaluer les décisions prises par les plateformes par le biais d'un examen par les pairs et/ou mener une enquête spécifique sur la base d'une plainte.
- Publier les résultats de ces évaluations dans un effort de transparence permettant, d'une part, aux plateformes d'accroître leur légitimité et, d'autre part, aux utilisateurs de prendre une décision éclairée sur le réseau social qu'ils souhaitent utiliser.

3. Contrôle de l'obligation des États à respecter la liberté d'expression

- Publier, de manière régulière, un classement des États en fonction de leur respect des droits de l'homme et des libertés fondamentales propres aux journalistes et aux défenseurs des droits de l'homme. Ce classement aiderait les plateformes à déterminer si elles peuvent révéler l'identité d'un détenteur de compte.
- Étudier l'activité des autorités de régulation nationales, afin de vérifier qu'elles bénéficient des garanties indispensables pour exercer leurs fonctions de manière indépendante, impartiale et transparente.

GOUVERNANCE

L'entité transnationale devrait comprendre des représentants de toutes les parties prenantes de l'espace de l'information et de la communication.

Il doit être composé des organes suivants :

- Un **organe plénier** réunissant des représentants de toutes les parties concernées.

- Un **organe administratif** composé de personnes ayant une expertise dans le domaine, respectant une représentation géographique et de genre équitable.
- Un **organe d'experts** qui devrait être mobilisé pour superviser les plateformes et être consulté lors des procédures de conception de normes et de lignes directrices.
 - > Les membres de l'organe d'experts doivent être élus par l'organe plénier pour leurs compétences spécialisées et respecter un principe d'indépendance.

4.2.2 RÈGLEMENT DES DIFFÉRENDS INTERNATIONAUX

Des affaires célèbres ont souligné l'importance de fournir aux plateformes un mécanisme de règlement des différends plus souple que les procédures judiciaires nationales pour déterminer l'application de la loi. Nous recommandons la création d'un mécanisme international de règlement des différends fondé sur les principes de l'arbitrage.



MISE EN ŒUVRE PAR LES ÉTATS

- > **Nous recommandons la création d'un système d'arbitrage international⁴¹** au sein de l'entité de régulation transnationale.
- > **Il devrait fournir :**
 - une liste d'**arbitres** : cette liste devrait être établie en tenant compte de leurs **compétences**, et pourrait être approuvée annuellement par l'organe plénier.
 - un **guide de procédure** : l'organisation doit proposer un guide de procédure que les parties pourraient ensuite adapter.
- > Les parties devraient s'engager à ce que la décision soit contraignante.

Assistance aux enquêtes

Les informations publiées sur les plateformes peuvent être utilisées comme preuves de violations des droits de l'homme ou de crimes. Cependant, l'obligation de modération incombant aux plateformes peut conduire à la perte de ces preuves.

Même si les plateformes acceptent de conclure des accords avec les mécanismes d'enquête internationaux et de transmettre les informations postées sur leurs services, ces accords sont ponctuels et limités dans le temps. Ils n'empêchent pas la perte de preuves en l'absence de mécanismes internationaux.

La capacité d'assistance en matière d'enquête de l'organisme international devrait servir d'organe chargé d'établir les faits et d'aider à trancher les affaires, ou encore d'aider d'autres organismes, nationaux ou internationaux. Cela devrait se faire en lien avec le processus d'arbitrage et d'autres organismes déjà existants, ou à créer, chargés de rassembler des preuves sur les crimes.

41 Strowel, Alain & Laura Somaini (2021).



MISE EN ŒUVRE PAR LES ÉTATS

- > **Nous recommandons la création d'un organe d'assistance aux enquêtes** au sein de l'entité transnationale.
- > Cet organe devrait être composé de personnes compétentes : juristes, chercheurs et analystes, entre autres, recommandées par un responsable de l'organe transnational indépendant et élu par l'organe plénier.
- > Cet organe devrait se voir confier le mandat suivant :
 - **recueillir, consolider et conserver** les informations et les preuves de violations et d'abus des droits de l'homme ;
 - **analyser** les preuves collectées et préparer les dossiers afin de faciliter et d'accélérer les procédures pénales équitables et indépendantes ;
 - **partager** les informations et les preuves collectées ainsi que le travail analytique produit avec les tribunaux nationaux, régionaux et internationaux.

ANNEXE 1

RÉGIME DE RESPONSABILITÉ EN COURS D'ÉLABORATION EN EUROPE : LE DIGITAL SERVICES ACT (DSA)

Le présent rapport souligne la nécessité de mettre en place des régimes de responsabilité pour les réseaux sociaux et leurs utilisateurs et propose une voie à suivre. L'Union européenne s'est engagée dans cette voie avec la présentation de la législation sur les services numériques (plus connue sous son nom anglais de Digital Services Act, DSA). Le DSA constitue un cadre réglementaire sans précédent qui fixera des normes sur la responsabilité des plateformes numériques concernant les contenus illégaux et néfastes et qui assurera une meilleure protection aux utilisateurs d'Internet et à leurs droits fondamentaux. Le DSA présente une des approches réglementaires possibles dans ce domaine, et pourrait servir de modèle à d'autres juridictions, à mesure que leurs pays progressent dans la mise en œuvre de garanties démocratiques dans l'espace de l'information et de la communication.

Le DSA concernera les sujets suivants, parmi ceux qui sont également traités par le présent rapport :

Pour les utilisateurs :

- La création de procédures claires et faciles pour signaler des contenus illégaux.
- Les utilisateurs auront accès à des mécanismes de règlements des différends dans leur pays.
- De nouvelles protections pour les mineurs.
- L'accès aux données des plateformes pour les chercheurs afin de comprendre les risques pour la société et les droits fondamentaux.
- Le droit des utilisateurs de refuser les recommandations de contenu basées sur le profilage.

Pour les plateformes :

- La transparence des conditions générales d'utilisation pour les plateformes.
- Des mécanismes de réponse rapide aux crises, avec des mesures supplémentaires de gestion des risques pour les crises de santé publique et de sécurité.
- Des audits indépendants sur leur gestion des risques, dont leurs systèmes algorithmiques.
- La transparence des règles de modération des contenus.
- Des informations claires sur les raisons pour lesquelles un contenu est recommandé aux utilisateurs.

REMERCIEMENTS

Le Forum tient à remercier chaleureusement les membres du comité de pilotage de ce groupe de travail, notamment ses co-présidents, ainsi que les plus de 50 experts que nous avons interrogés, pour leur temps, leurs avis éclairants et leur soutien.

- Camille Alloing, professeur de relations publiques, Université du Québec
- Enrique Armijo, professeur de droit à la Elon University School of Law et membre du corps professoral au Yale Law School Society Project.
- Farzaneh Badiei, fondatrice, Digital Medusa
- Romain Badouard, maître de conférences, Paris 1
- Ian Barber, responsable juridique, Global Partners Digital
- Dominique Cardon, directeur, Médialab de Sciences Po
- Anupam Chander, professeur en droit et technologie, Scott K. Ginsburg, Georgetown University Law Center
- Paul Coppin, responsable du service juridique, Reporters sans frontières (RSF)
- Alexandre Coutant, professeur de communication sociale et publique, Université du Québec
- Vera Eidelman, avocate, Speech, Privacy, and Technology Project de l'ACLU
- Olivier de Frouville, directeur du Centre de Recherche sur les Droits de l'Homme et le Droit Humanitaire de Paris (CRDH)
- Anna D. Gibson, doctorante, Département de communication, Université de Stanford
- Eric Goldman, professeur de droit, faculté de droit de l'université de Santa Clara
- Vincent F. Hendricks, professeur et directeur du Center for Information and Bubble studies, Université de Copenhague
- Matias Jackson, chercheur juridique et consultant en droit de la technologie et de l'internet, Université de la République d'Uruguay
- Uruguay
- David Kaye, ancien rapporteur spécial sur la liberté d'expression et d'opinion (2014-2020)
- Julien Kahn, avocat, JK Avocats
- Théophile Lenoir, docteur en droit, Université de Leeds
- Richard Malka, avocat, Malka & Associés
- Barthélémy Michalon, docteur en droit, Sciences Po
- Charles Miraillé, conseiller juridique, Commission nationale consultative des droits de l'homme
- Gabriel Nicholas, chargé de recherche, Center for Democracy & Technology
- Rasmus Nielsen, directeur, Institut Reuters pour l'étude du journalisme
- Javier Pallero, directeur politique, Access Now
- Courtney Radsch, directrice du plaidoyer, Committee to Protect Journalists (CPJ)
- Aurore Raoux, responsable des affaires publiques, News Media Europe
- Emilie Rannou, responsable de la stratégie et de la science des données, Ekimetrics
- Alain Strowel, professeur, Université Saint-Louis et UCLouvain
- Nicolas Suzor, professeur, département de droit de l'Université de technologie du Queensland
- Hélène Tigroudja, professeur de droit international public, Université Aix Marseille
- Erik Tuchtfield, chargé de recherche, Institut Max Planck de droit public comparé et de droit international
- Serena Villata, chercheuse titulaire, Chaire CNRS en intelligence artificielle
- Wout van Wijk, directeur, News Media Europe
- Richard Wingfield, responsable juridique, Global Partners Digital
- Jillian York, directrice pour la liberté d'expression internationale, Electronic Frontier Foundation
- Célia Zolynski, membre, Comité national de pilotage pour l'éthique numérique

Pour renforcer la participation d'experts de différentes régions du monde au groupe de travail, le Forum a conclu un partenariat avec six organisations locales de la région Moyen-Orient-Afrique du Nord et en Afrique de l'Ouest.

- La Fondation Maharat, une organisation non gouvernementale pour la liberté d'expression et le développement des médias, basée au Liban
- Le Groupe de travail pour le journalisme et la démocratie, une organisation qui promeut le journalisme et le développement des médias numériques en Tunisie.
- L'Association de lutte contre le racisme, l'ethnocentrisme et le régionalisme (ALCRER), une organisation non gouvernementale de défense des droits de l'homme, de la démocratie et de la bonne gouvernance au Bénin.
- Le Réseau des professionnels de la presse en ligne de Côte d'Ivoire (REPPRELICI), première organisation professionnelle de la
- la presse numérique en Côte d'Ivoire
- Le Syndicat des professionnels de l'information et de la communication du Sénégal (SYNPICS), une organisation professionnelle en charge du secteur des médias au Sénégal
- Nouveaux Droits de l'Homme (NDH-Burkina), association de défense des droits humains et de la liberté de la presse au Burkina Faso.

Grâce à leur soutien, 53 experts ont été mobilisés et ont participé aux ateliers pour élaborer les recommandations présentées dans ce rapport :

Liban :

- Layal Bahnam, responsable de programme, Fondation Maharat
- Dr Ali Rammal, professeur de médias, Université libanaise
- Dr Georges Sadaka, professeur de médias et ancien doyen de la faculté des médias, Université Libanaise
- Abed Kataya, responsable du contenu numérique, SMEX pour les droits numériques
- Dr Maria Abou Zeid, professeure associée et présidente du département d'études des médias de l'Université Notre-Dame-de-Louaizé
- Zeina Bou Harb, responsable de la coopération internationale, OGERO Telecom, qui dirige actuellement le secrétaire général du Lebanese Internet Governance Forum
- Abir Chebaro, spécialiste des questions de genre et ancien conseiller du Premier ministre pour les affaires relatives aux droits des femmes
- Layal Jebran, expert en technologie et entrepreneur, PDG et cofondateur de Moubarmij, ancien membre du conseil exécutif de l'Internet Society Lebanon
- Charbel Chbeir, avocat, membre du Barreau de Beyrouth et rapporteur du Centre sur l'informatique et la technologie du Barreau de Beyrouth
- Layal Sakr, avocat et fondateur de l'ONG SEEDS for Legal Initiatives
- Tony Mikhael, avocat et expert juridique de l'ONG Fondation Maharat

Tunisie :

- Neji Bghouri, chef du département des projets de l'UNESCO à Tunis et ancien président de l'Union nationale des journalistes tunisiens.
- Dr. Sadok Hammami, professeur à l'Institut de la presse et des sciences de l'information et membre du conseil de la presse
- Dr. Nouha Belaid, docteur en sciences de l'information et de la communication et professeur à l'Université Centrale de Tunis
- Karim Wannes, journaliste à Mosaïque FM, responsable de la déontologie au Syndicat national des journalistes tunisiens
- Rabeb Aloui, rédacteur en chef du média électronique spécialisé dans le fact-checking, BN Check

- Lasaad Dahech, expert auprès de l'ASBU (Arab States Broadcasting Union) et ancien PDG de la télévision nationale
- Réal Barnabé, expert en médias pour la Commission européenne, ancien chef de Radio Canada
- Tawfiq Omrane, journaliste et caricaturiste

Bénin :

- Gervais Loko, expert en gouvernance et responsable de programme à ALCRER
- Hugues Tossa, juriste spécialisé sur le numérique et chef de projet pour l'ONG Changement Social Bénin
- Renaud Sédjro Akpahouka, expert en conception pour CivicTech Solutions et responsable TIC à l'ONG Social Watch Bénin
- Maurice Thantan, chef de projet numérique et ancien président de l'Association des blogueurs du Bénin
- Donald Azokli Assogba, expert SEO (Search Engine Optimization)
- Hadjara Idris, expert digital et PDG de Mara Academy
- Constant Langnito, expert consultant Web et Mobile
- Joël Arsène Noumonvi, président de l'Association des blogueurs du Bénin
- Hervé Mekoun, sociologue et professeur d'université

Côte d'Ivoire :

- Lassina Serme, président du Réseau des professionnels de la presse en ligne de Côte d'Ivoire (REPPRELICI)
- Kébé Mamady, rédacteur en chef de la plateforme contre les fake news Ivoirecheck
- Marie Flore Begou, coordinatrice de l'ONG Citoyens engagés pour le leadership et la démocratie en Côte d'Ivoire (CELDICI),
- Marlyse Konan, présidente de l'Association des Femmes Journalistes de Côte d'Ivoire (AFJCI)
- Dr Karim Wally, enseignant-chercheur à l'Université Félix Houphouët-Boigny d'Abidjan
- Aboubacar Cidiq Sawadogo, président de l'Initiative citoyenne pour l'éducation aux médias sociaux en Côte d'Ivoire (ICEM-CI).
- Evelyne Déba, journaliste
- Wally Karim, directeur de la rédaction de Nord-Sud Quotidien
- Elisée Bolougbeu, directeur de publication de Destinationci

Sénégal :

- Bamba Kassé, journaliste et président de SYNPICS
- Adama Sow, journaliste à la Commission de protection des données personnelles du Sénégal (CDP)
- Moustapha Fall, avocat spécialisé en droit public, rédacteur du projet de loi sur la régulation des réseaux sociaux
- Mamadou Thior, président du Cored (Conseil de régulation de l'éthique et de la déontologie)
- Ibrahima Lissa Faye, président de l'APPEL (Association pour la promotion de la presse en ligne)
- Emmanuel Maurice Diokh, techno-pédagogue, formateur juridique et président d'Internet Sans Frontières
- Aminatou Diop, journaliste membre du CORED (Conseil pour l'observation des règles d'éthique et de déontologie dans les médias au Sénégal).
- Jean Louis Correa, professeur de droit associé à l'Université virtuelle du Sénégal

Burkina Faso :

- Idrissa Birba, journaliste, rédacteur en chef du journal Le Quotidien et président de NDH-Burkina.
- Boukari Bogré, enseignant et journaliste
- Frédéric Ilboudo, journaliste et chargé de communication à l'ONG Hunger Free World (HFW)
- Daouda Kinda, journaliste, responsable de la presse en ligne, Le Monde Rural
- Madi Bazié, journaliste, responsable de la presse en ligne, Ouest-info
- Élie Kaboré, journaliste au journal L'économiste du Faso
- Philippe Bationo, journaliste, responsable du média en ligne Le Talon
- Zalle Ibrahim, journaliste

BIBLIOGRAPHIE

- Access Now (2020). 26 Recommendations on Content Governance. Consulté à l'adresse : <https://www.accessnow.org/cms/assets/uploads/2020/03/Recommendations-On-Content-Governance-digital.pdf>
- ACLU (2021). The Problem With Censoring Political Speech Online – Including Trump's. Consulté à l'adresse : <https://www.aclu.org/news/freespeech/the-problem-with-censoring-political-speech-online-including-trumps/>
- American bar association (2018). The Ongoing Challenge to Define Free Speech. Consulté à l'adresse : https://www.americanbar.org/groups/crsj/publications/human_rights_magazine_home/the-ongoing-challenge-to-define-free-speech/
- Article 19 (2021). At a glance: Does the EU Digital Services Act protect freedom of expression? Consulté à l'adresse : <https://www.article19.org/resources/does-the-digital-services-act-protect-freedom-of-expression/>
- Aspen Institute (2021). Commission on Information Disorder. Consulté à l'adresse : https://www.aspeninstitute.org/wp-content/uploads/2021/11/Aspen-Institute_Commission-on-Information-Disorder_Final-Report.pdf
- Atlantic Council (2021). Free speech and online content: What can the US learn from Europe? Consulté à l'adresse : <https://www.atlantic-council.org/blogs/new-atlanticist/free-speech-and-online-content-what-can-the-us-learn-from-europe/>
- Badouard, Romain (2021). Les enjeux de la modération des contenus sur le web. Consulté à l'adresse : <https://la-rem.eu/2021/11/les-enjeux-de-lamoderation-des-contenus-sur-le-web/>
- Banko, Michele, Brendon MacKeen & Laurie Ray (2020). A Unified Typology of Harmful Content. Consulté à l'adresse : <https://aclanthology.org/2020.alw-1.16.pdf>
- Bayer, J.A (2019). The recommended responsibilities and duties of social media platform companies. Consulté à l'adresse : https://www.researchgate.net/publication/344785231_The_recommended_responsibilities_and_duties_of_social_media_platform_companies
- Berinsky, A. J (2015). Rumors and Health Care Reform: Experiments in Political Misinformation, Cambridge University Press, Vol. 47, pp.241 - 262. Consulté à l'adresse : <https://www.cambridge.org/core/journals/british-journal-of-political-science/article/rumors-and-health-care-reform-experiments-in-political-misinformation/8B88568CD057242D2D97649300215CF2>
- Boyd, Danah M. & Nicole B. Ellison (2007). Social Network Sites: Definition, History, and Scholarship, Journal of Computer-Mediated Communication, Vol. 13, pp. 210–230. Consulté à l'adresse : <https://academic.oup.com/jcmc/article/13/1/210/4583062>
- Brennan Center for Justice (2021). Double Standards in Social Media Content Moderation. Consulté à l'adresse : https://www.brennancenter.org/sites/default/files/2021-08/Double_Standards_Content_Moderation.pdf
- Buiten, Miriam, (2021). The Digital Services Act: From Intermediary Liability to Platform Regulation, SSRN. Consulté à l'adresse : https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3876328
- Burgorgue-Larsen, Laurence (2008). Les nouvelles tendances dans la jurisprudence de la Cour interaméricaine des droits de l'homme, Cursos de Derecho Internacional y Relaciones Internacionales de Vitoria Gasteiz, Universidad del País Vasco, pp. 149-180. Consulté à l'adresse : <https://hal.archives-ouvertes.fr/hal-01744225/document>
- Business for Social Responsibility (2021). A Human Rights-Based Approach to Content Governance. Consulté à l'adresse : https://www.bsr.org/reports/A_Human_Rights-Based_Approach_to_Content_Governance.pdf
- Cassella, Sarah, (2018). Rapport introductif, Le standard de due diligence et la responsabilité internationale. Consulté à l'adresse : <https://www.sfdi.org/publications/le-standard-de-due-diligence-et-la-responsabilite-internationale-journee-detude-franco-italienne-du-mans/>
- CIGI (2021). How “Free Speech” Kills Internet Regulation Debates: Part Two. Consulté à l'adresse : <https://www.cigionline.org/articles/how-free-speech-kills-internet-regulation-debates/>
- CNIL (2019). La forme des choix, Cahier IP innovation & Prospective, Num.06. Consulté à l'adresse : https://www.cnil.fr/sites/default/files/atoms/files/cnil_cahiers_ip6.pdf
- CNIL (2022). Contrôle de l'âge sur les sites web: la CNIL invite à développer des solutions plus efficaces et respectueuses de la vie privée. Consulté à l'adresse : <https://www.cnil.fr/fr/controle-de-lage-sur-les-sites-web-la-cnil-invite-developper-des-solutions-plus-efficaces-et>
- Common, MacKenzie, (2019). The Importance of Appeals Systems on Social Media Platforms, LSE Law - Policy Briefing Paper, Vol. 40, pp. 1-5. Consulté à l'adresse : https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3462770
- Congressional Research Service (2019). Free Speech and the Regulation of Social Media Content. Consulté à l'adresse : <https://sgp.fas.org/crs/misc/R45650.pdf>
- Congressional Research Service (2021). Social Media: Misinformation and Content Moderation Issues for Congress. Consulté à l'adresse : <https://crsreports.congress.gov/product/pdf/R/R46662>
- Conseil de l'Europe (2014). La prééminence du droit sur l'internet et dans le monde numérique en général. Consulté à l'adresse : <https://rm.coe.int/la-preeminence-du-droit-sur-l-internet-et-dans-le-monde-numerique-en-g/16808b3e6f juin 2021>
- Council of Europe (2020). Prioritization uncovered: the Discoverability of Public Interest Content Online <https://rm.coe.int/publication-content-prioritisation-report/1680a07a57>
- Council of Europe (2021). Content Moderation: Best Practices towards Effective Legal and Procedural Frameworks for Self-regulatory and Co-regulatory Mechanisms of Content Moderation. Consulté à l'adresse : <https://edoc.coe.int/fr/internet/10198-content-moderation-guidance-note.html>
- D'Ambrosio, Luca (2020). Le devoir de vigilance: une innovation juridique entre continuité et ruptures, Droit et Société, Vol. 106, pp. 633-647. Consulté à l'adresse : <https://www.cairn.info/revue-droit-et-societe-2020-3-page-633.htm>
- Data & Society (2018). Content or context moderation: Artisanal, Community-Reliant, and Industrial Approaches. Consulté à l'adresse : https://datasociety.net/wp-content/uploads/2018/11/DS_Content_or_Context_Moderation.pdf

Desmaris, Sacha, Pierre Dubreuil & Benoît Loutrel (2019), Créer un cadre français de responsabilisation des réseaux sociaux: agir en France avec une ambition européenne. Consulté à l'adresse : <https://la-rem.eu/2021/11/les-enjeux-de-la-moderation-des-contenus-sur-le-web/>

Douek, Evelyn (2020). The Rise of Content Cartels. Consulté à l'adresse : <https://knightcolumbia.org/content/the-rise-of-content-cartels>

Douek, Evelyn (2021). Governing online speech: from "post-as-trump" to proportionality and probability, Columbia Law Review, Vol. 121, pp. 759-834. Consulté à l'adresse : <https://columbialawreview.org/content/governing-online-speech-from-posts-as-trumps-to-proportionality-and-probability/>

Dugoin-Clément, Christine, (2021). Donald Trump banni des grands réseaux sociaux: et maintenant ? Consulté à l'adresse : <https://theconversation.com/donald-trump-banni-des-grands-reseaux-sociaux-et-maintenant-152950>

ECFR (2020). Regulation and accountability: How to save the internet. Consulté à l'adresse : https://ecfr.eu/article/commentary_regulation_and_accountability_how_to_save_the_internet/

Electronic Frontier Foundation (2021). Beyond Platforms: Private Censorship, Parler, and the Stack. Consulté à l'adresse : <https://www.eff.org/deeplinks/2021/01/beyond-platforms-private-censorship-parler-and-stack>

Electronic Frontier Foundation (EFF). How To Appeal. Consulté à l'adresse : <https://www.onlinecensorship.org/en/pages/1214ccb8-e1d5-472a-91e8-745d895b4dff>

Electronic Frontier Foundation (2021). UN Human Rights Committee Criticizes Germany's NetzDG for Letting Social Media Platforms Police Online Speech. Consulté à l'adresse : <https://www.eff.org/deeplinks/2021/11/un-human-rights-committee-criticizes-germanys-netzdg-letting-social-media>

Ermoshina, Ksenia (2022). Safe space by design? Federated architectures and alternative socio-technical models for content moderation, Alternative content regulation on social media. Consulté à l'adresse : <https://carism.u-paris2.fr/fr/pre-conference-ica>

Ertzscheid, Olivier, (2017). L'appétit des géants. Pouvoir des algorithmes, ambitions des plateformes, C&F éd.

European Federation of Journalists (1971). Munich Declaration of the Duties and Rights of Journalists. Consulté à l'adresse : <https://training.rsf.org/appendix-iii-declaration-of-rights-and-obligations-of-journalists/>

European Liberal Forum (2022). The new Digital Service Act and the protection of freedom and rights in the digital public square. Consulté à l'adresse : <https://liberalforum.eu/2022/01/digitising-europe-the-new-digital-service-act-and-the-protection-of-freedom-and-rights-in-the-digital-public-square/>

European Parliament (2022). Digital Services Act. Consulté à l'adresse : https://www.europarl.europa.eu/doceo/document/TA-9-2022-0269_EN.html

European Parliamentary Research Service (2021). Liability of online platforms. Consulté à l'adresse : [https://www.europarl.europa.eu/RegData/etudes/STUD/2021/656318/EPRS_STU\(2021\)656318_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2021/656318/EPRS_STU(2021)656318_EN.pdf)

European Union (2019), Directive (EU) 2019/790. Consulté à l'adresse : <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32019L0790>

European Union (2000), Directive 2000/31/EC. Consulté à l'adresse : <https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=celex%3A32000L0031>

Fagan, Frank, (2018). Systemic social media regulation? Duke Law & Technology Review, Vol. 16, pp. 393. Consulté à l'adresse : https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3160781

Forum on Information & Democracy: Working group on infodemics (2020). How to end infodemics. Consulté à l'adresse : https://informationdemocracy.org/wp-content/uploads/2020/11/ForumID_Report-on-infodemics_101120.pdf

Forum on Information & Democracy: Working group on the sustainability of journalism (2021). A New Deal for journalism. Consulté à l'adresse : https://informationdemocracy.org/wp-content/uploads/2021/06/ForumID_New-Deal-for-journalism_16Jun21.pdf

Fukuyama, Francis, Barak Richman & Ashish Goel (2021). How to Save Democracy From Technology. Consulté à l'adresse : <https://www.foreignaffairs.com/articles/united-states/2020-11-24/fukuyama-how-save-democracy-technology>

Gillespie, Tarleton, (2018). Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions that Shape Social Media, Yale University Press. Consulté à l'adresse : <https://yalebooks.yale.edu/book/9780300173130/custodians-internet>

GPPI (2017). Internet Companies Cannot Be Judges of Free Speech. Consulté à l'adresse : <https://www.gppi.net/2017/04/20/internet-companiescant-be-judges-of-free-speech>

Grimmelmann, James, (2019). Internet Law: cases & problems, Semaphore Press.

G'sell, Florence, (2021). Le conseil de surveillance de Facebook et l'affaire Trump les limites du concept de 'tribunal d'entreprise'. La semaine du droit, éd. générale. n22°.

Guéniat, Marc, (2020). Les géants du numérique deviennent les arbitres de la liberté d'expression. Consulté à l'adresse : <https://www.heidi.news/innovation-solutions/les-reseaux-sociaux-mettent-les-internautes-dans-des-bulles-d-enfermement-culturelles>

Hale, Thomas, David Held (2011). Handbook of Transnational Governance Institutions & Innovations, Polity Press.

Haupt, Claudia E, (2021). Regulating Speech Online: Free Speech Values in Constitutional Frames, Washington University Law Review. Consulté à l'adresse : https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3794884#

Hausalter, Louis, (2021). Liberté d'expression et pluralisme: face à Twitter et Facebook, l'impuissance des politiques. Consulté à l'adresse : <https://www.marianne.net/politique/liberte-d-expression-et-pluralisme-face-a-twitter-et-facebook-l-impuissance-des-politiques>

Heldt, Amélie & Stephan Dreyer (2021). Competent Third Parties and Content Moderation on Platforms: Potentials of Independent Decision-Making Bodies From A Governance Structure Perspective, Journal of Information Policy, Vol. 11, pp. 266-300. Consulté à l'adresse : https://leibniz-hbi.de/uploads/media/default/cms/media/6134zt1_Heldt-Amelie_jinfopoli.11.2021.0266.pdf

Heinrich-Böll-Stiftung (2019). Tackling hate speech requires the protection of freedom of expression. Consulté à l'adresse : <https://ng.boell.org/en/2019/12/17/tackling-hate-speech-requires-protection-freedom-expression>

Hennebel, Ludovic, H  l  ne Tigroudja (2018). Trait   de droit international des droits de l'homme, Pedone.

Hirgen, Bezhani, (2021). Online Platforms and Freedom of Expression: The EU's Digital Services Act. Consulté à l'adresse : <https://www.duo.uio.no/handle/10852/89284>

Hoog, Emmanuel, (2019). Confiance et liberté - Vers la création d'une instance d'autorégulation et de médiation de l'information. Consulté à l'adresse : <https://www.culture.gouv.fr/Espace-documentation/Rapports/Confiance-et-liberte-Vers-la-creation-d-une-instance-d-autoregulation-et-de-mediation-de-l-information>

Horwitz, Jeff (2021). Facebook Says Its Rules Apply to All. Company Documents Reveal a Secret Elite That's Exempt, The Wall Street Journal. Consulté à l'adresse : <https://www.wsj.com/articles/facebook-files-xcheck-zuckerberg-elite-rules-11631541353>

Indian government (2020). Central government Act, Section 79 in the Information Technology Act.

Institut Montaigne (2020). Disinformation: The German Approach and What to Learn From It. Consulté à l'adresse : <https://www.institut-montaigne.org/en/blog/disinformation-german-approach-and-what-learn-it>

Institut Montaigne (2021). Internet: le péril jeune ? Consulté à l'adresse : https://www.institutmontaigne.org/ressources/pdfs/publications/rapport-internet-le-peril-jeune_0.pdf

Institute for Governance and Policy Studies (2021). Regulating Harmful Communication: Current legal frameworks. After Christchurch: Hate, harm and the limits of censorship working paper series. Consulté à l'adresse : https://www.wgtn.ac.nz/_data/assets/pdf_file/0008/1926593/WP-21-05-regulating-harmful-communication-legal-frameworks.pdf

Inter-American Dialogue (2021). Content Moderation and Self-Regulation Mechanisms: The Facebook Oversight Board and Its Implications for Latin America. Consulté à l'adresse : <https://www.foreignaffairs.com/reviews/capsule-review/2021-12-14/content-moderation-and-self-regulation-mechanisms-facebook>

Isaac, Henri & Louis-Victor de Franssu (2022). Les utilisateurs doivent être intégrés à la régulation numérique, Acteurs publics. Consulté à l'adresse: <https://acteurspublics.fr/articles/les-utilisateurs-doivent-etre-integres-a-la-regulation-numerique>

Isaac, Stanley-Becker, Elizabeth Dwoskin (2020). Trump allies, largely unconstrained by Facebook's rules against repeated falsehoods, cement pre-election dominance. Consulté à l'adresse : <https://www.washingtonpost.com/technology/2020/11/01/facebook-election-misinformation/>

Jourová, Věra (2021). Après Trump, il est urgent d'encadrer les géants du numérique, Libération. Consulté à l'adresse : https://www.libération.fr/debats/2021/01/22/apres-trump-il-est-urgent-d-encadrer-les-geants-du-numerique_1818153

Jurcys, Paulius, Christopher Donewald, Mark Fenwick, Markus Lampinen, Vytautas Nekrošius, & Andrius Smaliukas (2020). Ownership of User-Held Data: Why Property Law is the Right Approach, SSRN Electronic Journal. Consulté à l'adresse : <https://jolt.law.harvard.edu/assets/digestImages/Paulius-Jurcys-Feb-19-article-PJ.pdf>

Kaspar, Lea, Richard Wingfield (2021). Digital Rights at a Crossroads - recommendations for advancing human rights and social justice in the post-2020 era. Consulté à l'adresse : <https://www.gp-digital.org/wp-content/uploads/2021/03/Digital-Rights-at-a-Crossroads.pdf>

Kave, David. (2019). Speech Police. Columbia Global Reports.

Khan, Irene, (2021). Disinformation and freedom of opinion and expression Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression. Consulté à l'adresse : https://eos.cartercenter.org/uploads/document_file/path/985/Report_of_the_Special_Rapporteur_on_Disinformation_and_Freedom_of_Opinion_and_Expression_E.pdf

Khana, Tanveer, Antonis Michalalasab & Adnan Akhunzadac (2021). Fake news outbreak 2021: Can we stop the viral spread? Consulté à l'adresse : <https://www.sciencedirect.com/science/article/pii/S1084804521001326>

Lefoulli, Yassine & Leonardo Madio (2021). The Economics of Platform Liability. Consulté à l'adresse : https://www.tse-fr.eu/sites/default/files/TSE/documents/sem2021/eco_platforms/madio.pdf

Legros, Claire, (2021). La liberté d'expression à l'heure du numérique ou la difficile quête de l'équilibre sur les réseaux sociaux. Consulté à l'adresse : https://www.lemonde.fr/idees/article/2021/04/02/reseaux-sociaux-et-liberte-d-expression-inventer-des-dispositifs-pour-protger-nosdemocraties_6075320_3232.html

Lingibé, Patrick, (2018). Quelles responsabilités sur les réseaux sociaux ? Consulté à l'adresse : <https://www.village-justice.com/articles/quelles-responsabilites-sur-les-reseaux-sociaux,29033.html>

Loutrel, Benoît, Pierre Dubreuil & Sacha Desmaris (2019). Ce que peut l'Etat face aux plateformes. Consulté à l'adresse : <https://www.numerique.gouv.fr/uploads/rapport-mission-regulation-reseaux-sociaux.pdf>

Mercatus center (2020). Do Our Leaders Believe in Free Speech and Online Freedom Anymore? Consulté à l'adresse : <https://www.mercatus.org/bridge/commentary/do-our-leaders-believe-free-speech-and-online-freedom-anymore>

Musiani, Francesca, (2013). Governance by algorithm, Internet Policy Review, Vol.2. Consulté à l'adresse : <https://doi.org/10.14763/2013.3.188>

Natalegawa, Andreyka & Kyra Jasper (2021). Controlling the Information Space: Big Tech and Free Speech in Southeast Asia, CSIS. Consulté à l'adresse : <https://www.csis.org/blogs/new-perspectives-asia/controlling-information-space-big-tech-and-free-speech-southeast-asia>

Ndior, Valère, (2022). Les réseaux sociaux et l'administration Trump, Hérodote, Vol. 184-185, pp. 201 à 215. Consulté à l'adresse : <https://www.cairn.info/revue-herodote-2022-1-page-201.htm>

Nicholas, Gabriel, (2022). Shedding Light on Shadowbanning, Center for Democracy & Technology. Consulté à l'adresse : <https://cdt.org/wp-content/uploads/2022/04/remediated-final-shadowbanning-final-050322-upd-ref.pdf>

Observacom (2022). Twitter suspendió la cuenta un portal ecuatoriano de periodismo de investigación, luego de una publicación vinculada al narcotráfico. Consulté à l'adresse : <https://www.observacom.org/twitter-suspendio-la-cuenta-un-portal-ecuatoriano-de-periodismo-de-investigacion-luego-de-una-publicacion-vinculada-al-narcotrafico/irisspecial-2021-01fr-decrypter-la-legislation-sur-les-services-numer/1680a4d2ba>

OCDE (2021). Transparency reporting on terrorist and violent extremist content online: An update on the global top 50 content sharing services. Éditions OCDE. Consulté à l'adresse : <https://doi.org/10.1787/8af4ab29-en>

OFCOM (2021). Intermédiaires et plateformes de communication - Effets sur la communication publique et approches de gouvernance. Consulté à l'adresse : <https://www.bakom.admin.ch/bakom/fr/page-daccueil/suisse-numerique-et-internet/communication-numerique/plateformes-de-communication.html>

Office of the United Nations High Commissioner for Human Rights (1966). International Covenant on Civil and Political Rights. Consulté à l'adresse : <https://www.ohchr.org/sites/default/files/Documents/ProfessionalInterest/ccpr.pdf>

Oversight Board (2022). Interview Quasi-Judicial Oversight Mechanisms for Social Platforms. Consulté à l'adresse : <https://www.nomos-elibrary.de/10.5771/2699-1284-2021-3-254/quasi-judicial-oversight-mechanisms-for-social-platforms-a-conversation-with-catalina-botero-marino-co-chair-of-the-oversight-board-volume-2-2021-issue-3?page=1>

Pew Research Center (2017). The Future of Free Speech, Trolls, Anonymity and Fake News Online . Consulté à l'adresse : <https://www.pewresearch.org/internet/2017/03/29/the-future-of-free-speech-trolls-anonymity-and-fake-news-online/>

Pisillo-Mazzeschi, Riccardo (2008). Responsabilité de l'État pour violation des obligations positives relatives aux droits de l'homme, Collected Courses of the Hague Academy of International Law, Vol. 333. Consulté à l'adresse : http://dx.doi.org/10.1163/1875-8096_pplrdc_A9789004172845_02

Reisach, Ulrike (2021). The responsibility of social media in times of societal and political manipulation. European Journal of Operational Research. Consulté à l'adresse : <https://www.sciencedirect.com/science/article/pii/S0377221720308249>

Reporters sans frontières (2016). RSF deplores suspension of French journalist's Facebook account. Consulté à l'adresse : <https://rsf.org/en/rsf-deplores-suspension-french-journalist-s-facebook-account>

Reporters sans frontières (2021). India: RSF asks Twitter to restore The Kashmir Walla's account. Consulté à l'adresse : <https://rsf.org/en/india-rsf-asks-twitter-restore-kashmir-walla-s-account>

Rest of world NY media (2021). How social media companies help African governments abuse "disinformation laws" to target critics. Consulté à l'adresse : <https://restofworld.org/2021/social-media-africa-democracy/>

Revue internationale de droit économique (2019). Plateformes digitales et États: la corégulation par les données. Le cas des requêtes gouvernementales. Revue internationale de droit économique, Vol.33, pp. 363-381. Consulté à l'adresse : <https://www.cairn.info/revue-internationale-de-droit-economique-2019-3-page-363.htm>

Rikke, F.J. & Lumi Zuleta (2020). Private Governance of Freedom of Expression on Social Media Platforms: EU content regulation through the lens of human rights standards. Nordicom Review, Vol. 41, pp. 51-67. Consulté à l'adresse : <https://www.proquest.com/open-view/a07f1352f6804ab21b66ee2bd5cd2e9e/1?pq-origsite=gscholar&cbl=2040374>

Pen America (2021). No Excuse No Abuse: What Social Media Companies Can Do Now to Combat Online Harassment and Empower Users. Consulté à l'adresse : <https://pen.org/report/no-excuse-for-abuse/>

Sander, Barrie (2020). Freedom of Expression in the Age of Online Platforms: the Promise and Pitfalls of a Human Rights-Based Approach to Content Moderation. Fordham International Law Journal, Vol. 43, pp. 984-985

Stella, Elie, (2019). L'adaptation du droit pénal aux réseaux sociaux en ligne. Consulté à l'adresse : <https://hal.univ-lorraine.fr/tel-02985468/document>

Strowel, Alain & Laura Somaini (2021). Towards a robust framework for algorithmic transparency to tackle the dissemination of illegal and harmful content on online platforms: Proposals for the Digital Services Act and beyond. Consulté à l'adresse : https://cdn.uclouvain.be/groups/cms-editors-crides/droit-intellectuel/CRIDES_WP_2_2021_Alain%20Strowel%20and%20Laura%20Somaini.pdf

Sudre, Frédéric, Hélène Surrel, Laure Milano & Béatrice Pastre-Belda (2021). Droit européen et international des droits de l'homme, PUF.

Swedish Defence Research Agency (2019). Supervised Classification of Twitter Accounts Based on Textual Content of Tweets. Consulté à l'adresse : https://www.foi.se/download/18.46c3b221170ccb3336f1c6/1585826754727/Supervised-classification-of-twitter_FOI-S--6118--SE.pdf

Moore, Martin, and Tambini, Damian, (eds.) (2021). Regulating Big Tech: Policy Responses to Digital Dominance, Oxford University Press.

Timberg Craig, Elizabeth Dwoskin & Reed Albergotti (2021). Inside Facebook, Jan. 6 violence fueled anger, regret over missed warning signs. Consulté à l'adresse : <https://www.washingtonpost.com/technology/2021/10/22/jan-6-capitol-riot-facebook/>

The Heartland Institute (2021). Six Principles for State Legislators Seeking to Protect Free Speech on Social Media Platforms. Consulté à l'adresse : <https://www.heartland.org/publications-resources/publications/six-principles-for-state-legislators-seeking-to-protect-free-speech-on-social-media-platforms>

The Future of Free Speech (2021). A framework of first reference: Decoding a human rights approach to content moderation in the era of "platformization". Consulté à l'adresse : https://futurefreespeech.com/wp-content/uploads/2021/11/Report_A-framework-of-first-reference.pdf

The Future of Free Speech (2021). Rushing to Judgment: Are Short Mandatory Takedown Limits for Online Hate Speech Compatible with The Freedom of Expression? Consulté à l'adresse : https://futurefreespeech.com/wp-content/uploads/2021/01/FFS_Rushing-to-Judgment-3.pdf

The Paradigm Initiative (2020). Digital rights and data privacy in Nigeria. Consulté à l'adresse : https://ng.boell.org/sites/default/files/2020-08/Digital%20Rights%20and%20Privacy%20in%20Nigeria_0.pdf

The Santa Clara principles (2018). The Santa Clara Principles: On Transparency and Accountability in Content Moderation. Consulté à l'adresse : <https://santaclaraprinciples.org/>

The World Wide Web Consortium (2022). Privacy Principles. Consulté à l'adresse : <https://www.w3.org/TR/2022/DNOTE-privacy-principles-20220512/>

Tony Blair Institute (2021). Social Media Futures: What Is Brigading? Consulté à l'adresse : <https://institute.global/policy/social-media-futures-what-brigading>

Tsamados, Andreas, Nikita Aggarwal, Josh Cows, Jessica Morley, Huw Roberts, Mariarosaria Taddeo & Luciano Floridi (2022). The ethics of algorithms: key problems and solutions. AI & Soc, Vol. 37, pp. 215-230. Consulté à l'adresse : <https://doi.org/10.1007/s00146-021-01154-8>

Twitter (2019). World Leaders on Twitter: Principles & Approach. Consulté à l'adresse : https://blog.twitter.com/en_us/topics/com-

[pany/2019/worldleaders2019](#)

UN Human Rights Council (2018). Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression. Consulté à l'adresse : <https://digitallibrary.un.org/record/1631686/usage>

UNESCO (2018). Journalism, 'Fake News' and Disinformation: Handbook for Journalism Education and Training. Consulté à l'adresse : https://unesdoc.unesco.org/ark:/48223/pf0000265552_eng

United Nations (2011). Guiding Principles on Business and Human Rights: Implementing the United Nations "Protect, Respect and Remedy" Framework. Consulté à l'adresse : https://www.ohchr.org/sites/default/files/documents/publications/guidingprinciplesbusinesshr_en.pdf

United Nations (1969). Vienna Convention on the Law of Treaties. Consulté à l'adresse : https://legal.un.org/ilc/texts/instruments/english/conventions/1_1_1969.pdf

United States Congress (2020). Eliminating Abusive and Rampant Neglect of Interactive Technologies Act of 2019, S.3398, 116th. Consulté à l'adresse : <https://www.congress.gov/bill/116th-congress/senate-bill/3398/text>

Vengattil, Munsif & Elizabeth Culliford (2022). Facebook allows war posts urging violence against Russian invaders, Reuters. Consulté à l'adresse : <https://www.reuters.com/world/europe/exclusive-facebook-instagram-temporarily-allow-calls-violence-against-russians-2022-03-10/>

Vermeule, Adrian & Cass Sunstein (2009). Conspiracy Theories: Causes and Cures. The Journal of Political Philosophy, Vol.17, pp. 202-227. Consulté à l'adresse : <https://onlinelibrary.wiley.com/doi/10.1111/j.1467-9760.2008.00325.x>

Wong, Cynthia & James X. Dempsey (2011). The media and liability for content on internet. Consulté à l'adresse : <https://www.opensocietyfoundations.org/uploads/ab948019-f717-436d-bdb8-b45893e7af08/mapping-digital-media-liability-content-internet-20110926.pdf>

Soutenu par :



Forum Information & Démocratie

Contact : contact@informationdemocracy.org