



Typologie des modèles réglementaires des réseaux sociaux pour lutter contre les discours de haine en ligne

Rapport de recherche

Soumis en réponse à un appel à projet pour l'année 2023 dans le domaine du discours haine en ligne financé par l'Office fédéral de la communication

Auteurs du rapport

Prof. Dr. Sophie Weerts

sophie.weerts@unil.ch

Mathieu Fasel, MLaw / MA

mathieu.fasel@unil.ch

Université de Lausanne



Institut de hautes études en administration publique



Institut de hautes études en administration publique
Swiss Graduate School of Public Administration

Lausanne, février 2024

Typologie des modèles réglementaires des réseaux sociaux

Prof. Dr. Sophie Weerts & Mathieu Fasel

Université de Lausanne

Institut de hautes études en administration publique (IDHEAP)

Résumé :

Ce rapport de recherche propose une typologie des modèles réglementaires permettant de lutter contre les discours de haine en ligne, lesquels constituent un enjeu tant du point de vue des droits humains que de la démocratie. Cette typologie montre que la réponse réglementaire a évolué sur le plan temporel et territorial. Ainsi, les entreprises de réseaux sociaux ont initialement été considérées comme des destinataires directs de la réglementation (modèle d'origine américaine et fondé sur l'idée d'exonération de la responsabilité des entreprises de réseaux sociaux), pour ensuite être appréhendées comme des intermédiaires entre l'Etat et les utilisateurs (modèle d'origine européenne et fondé sur l'idée que les entreprises de réseaux sociaux doivent contribuer à la mise en œuvre de la lutte contre les discours de haine en ligne). Sur le plan territorial, le modèle européen de la réglementation des entreprises de réseaux sociaux présente aussi des nuances, avec des mesures qui ont été progressivement renforcées tant du point de vue des rapports entre entreprises de réseaux sociaux et autorités publiques qu'entre entreprises de réseaux sociaux et utilisateurs. Cette typologie résulte du croisement de deux types d'analyse sur deux types de données. Une première analyse diachronique est fondée sur l'analyse documentaire des travaux des rapporteurs sur la liberté d'expression et d'information des Nations Unies entre 1995 et 2019. Elle montre comment le discours de haine a été progressivement défini et quels instruments et approches ont été privilégiés pour lutter contre le discours de haine. La seconde analyse, de type synchronique, est fondée sur la comparaison de cinq législations (Allemagne, France, Autriche, dans l'Union Européenne et au Royaume-Uni) réglementant les entreprises de réseaux sociaux. Elle montre l'expansion des outils mobilisés pour faire des entreprises des acteurs de la régulation du discours en ligne, mais aussi pour encadrer cette activité. Au final, ce rapport permet d'éclairer les autorités qui envisageraient de réglementer le fonctionnement des entreprises de réseaux sociaux.

Mots-clés : *Discours de haine en ligne, réglementation des entreprises de réseaux sociaux, Typologie des modèles, Netzwerkdurchsetzungsgesetz, Loi Avia, Kommunikationsplattformengesetz, Digital Services Act, Online Safety Act*

Typologie der Regulierungsmodelle für soziale Netzwerke

Prof. Dr. Sophie Weerts & Mathieu Fasel

Universität Lausanne

Institut für öffentliche Verwaltung (IDHEAP)

Abstrakt:

Dieser Forschungsbericht enthält eine Typologie regulatorischer Modelle zur Bekämpfung von Online-Hassreden, die sowohl aus menschenrechtlicher als auch aus demokratischer Sicht eine Herausforderung darstellen. Diese Typologie zeigt, dass sich die regulatorischen Reaktionen zeitlich und territorial verändert haben. So wurden soziale Netzwerkunternehmen zunächst als direkte Adressaten der Regulierung betrachtet (ein Modell, das aus den USA stammt und das auf der Idee der Haftungsfreistellung für soziale Netzwerkunternehmen basiert ist), um dann als Vermittler zwischen dem Staat und den Nutzern verstanden zu werden (ein Modell, das aus Europa stammt und das auf der Idee basiert ist, dass soziale Netzwerkunternehmen zur Umsetzung der Bekämpfung von Online-Hassreden beitragen müssen). Auf territorialer Ebene weist das europäische Modell der Regulierung von Social-Networking-Unternehmen ebenfalls Nuancen auf, wobei die Maßnahmen sowohl im Hinblick auf die Beziehungen zwischen Social-Networking-Unternehmen und staatlichen Behörden als auch zwischen Social-Networking-Unternehmen und Nutzern schrittweise verstärkt wurden. Diese Typologie ergibt sich aus der Kreuzung von zwei Analysearten auf zwei Datentypen. Eine erste diachrone Analyse basiert auf der Dokumentenanalyse der Arbeiten der Berichtersteller der Vereinten Nationen für Meinungs- und Informationsfreiheit zwischen 1995 und 2019. Sie zeigt, wie Hassrede nach und nach definiert wurde und welche Instrumente und Ansätze zur Bekämpfung von Hassrede bevorzugt wurden. Die zweite Analyse ist synchron und basiert auf einem Vergleich von fünf Gesetzen (Deutschland, Frankreich, Österreich, in der Europäischen Union und im Vereinigten Königreich), die Unternehmen in sozialen Netzwerken regulieren. Sie zeigt die Ausweitung der Instrumente, die mobilisiert wurden, um die Unternehmen zu Akteuren der Regulierung des Online-Diskurses zu machen, aber auch um diese Aktivität selbst zu regulieren. Letztendlich kann dieser Bericht den Behörden, die eine Regulierung der Tätigkeit von Unternehmen, die soziale Netzwerke betreiben, in Erwägung ziehen, als Orientierungshilfe dienen.

Stichwörter: Online-Hassreden, Regulierung von Social-Media-Unternehmen, Typologie der Modelle, Netzwerkdurchsetzungsgesetz, Loi Avia, Kommunikationsplattformengesetz, Digital Services Act, Online Safety Act

Typology of regulatory models for social media

Prof. Dr. Sophie Weerts & Mathieu Fasel

University of Lausanne

Swiss Graduate School of Public Administration (IDHEAP)

Abstract:

This research report proposes a typology of regulatory models for combating online hate speech, an issue for both human rights and democracy. The typology shows that the regulatory response has evolved over time and in different territories. Initially, social media companies were seen as direct addressees of regulation (a model of American origin based on the idea that social media companies are exempt from liability). Later, they were seen as intermediaries between the State and users (a model of European origin based on the idea that social media companies should contribute to the fight against online hate speech). At a territorial level, the European model for regulating social media companies also has its nuances, with measures that have been progressively strengthened both from the point of view of relations between social media companies and public authorities and between social media companies and users. This typology results from the intersection of two types of analysis on two types of data. A first diachronic analysis is based on a documentary analysis of the work of the United Nations rapporteurs on freedom of expression and information between 1995 and 2019. It shows how hate speech has been progressively defined, and what instruments and approaches have been favoured to combat it. A second analysis, of a synchronic nature, is based on a comparison of five pieces of legislation (Germany, France, Austria, the European Union and the United Kingdom) regulating social media companies. It shows the expansion of tools mobilized to make companies players in the regulation of online discourse, but also to frame this activity. All in all, this report is a useful guide for authorities considering regulating the operations of social media companies.

Keywords: Online hate speech, Regulation of social media, companies, Typology of models, Netzwerkdurchsetzungsgesetz, Loi Avia, Kommunikationsplattformengesetz, Digital Services Act, Online Safety Act

Sommaire

1. Introduction	8
2. État des lieux : Discours de haine et environnement en ligne	11
2.1. Le discours de haine en droit international des droits humains	11
2.1.1. Dans les normes et standards onusiens.....	11
2.2.2. Dans les normes et standards du Conseil de l'Europe.....	15
2.2.3. Caractéristiques communes pour appréhender la notion de discours de haine	17
2.2. L'environnement numérique et l'approche économique de la réglementation des intermédiaires en ligne	18
2.1.1. Les acteurs de l'écosystème numérique	18
2.2.2. L'approche économique de la réglementation des entreprises en ligne : le régime juridique de responsabilité limitée pour les propos diffusés	19
3. Approches, données et méthode d'analyse	22
4. Conceptualisation du discours de haine et sa régulation à la lumière du droit international des droits humains	25
4.1. Les années 1990 et 2000 : l'Etat régulateur de la liberté d'expression et l'absence de définition du discours de haine	25
4.2. Au cours de la première moitié des années 2010 : l'émergence des entreprises comme intermédiaires de la régulation et la conception restrictive du discours de haine à l'aune du droit pénal international	29
4.3. Au cours de la seconde moitié des années 2010 : la montée en puissance des entreprises privées comme intermédiaires de la régulation et l'élargissement de la conception du discours de haine	34
5. Analyse des réglementations applicables aux entreprises en ligne	40
5.1. Allemagne : <i>Netzwerkdurchsetzungsgesetz</i> (NetzDG)	40
5.1.1. Contexte	41
5.1.2. Champ d'application	41
5.1.3. Règles organisationnelles applicables à l'activité des entreprises de réseaux sociaux.....	42
5.1.4. Règles de procédure applicables dans le rapport entre entreprises et utilisateurs.....	43
5.1.5. Autorités spéciales	44
5.2. France : Loi en vue de lutter contre les discours de haine sur les réseaux sociaux (Loi Avia)....	45
5.2.1. Contexte	45
5.2.2. Champ d'application	47
5.2.3. Règles organisationnelles applicables à l'activité des entreprises de réseaux sociaux.....	48
5.2.4. Règles de procédure applicables dans le rapport entre entreprises et utilisateurs.....	49
5.2.5. Autorités spéciales	49

5.3.	Autriche : <i>Kommunikationsplattformengesetz</i> (Ko-Pl)	49
5.3.1.	Contexte	50
5.3.2.	Champ d'application	51
5.3.3.	Règles organisationnelles applicables aux activités des entreprises de réseaux sociaux ...	51
5.3.4.	Règles de procédure applicables aux rapports entre entreprises et utilisateurs	52
5.3.5.	Autorités spéciales	52
5.4.	Union européenne : Règlement sur les services numériques (DSA)	53
5.4.1.	Contexte	53
5.4.2.	Champ d'application	54
5.4.3.	Règles organisationnelles relatives aux activités des entreprises de réseaux sociaux .	54
5.4.4.	Règles de procédure applicables aux rapports entre entreprises et utilisateurs	55
5.4.5.	Autorités spéciales	55
5.5.	Royaume Uni : Online Safety Act	56
5.5.1.	Contexte	56
5.5.2.	Champ d'application	59
5.5.3.	Règles organisationnelles relatives à l'activité des entreprises de réseaux sociaux	59
5.5.4.	Règles de procédure applicables aux rapports entre entreprises et utilisateurs	60
5.5.5.	Autorités spéciales	60
5.6.	Synthèse : mise en œuvre et tableau récapitulatif	61
6.	Conclusions et remarques prospectives	68
	Bibliographie	71

I. Introduction

L'essor d'Internet, à partir des années 1990, a ouvert un nouvel espace de diffusion d'informations et de communication¹ et l'arrivée de nouveaux acteurs, les entreprises de réseaux sociaux². L'utilisation de ces réseaux sociaux permet la propagation à très large échelle de propos appréciables mais aussi inacceptables du point de vue des valeurs de dignité humaine et d'égalité qui sont au cœur des droits humains. Dans le contexte de l'environnement en ligne, la notion de discours de haine est souvent utilisée, sans que l'on sache précisément ce que recouvre une telle notion. À ce stade, nous dirons ici que constitue un discours tout propos ou comportement qui cherche à décrire, convaincre, justifier ou expliquer. Ce discours présente une dimension de haine dès lors qu'il repose sur une vision négative des rapports interindividuels selon laquelle les individus ne seraient pas égaux entre eux en raison de caractéristiques physiques, culturelles, sociales ou politiques.

La propagation de propos attentatoires aux valeurs des droits humains dans l'écosystème d'Internet est particulièrement problématique, car l'expérience de la seconde guerre mondiale ou d'autres conflits plus récents ont montré que la diffusion de tels discours était précurseur de violences physiques contre des groupes ou des personnes appartenant à ces groupes³. Or, ce type de propos continue d'exister, connaît une accentuation dans l'environnement numérique et s'articule avec des logiques de désinformation et mésinformation. En l'occurrence, en Suisse, le dernier rapport du Comité pour l'élimination de la discrimination raciale (CERD) des Nations Unies, publié en 2021, a souligné comment des personnes d'ascendance africaine, des juifs, des musulmans et des Roms continuent d'être la cible de discours de haine⁴. Plus récemment, il a été montré que le phénomène touche aussi de manière très inquiétante les personnes de confession juive⁵. De tels discours ont des implications d'un point de vue individuel (ils portent atteinte à la dignité humaine, à l'intégrité psychique des personnes visées ainsi qu'à l'idée d'égalité entre les individus) et d'un point de vue collectif (ils

¹ Weber, R. H. (2021). *Internet Governance at the Point of No Return*. EIZ Publishing, pp. 2-3.

² Dans son étude préparatoire pour le Digital Services Act, la Commission européenne note l'explosion de l'importance d'Internet et des réseaux sociaux depuis les années 2000. Selon la Commission européenne, plus de 76% des européens utilisent des plateformes de partage de vidéos ou de musique, plus de 72% font du shopping en ligne et plus de 70% sont des utilisatrices et utilisateurs de réseaux sociaux. Cette situation marque un tournant majeur par rapport au début du millénaire et légitime l'adoption d'un nouveau cadre réglementaire. Commission européenne. (2020). *Impact assessment accompanying the proposal for a regulation of the European parliament and of the council*. Consulté le 15 janvier 2023. Disponible sous <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52020SC0348>.

³ Nations Unies. *Discours de haine*. Consulté le 5 septembre 2023. Disponible sous : <https://www.un.org/fr/hate-speech/understanding-hate-speech/what-is-hate-speech>

⁴ Nations Unies, Comité pour l'élimination de la discrimination raciale. (2023). Rapport CERD/C/CHE/10-12. Consulté le 7 septembre 2023. Disponible sous : <https://www.ungeneva.org/fr/news-media/meeting-summary/2021/11/examen-du-rapport-de-la-suisse-devant-le-cerd-la-montee-des>

⁵ Le Temps. (2024). Le nombre d'actes antisémites a explosé en 2023. Edition du 15 février 2024. Consulté le 15 février 2024. Disponible sous : [Le nombre d'actes antisémites a explosé en 2023 - Le Temps](#)

peuvent également décourager les personnes visées de participer aux débats publics⁶) et peuvent contribuer à un climat de violence à l'égard de groupes déterminés. Autrement dit, la libre circulation de ce type de discours constitue un problème public et, sur le plan juridique, peut violer ou restreindre les droits humains des personnes visées. D'un point de vue des droits fondamentaux, une telle situation requiert alors une intervention de l'Etat. Elle pose alors la question du type d'intervention étatique, et notamment de la pertinence de l'intervention du législateur et de ses modalités d'intervention.

La réglementation, la sanction pénale ou la réparation civile, constitue une réponse en matière de propos haineux. Toutefois, l'environnement internet génère de nouveaux défis du point de vue de la diffusion des discours de haine. Premièrement, la communication entre l'émetteur d'un message et son récepteur intègre un nouveau type d'acteurs, à savoir les entreprises de réseaux sociaux (ci-après les entreprises en ligne). Or, dans le paysage économique actuel, les entreprises en ligne se caractérisent par leur dimension transnationale, en ce sens que leur champ d'activité dépasse les frontières classiques de l'Etat dans lequel elles sont institués⁷. Les entreprises en ligne s'inscrivent aussi dans une logique décentralisée dès lors que l'information provient majoritairement des utilisatrices et utilisateurs eux-mêmes⁸. Cette décentralisation de la production de l'information remet en cause l'idée de contrôle éditorial qui accompagnait jusqu'ici la production de l'information, en principe réservée aux larges médias journalistiques ou de télévision⁹. Le fonctionnement de la technologie à la base des réseaux sociaux permet aussi une anonymisation des auteurs et autrices des discours postés en ligne¹⁰ et pose la délicate question de l'identification du critère de rattachement territorial entre l'infraction, son auteur et sa victime¹¹. Ceci remet en question l'adéquation de la réponse juridique classique – pénale et civile - face à cette situation et nécessite d'investiguer d'autres réponses régulatrices¹².

Dans ce contexte, le présent rapport de recherche vise à répondre à la question de recherche suivante : **quels sont les modèles régulateurs pour lutter contre les discours de haine en ligne ?** À ce titre, signalons que la notion de modèle vise un idéal type, lequel se caractérise par une logique sous-jacente qui doit être mise au jour. Le terme de régulation s'entend au sens juridique, et concerne les textes législatifs adoptés par des autorités publiques (droit dur). La mise au jour de la philosophie réglementaire (modèle) permettra alors de catégoriser les

⁶ Botero Arcila, B. & Griffin, R. (2023). *Social media platforms and challenges for democracy, rule of law and fundamental rights*. Etude pour le parlement européen. Disponible sous : [https://www.europarl.europa.eu/RegData/etudes/STUD/2023/743400/IPOL_STU\(2023\)743400_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2023/743400/IPOL_STU(2023)743400_EN.pdf)

⁷ Thiel, M., Maslanik, J. (2010). *Transnational actors*. Oxford Research Encyclopedias: International studies. <https://doi.org/10.1093/acrefore/9780190846626.013.105>

⁸ Benkler, Y. (2006) *The Wealth of Networks*. Yale University Press, New Haven, p. 33.

⁹ Botero Arcila & Griffin, *idem* n3, p. 20.

¹⁰ Banks, J. (2010). Regulating hate speech online. *International Review of Law, Computers & Technology*, 24(3), 233-239. <https://doi.org/10.1080/13600869.2010.522323>

¹¹ Alkiviadou, N. (2019). Hate speech on social media networks: towards a regulatory framework? *Information & Communications Technology Law*, 28(1), 19-35. <https://doi.org/10.1080/13600834.2018.1494417>

¹² Stockmann, D. (2023). Tech companies and the public interest: the role of the state in governing social media platforms. *Information, Communication & Society*, 26(1), 1-15. <https://doi.org/10.1080/1369118X.2022.2032796>

réponses législatives. Elle sera opérée à l'aune d'une double analyse (synchronique et diachronique) dans une perspective de politique publique et de droit comparé. La perspective synchronique sera utilisée dans le cadre de l'étude des rapports des rapporteurs spéciaux des Nations unies afin de cerner le cadrage du problème public (perspective de politique publique). La perspective diachronique permettra de comparer les réponses législatives qui s'inscrivent dans le sillage de la mise en œuvre du droit international et plus particulièrement des règles et standards en matière de droits humain (perspective de droit comparé). Les résultats de la recherche doivent ainsi permettre de contribuer à la connaissance du droit positif dans une perspective comparée et internationale, mais aussi aux études de politiques publiques qui s'intéressent à la circulation des idées dans le cadre de la mise en œuvre du droit international des droits de l'homme. Enfin, il faut dire que cette étude repose sur le postulat qu'entre le droit international et le droit national et comparé se joue un jeu d'influences verticale (entre le niveau international et national) et horizontale (entre les législations nationales), lequel génère une circulation des idées permettant l'amélioration de la réponse législative dans une perspective d'analyse préliminaire des mesures et des effets.

Par ailleurs, il convient ici de préciser les limites de cette étude. En effet, la question de la réglementation des entreprises de réseaux sociaux étant relativement récente, cette étude ne peut pas avoir une dimension évaluative, mais elle s'inscrit dans une démarche de type descriptive et analytique. Elle n'a pas non plus la prétention d'être fondée sur une analyse exhaustive de tous les acteurs ni de tous les instruments déployés par les Etats pour lutter contre le phénomène du discours de haine en ligne (dans une perspective de politique publique). Comme nous l'avons indiqué, elle entend se focaliser sur la réponse législative (instrument de droit dur du point de vue des instruments de politiques publiques).

Dès lors, l'étude est structurée de la manière suivante. Nous commencerons par présenter les deux éléments qui sont au cœur de cette étude, à savoir la notion de discours de haine et l'environnement « en ligne » que constitue l'écosystème internet (chapitre 2). Ensuite, nous présenterons notre méthodologie ainsi que la collecte des données que nous avons mobilisées dans notre étude (chapitre 3). Puis, à l'aune d'une analyse chronologique des documents produits au sein des Nations Unies, nous montrerons comment la notion de discours de haine a été clarifiée et quelles solutions réglementaires ont été esquissées à l'attention des Etats (chapitre 4). Dans une logique comparative (diachronique), nous analyserons alors les différentes réponses législatives adoptées en Allemagne, France, Autriche, dans l'Union européenne et au Royaume-Uni dans la foulée des propositions onusiennes (chapitre 5). Dans notre conclusion, nous mettrons en lumière la diversité des réponses réglementaires et évoquerons la question des développements futurs à l'échelle des Etats (chapitre 6).

Le présent rapport a été rédigé dans le cadre de l'obtention d'un financement de recherche dans le domaine du discours haineux en ligne, octroyé par l'Office fédéral de la communication (OFCOM) pour l'année 2023.

2. État des lieux : Discours de haine et environnement en ligne

Au départ de cette étude, deux éléments doivent être clarifiés. Le premier concerne la notion de discours de haine, laquelle est antérieure au développement des technologies de l'information et des communications et a été généralement appréhendée sous l'angle des droits humains (2.1.). Le second, plus récent, concerne cet environnement spécifique que constitue le contexte numérique et vise à clarifier la chaîne des acteurs impliqués dans ce contexte ainsi que la manière dont ceux-ci ont été cadrés juridiquement (2.2).

2.1. Le discours de haine en droit international des droits humains

Au regard du droit international des droits de l'homme, aucun traité ou convention ne mentionne la notion de discours de haine¹³. La notion de haine, elle, est bien mentionnée tant dans des règles de droit international que dans la pratique des organes de surveillance. Cette section a donc pour objet de cerner les contours de cette notion qui est au centre de cette étude. Elle s'appuie sur une analyse des textes et recommandations des Nations unies (2.1.1.) ainsi que sur les documents produits dans le cadre du Conseil de l'Europe (2.2.2).

2.1.1. Dans les normes et standards onusiens

Sur le plan des règles de droit conventionnel, le terme 'haine' figure dans deux conventions onusiennes : la **Convention sur l'élimination de toutes les formes de discrimination raciale de 1965** et le **Pacte sur les droits civils et politiques de 1966**.

Dans la Convention sur l'élimination de toutes les formes de discriminations raciales, l'article 4 dispose que :

« Les Etats parties condamnent toute propagande et toutes organisations qui s'inspirent d'idées ou de théories fondées sur la supériorité d'une race ou d'un groupe de personnes d'une certaine couleur ou d'une certaine origine ethnique, ou qui prétendent justifier ou encourager toute forme de haine et de discrimination raciales; ils s'engagent à adopter immédiatement des mesures positives destinées à éliminer toute incitation à une telle discrimination, ou tous actes de discrimination, et, à cette fin, tenant dûment compte des principes formulés

¹³ La question du discours de haine renferme une double question. La première relève de l'interprétation. Il s'agit de déterminer quand un propos ou un acte relève de cette notion qu'est le discours de haine. Elle relève alors du pouvoir des autorités d'application et du travail de dogmatique juridique opéré par la doctrine juridique. Elle ne sera pas évoquée dans le cadre de cette étude. La seconde question, qui précède la première, implique de déterminer quelles sont les caractéristiques de cette notion de discours de haine. Cette question-ci est au cœur de cette section.

dans la Déclaration universelle des droits de l'homme et des droits expressément énoncés à l'article 5 de la présente Convention. Ils s'engagent notamment :

- a) À déclarer délits punissables par la loi toute diffusion d'idées fondées sur la supériorité ou la haine raciale, toute incitation à la discrimination raciale, ainsi que tous actes de violence, ou provocation à de tels actes, dirigés contre toute race ou tout groupe de personnes d'une autre couleur ou d'une autre origine ethnique, de même que toute assistance apportée à des activités racistes, y compris leur financement ;
- b) À déclarer illégales et à interdire les organisations ainsi que les activités de propagande organisée et tout autre type d'activité de propagande qui incitent à la discrimination raciale et qui l'encouragent et à déclarer délit punissable par la loi la participation à ces organisations ou à ces activités ;
- c) À ne pas permettre aux autorités publiques ni aux institutions publiques, nationales ou locales, d'inciter à la discrimination raciale ou de l'encourager. »¹⁴.

Les travaux du Comité contre la discrimination raciale (ci-après le Comité), en sa qualité d'organisme de surveillance du respect de la convention, ont apporté plusieurs clarifications concernant la manière de lutter contre les propos haineux.

Ainsi, le Comité a indiqué que l'article 4 n'est pas directement applicable. Il invite dès lors les Etats à mettre en œuvre dans leur système juridique une série de mesures. Concrètement, les Etats sont invités à ériger en délits punissables par la loi et à sanctionner efficacement « a) toute *diffusion d'idées* fondées sur la supériorité ou la haine raciale ou ethnique, par quelque moyen que ce soit ; b) *l'incitation à la haine* ou à la discrimination envers des membres d'un groupe racial ou ethnique en raison de la race, la couleur, l'ascendance ou l'origine nationale ou ethnique ; c) les menaces ou l'incitation à la violence contre des personnes ou des groupes pour les motifs énoncés ci-dessus ; d) l'expression d'insultes, de moqueries ou de calomnies à l'égard de personnes ou de groupes, ou la justification de la haine, du mépris ou de la discrimination pour les motifs énoncés à l'alinéa b, lorsque ces actes s'apparentent clairement à de l'incitation à la haine ou à la discrimination ; e) la participation à des organisations ou des activités qui encouragent la discrimination raciale ou y incitent »¹⁵. Cela étant, le comité indique que, dans la poursuite de ces infractions, les Etats doivent prendre en considération non seulement les éléments comme le contenu et la forme du discours, mais également le climat économique, social et politique, la position et le statut de l'orateur, la portée du discours –

¹⁴ Nations Unies, Convention internationale sur l'élimination de toutes les formes de discrimination raciale. Disponible sous : <https://www.ohchr.org/fr/instruments-mechanisms/instruments/international-convention-elimination-all-forms-racial>. Cette convention a notamment été signée et ratifiée par la Suisse, l'Allemagne, l'Autriche, la France, le Royaume-Uni et les Etats-Unis.

¹⁵ Nations Unies. (2013). Recommandation générale n°35 sur la Convention internationale sur l'élimination de toutes les formes de discrimination raciale. CERD/C/GC/35, pp. 4-5, n°13. Disponible sous : <https://undocs.org/Home/Mobile?FinalSymbol=CERD%2FC%2FGC%2F35&Language=E&DeviceType=Desktop&LangRequested=False>

notamment la nature de l'audience et les modes de transmissions – et les objectifs du discours¹⁶.

En outre, le Comité a souligné que la lutte contre les différents types de propos pouvait être problématique au regard de la liberté d'expression. En effet, l'article 5 énonce l'obligation pour les Etats parties d'interdire et d'éliminer la discrimination raciale et de garantir le droit de chacun à l'égalité devant la loi dans la jouissance des droits civils politiques, économiques, sociaux et culturels, parmi lesquels notamment la liberté d'opinion et d'expression. À ce titre, le Comité estime que « la liberté d'expression ne devrait pas viser la destruction des droits et libertés d'autrui, notamment celle du droit à l'égalité et à la non-discrimination », et que « le droit fondamental à la liberté d'expression ne doit pas aller à l'encontre des principes d'égalité et de non-discrimination, car l'exercice du droit à la liberté d'expression comporte des responsabilités particulières, parmi lesquelles l'obligation de ne pas diffuser d'idées fondées sur la supériorité ou la haine raciale »¹⁷.

Enfin, le Comité a indiqué que les dispositions conventionnelles ne doivent pas être interprétées isolément. Ainsi, l'article 4 doit être interprété avec l'article 7, lequel requiert que les Etats s'engagent à prendre des mesures immédiates et efficaces, notamment dans les domaines de l'enseignement, de l'éducation, de la culture et de l'information, pour lutter contre les préjugés conduisant à la discrimination raciale et « promouvoir les principes universels des droits de l'homme »¹⁸. Les Etats parties doivent donc aller au-delà de la réponse dans le droit pénal. À cet égard, le Comité recommande l'adoption de stratégies pour réaliser les engagements prévus dans l'article 7¹⁹. Autrement dit, le comité appelle les Etats à développer une politique publique de lutte contre les propos haineux et d'incitation à la haine raciale.

Dans le **Pacte international relatif aux droits civils et politiques**, l'article 20 prévoit que :

« L'appel à la haine nationale, raciale ou religieuse qui constitue une incitation à la discrimination, à l'hostilité ou à la violence est interdit par la loi »²⁰.

Dans son observation générale, adoptée en 1983, l'organe de surveillance du pacte, le Comité des droits civils et politiques, a également mis en évidence la tension entre interdiction d'appel

¹⁶ *Ibid.*, p.p. 5-6, n° 15.

¹⁷ Recommandation générale n°35, *idem* n°15, p.8, n°26.

¹⁸ *Ibid.*, pp. 8-9, n°31.

¹⁹ Il faut noter ici que le Comité s'adresse aussi aux acteurs privés, notamment les fournisseurs d'accès internet qu'il encourage à s'autoréglementer et à se conformer à des codes d'éthique, comme cela avait été souligné dans la déclaration et le programme d'action de Durban. *Ibid.*, p. 10, n°42 ; Nations Unies. (2002). Déclaration et Programme d'action de Durban, pp. 108-109, par. 147. Disponible sous : https://www.ohchr.org/sites/default/files/Documents/Publications/DurbanDecProgAction_fr.pdf

²⁰ Nations Unies, Pacte international relatif aux droits civils et politiques. Disponible sous : <https://www.ohchr.org/fr/instruments-mechanisms/instruments/international-covenant-civil-and-political-rights>

à la haine et liberté d'expression. À son tour, il a estimé que l'article 20 n'est pas incompatible avec l'article 19 qui garantit la liberté d'expression, car l'exercice de ce droit impose des devoirs spéciaux et des responsabilités spéciales (art. 19, al. 3)²¹. Le Comité des droits civils et politiques a alors précisé que « l'interdiction prévue au paragraphe 1 s'étend à toutes les formes de propagande menaçant d'entraîner ou entraînant un acte d'agression ou une rupture de la paix, en violation de la Charte des Nations Unies, tandis que le paragraphe 2 vise tout appel à la haine nationale, raciale ou religieuse qui constitue une incitation à la discrimination, à l'hostilité ou à la violence, que cette propagande ou cet appel ait des objectifs d'ordre intérieur ou extérieur par rapport à l'Etat intéressé »²². Sur le plan des actions, le Comité des droits civils et politiques a rappelé que les Etats ont l'obligation de « prendre les mesures pour s'acquitter des obligations énoncées à l'article, et qu'ils devraient eux-mêmes s'abstenir de toute propagande ou de tout appel de ce genre »²³. Enfin, concernant l'article 19, al. 3, qui garantit la liberté d'expression, le Comité des droits civils et politiques a indiqué que des restrictions à la liberté d'expression sont justifiées pour deux motifs : le premier a une dimension individuelle - le respect des droits ou de la réputation d'autrui - et le second a une dimension collective - la sauvegarde de la sécurité nationale, de l'ordre public, de la santé ou de la moralité publiques²⁴. Autrement dit, la liberté d'expression n'est pas une liberté absolue et des restrictions sont envisageables.

Par ailleurs, d'autres organes des Nations Unies ont également pris position sur la question de propos incitant à la haine, à travers le prisme de la lutte contre l'intolérance et de la discrimination. Ainsi, le Conseil des droits de l'homme des Nations unies a adopté la **résolution 16/18 pour la « Lutte contre l'intolérance, les stéréotypes négatifs, la stigmatisation, la discrimination, l'incitation à la violence et la violence visant certaines personnes en raison de leur religion ou de leur conviction »**²⁵. Le Haut-Commissariat des Nations unies aux droits de l'homme a également adopté en 2012 une stratégie et un plan d'action²⁶. Le **plan d'action de Rabat** porte précisément sur

²¹ Comité des droits de l'homme. (1983). Observation générale n°11, article 20 du Pacte international relatif aux droits civils et politiques, par. 2. Disponible sous : https://tbinternet.ohchr.org/_layouts/15/treatybodyexternal/TBSearch.aspx?Lang=fr&TreatyID=8&DocTypeID=11 ; Comité des droits de l'homme. (2011). Observation générale n°34, article 19 du Pacte international relatif aux droits civils et politiques, CCPR/C/GC/34. Disponible sous : <https://documents-dds-ny.un.org/doc/UNDOC/GEN/G11/453/32/PDF/G1145332.pdf?OpenElement>

²² Observation générale n°11, *idem* n21, par. 2.

²³ *Ibid.*

²⁴ Observation générale n°34, *idem* n21, par. 21.

²⁵ Nations Unies, Conseil des droits de l'homme. (2011). Résolution A/HRC/RES/16/18. Disponible sous : <https://undocs.org/Home/Mobile?FinalSymbol=A%2FHRC%2FRES%2F16%2F18&Language=E&DeviceType=Desktop&LangRequested=False>

²⁶ Nations-Unies, Haut-Commissariat aux droits de l'homme. (2011). Conclusions et recommandations issues des quatre ateliers d'experts organisés par le HCDH en 2011 et adoptés par les experts à Rabat, Maroc, le 5 octobre 2012. Disponible sous : https://www.ohchr.org/sites/default/files/Rabat_draft_outcome_FR.pdf

l'interdiction de tout appel à la haine nationale, raciale ou religieuse²⁷. Parmi les mesures recommandées, et dans le prolongement du Comité contre les discriminations raciales, il encourage tous les Etats à adapter leur législation nationale afin d'y consacrer l'interdiction de l'incitation à la haine. Il définit aussi les propos qu'il entend combattre, à savoir *“tout type de communication, orale ou écrite, ou de comportement, constituant une atteinte ou utilisant un langage péjoratif ou discriminatoire à l'égard d'une personne ou d'un groupe en raison de leur identité, en d'autres termes, de l'appartenance religieuse, de l'origine ethnique, de la nationalité, de la race, de la couleur de peau, de l'ascendance, du genre ou d'autres facteurs constitutifs de l'identité”*²⁸. Ce type de mesure doit alors s'inscrire dans le respect du régime de restriction à la liberté d'expression.

2.2.2. Dans les normes et standards du Conseil de l'Europe

L'autre acteur incontournable en matière de droits humains est le Conseil de l'Europe. À la différence des Nations Unies, le Conseil de l'Europe ne dispose pas de convention internationale consacrant l'interdiction de formes d'expression de propos haineux²⁹. Toutefois, à travers sa **jurisprudence sur la liberté d'expression**, la Cour européenne des droits de l'homme a été saisie à plusieurs reprises d'affaires évoquant des propos haineux. Ces différentes affaires lui ont permis d'opérer une distinction entre « crimes de haine » et « discours de haine ». Le premier a été décrit comme *“un acte criminel motivé par des préjugés à l'égard d'une personne ou d'un groupe de personnes, tandis que le discours de haine concerne diverses formes d'expression dirigées contre une personne ou un groupe de personnes en raison des caractéristiques personnelles ou du statut de cette personne ou de ce groupe de personnes. Lorsque le discours de haine prend la forme d'un comportement qui constitue en soi une infraction pénale - comme un comportement abusif harcelant ou insultant – il peut également être qualifié de crime de haine”*³⁰. Autrement dit, dans le discours de haine, il manque un acte physique qui est de nature à porter atteinte à l'intégrité physique ou psychique d'une ou plusieurs autres personnes.

En outre, dans la **résolution R(97)20**, le Comité des ministres du Conseil de l'Europe a défini, en 1997, le discours de haine comme *«toutes formes d'expression qui propagent, incitent, prouvent ou justifient la haine raciale, la xénophobie, l'antisémitisme ou d'autres formes de haine fondées sur l'intolérance, y compris l'intolérance qui s'exprime sous la forme de nationalisme agressif et d'ethnocentrisme, de discrimination et d'hostilité à l'encontre des minorités, des immigrants et des personnes issues de l'immigration»*³¹. Autrement dit, les discours de haine sont bien

²⁷ Nations Unies, Haut-Commissariat aux droits de l'homme. (2012). Plan d'action de Rabat. Disponible sous [La liberté d'expression contre l'incitation à la haine : le HCDH et le Plan d'action de Rabat | OHCHR](#).

²⁸ *Ibid.*

²⁹ Voir notamment le 'Guide sur l'article 10 de la convention européenne des droits de l'homme' qui explique comment la liberté d'expression peut entrer en conflit avec d'autres dispositions de la Convention. Disponible sous : https://www.echr.coe.int/documents/d/echr/Guide_Art_10_FRA

³⁰ Service de l'exécution des arrêts de la Cour Européenne des droits de l'homme. (2022). Crimes de haine et discours de haine. Disponible sous : <https://rm.coe.int/thematic-factsheet-hate-crime-fra-docx/1680a96866>

³¹ Comité des ministres du Conseil de l'Europe. (1997). Recommandation n°R(97)20 sur le « discours de haine ». Disponible sous :

problématiques parce qu'ils valident publiquement des idées fondées sur la haine raciale, la xénophobie, l'antisémitisme et d'autres formes de haine.

En 1993, le Conseil de l'Europe s'est doté d'une Commission européenne contre le racisme (ci-après ECRI) pour lutter contre le racisme, l'antisémitisme, la xénophobie et toutes autres formes de discriminations³². Dans sa **recommandation n° 15 du 8 décembre 2015**, l'ECRI a opté pour une définition beaucoup plus large que celle retenue par les instances onusiennes. Il a ainsi défini le discours de haine comme *“le fait de prôner, de promouvoir ou d'encourager sous quelque forme que ce soit, le dénigrement, la haine ou la diffamation d'une personne ou d'un groupe de personnes ainsi que le harcèlement, l'injure, les stéréotypes négatifs, la stigmatisation ou la menace envers une personne ou un groupe de personnes et la justification de tous les types précédents d'expression au motif de la “race”, de la couleur, de l'origine familiale, nationale ou ethnique, de l'âge, du handicap, de la langue, de la religion ou des convictions, du sexe, du genre, de l'identité de genre, de l'orientation sexuelle, d'autres caractéristiques personnelles ou de statut”*³³. Dans le préambule de la recommandation, l'ECRI précise que ce qui est visé est ainsi un discours qui « tend à refléter ou à promouvoir l'hypothèse injustifiée que l'auteur est de quelque manière que ce soit supérieur à la personne ou au groupe de personnes visées »³⁴ et que toutes les formes de médias, en ligne ou hors ligne, jouent un rôle particulier aussi bien pour diffuser le discours de haine que pour le combattre³⁵. Dans l'exposé des motifs, la recommandation n° 15 indique encore que « la conception du discours de haine diffère de celle de nombreux autres documents en ce qu'elle s'applique non seulement à l'appel ou à l'incitation, sous toute forme que ce soit, au dénigrement, à la haine ou à la diffamation, et l'encouragement de telles manifestations ainsi qu'au harcèlement, aux injures, aux stéréotypes négatifs, à la stigmatisation ou aux menaces ; aux usages qui ne visent pas seulement à inciter à la commission d'actes de violence, d'intimidation, d'hostilité ou de discrimination mais dont on peut raisonnablement attendre qu'ils aient cet effet et aux motifs autres que la race, la couleur de peau, la langue, la religion ou les convictions, la nationalité ou l'origine nationale ou ethnique ainsi que l'ascendance »³⁶. Quant à la forme d'expression, elle s'entend comme englobant les discours et les publications de toute forme, notamment par le biais des médias électroniques, ainsi que leur diffusion et leur conservation³⁷.

<https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=090000168050116d>

³² L'institution est chargée de combattre le racisme, la discrimination raciale, la xénophobie et l'antisémitisme à la lumière de la convention européenne des droits de l'homme. L'ECRI exécute une mission d'accompagnement des Etats parties au Conseil de l'Europe en analysant leur système juridique et son effectivité, en incitant les différentes autorités étatiques à agir et en formulant des recommandations

³³ Commission européenne contre le racisme et l'intolérance. (2015). Recommandation de politique générale n° 15. Disponible sous : <https://rm.coe.int/recommandation-de-politique-generale-n-15-de-l-ecri-sur-la-lutte-contr/16808b5b03>

³⁴ *Ibid.*, p. 3.

³⁵ *Ibid.*, p. 4.

³⁶ *Ibid.*, p. 17.

³⁷ *Ibid.*, p. 17.

2.2.3. Caractéristiques communes pour appréhender la notion de discours de haine

Les différents documents institutionnels évoqués montrent donc que les institutions du Conseil de l'Europe et les Nations unies ne disposent pas d'une définition commune du discours de haine. Toutefois, il est possible d'identifier les points de convergence et de divergence entre les deux organisations internationales. Du côté des Nations unies, le discours de haine présente les éléments suivants :

- Sur le plan de la forme : il s'agit de « tout type de communication, orale ou écrite, ou de comportement » ou encore « la participation à des organisations ou des activités encourageant la discrimination ou qui y incitent ».
- Sur le plan du contenu du propos : il s'agit de propos qui véhiculent la haine raciale ou ethnique, nationale ou religieuse.
- Sur le plan de la finalité : il y a une distinction entre :
 - les propos **menaçant d'entraîner ou entraînant un acte d'agression** ou une rupture de la paix, en violation de la Charte des Nations Unies et ;
 - Ceux qui sont une **incitation** à la discrimination, à l'hostilité ou à la violence.
 - Cette incitation à la discrimination peut prendre la forme d'insultes, moqueries, calomnies, justification de la haine, du mépris ou de l'incitation à la haine ou à la discrimination.

Les institutions du Conseil de l'Europe visent quant à elles :

- Sur le plan de la forme : ce qui relève des expressions écrites, verbales, visuelles ou encore gestuelles.
- Sur le plan du contenu : ces expressions sont fondées sur des idées hostiles envers certains groupes de personnes pour des motifs physiques ou qui appartiennent à des groupes sociaux et culturels spécifiques.
- Sur le plan de la finalité : les discours de haine peuvent poursuivre différentes finalités entre propager, inciter, promouvoir, encourager ou encore justifier.

Ainsi, on peut retenir que les Nations unies incluent la participation à des organisations qui contribuent à la violence ou à l'incitation à la violence. Sur le plan de discrimination, les Nations unies disposent en outre d'une conception plus restrictive puisqu'ils visent uniquement les propos qui véhiculent des messages de haine sur la base de la « race », de la religion et de la nationalité alors que le Conseil de l'Europe dispose d'un éventail beaucoup plus large de contenus qui relèvent du discours de haine puisque cette conception recouvre toutes idées hostiles à l'égard de groupes de personnes pour des motifs physiques, sociaux ou culturels permettant aussi d'inclure les propos à l'encontre des femmes, des minorités de genre ou encore des personnes en situation de handicap. Enfin, c'est sur le dernier point que les institutions sont les plus proches. Les Nations Unies proposent une liste d'actions spécifiques (moquerie, calomnie, ...) alors que le Conseil de l'Europe vise plus globalement les types de registres permettant de véhiculer le discours. Ces nuances entre les deux organisations internationales montrent ainsi les variations qui peuvent entourer la qualification de discours de haine et conduire à des réponses réglementaires variées.

2.2. L'environnement numérique et l'approche économique de la réglementation des intermédiaires en ligne

Dans cette sous-section, il s'agit d'une part de décrire l'environnement numérique en mettant en lumière qui sont les entreprises en ligne pertinentes dans le cadre de la diffusion du discours de haine (2.2.1.). D'autre part, considérant que plusieurs études juridiques ont montré que contrairement à certaines idées véhiculées dès les années 1990, le développement de l'écosystème numérique ne s'est pas fait en dehors du droit³⁸, il s'agira de montrer quel cadrage juridique a été privilégié avec la survenance des entreprises de réseaux sociaux (2.2.2).

2.1.1. Les acteurs de l'écosystème numérique

Les technologies de l'information et des communications ont permis la mise sur pied et l'accès à l'infrastructure internet. Dans ce cadre, les premiers acteurs d'importance sont ceux qui permettent l'accès à Internet, à savoir les fournisseurs de services de télécommunication et les fournisseurs d'accès à internet qui sont le plus souvent les mêmes³⁹. En outre, l'accès à internet n'est pas possible sans matériel. Ainsi, les sociétés de matériel informatique conçoivent et fabriquent les dispositifs informatiques qui connectent les particuliers à Internet : cela comprend non seulement les ordinateurs, mais aussi les objets connectés. Ces sociétés de matériel informatique sont également primordiales pour les fournisseurs de services de télécommunication et d'accès à internet qui achètent également du matériel et les autres composants de réseau qui constituent l'ossature physique des réseaux.

Ensuite, toute une série d'organisations permettent l'utilisation d'internet et garantissent son fonctionnement :

- L'ICANN, entité à but non lucratif, qui attribue les noms de domaine⁴⁰ ;
- Les organisations qui assurent des activités de stockage de données ou de messageries électroniques.
- Les organisations qui offrent des services de moteurs de recherche (Google, Ecosia, et), lesquelles permettent aux usagers de rechercher des informations et ceux qui les créent, les agrègent et les publient. À ce propos, les organisations qui exploitent les moteurs de recherche fonctionnent avec des algorithmes qui déterminent ce que les utilisateurs voient et dans quel ordre. Ces algorithmes peuvent restreindre ou hiérarchiser les contenus, influençant ainsi les informations auxquelles les utilisateurs

³⁸ Voir en particulier : Cohen, J. (2019). *Between Truth and Power – The Legal Constructions of Informational Capitalism*. Oxford University Press ; Biettri, E. (2023). A Genealogy of Digital Platforms Regulation. *Georgetown Law Technology Review*, 7(1), 1- 68. Disponible sous : [A Genealogy of Digital Platform Regulation by Elettra Biettri :: SSRN](https://www.ssrn.com/sol3/papers.cfm?abstract_id=4644444)

³⁹ Certaines sociétés sont parfois dans une situation hybride en ce sens qu'elles fournissent un accès à internet et d'autres services liés à internet (par exemple Google qui a également un service *Google Fiber*). Voir notamment : Colombier, N., M'chirgui, Z., Pénard, T. (2010). An Empirical Analysis of Internet Providers' Interconnection Strategies. *Revue d'économie industrielle*, 131(3), 25-50. <https://doi.org/10.4000/rei.4178>

⁴⁰ ICANN, <https://www.icann.org/fr>

ont accès. Le service de moteur de recherche offert par ces entreprises constitue donc un élément important dans la circulation de l'information⁴¹.

- Les organisations qui offrent une plateforme de médias sociaux et des réseaux professionnels. Elles permettent aux usagers d'accéder à de l'information et de communiquer entre eux.⁴²
- Enfin, les organisations qui collectent et revendent des données (*data broker*) à partir des services et activités produites sur Internet.⁴³

Dans cet écosystème, les entreprises fournisseuses de courrier électronique et celles de plateformes de médias sociaux et réseautages offrent non seulement une gamme de services qui connectent les utilisateurs, mais elles disposent aussi de fils d'actualité et de plateformes de partage d'informations et d'idées avec possibilités d'émettre des critiques, des commentaires et des discussions. Leur modèle d'affaires repose sur l'extraction des données personnelles de leurs utilisatrices et utilisateurs⁴⁴. Dès lors, plus ces derniers passent du temps sur les espaces proposés par ces entreprises, plus celles-ci peuvent extraire des données de ces utilisatrices et utilisateurs, et plus ces entreprises peuvent engranger de revenus soit directement (en leur proposant des publicités) soit indirectement (en revendant les données collectées).

Dans le cas des intermédiaires en ligne auxquels nous nous intéressons dans le cadre de cette étude – les entreprises de réseaux sociaux –, ce modèle d'affaire explique que ces entreprises mettent en place des outils technologiques qui favorisent l'amplification des contenus qui génèrent le plus de réactions⁴⁵. Dès lors, les différentes entreprises de réseaux sociaux jouent un rôle clé dans la circulation des idées, que celles-ci soient inoffensives ou haineuses.

2.2.2. L'approche économique de la réglementation des entreprises en ligne : le régime juridique de responsabilité limitée pour les propos diffusés

Les entreprises de fournisseur d'accès et réseaux sociaux sont principalement des entreprises établies aux Etats-Unis et qui ont été historiquement régies par le droit américain. Le développement de leurs activités à l'échelle planétaire a toutefois conduit à certains ajustements de la part d'autres régulateurs, et notamment de l'Union européenne. Enfin, soumises à la pression de l'opinion publique et de régulateurs, ces entreprises ont aussi développé des standards communs afin d'essayer de limiter leur impact sur les individus et la

⁴¹ Andersen, J. (2018). Archiving, ordering, and searching: search engines, algorithms, databases, and deep mediatization. *Media, Culture & Society*, 40(8), <https://doi.org/10.1177/0163443718754652>

⁴² Ortiz-Opsina, E. (2019). The rise of social media. Consulté le 12 décembre 2023. Disponible sous : <https://ourworldindata.org/rise-of-social-media>

⁴³ Crain, M. (2016). The limits of transparency: Data brokers and commodification. *New Media & Society*, 20(1). <https://doi.org/10.1177/1461444816657096>

⁴⁴ Zuboff, S. (2020). *L'âge du capitalisme de surveillance*. Zulma Essays.

⁴⁵ En l'occurrence, des chercheurs ont montré que les contenus contenant des discours haineux faisaient partie des contenus qui retenaient le plus l'attention des utilisateurs, qu'ils soient choqués ou non par les idées véhiculées. Voy. Maarouf, A., Pröllochs, N., Feurriegel, S. (2022). The Virality of Hate Speech on Social Media. <https://doi.org/10.48550/arXiv.2210.13770>

société. Elles ont ainsi mis en place la procédure de modération de contenu, pour laquelle elles ont aussi été encouragées – en tout cas à l'échelle de l'Union européenne – à s'auto-réguler en reprenant des principes formulés par la société civile ou en développant leurs propres conduites.

En effet, dans le développement d'Internet et de son écosystème au cours des années 1990, le législateur américain a notamment traité la question du régime de responsabilité applicable à ces acteurs pour les contenus dont ils permettent la diffusion. Il a ainsi modifié en 1996 la loi fédérale sur les communications aux Etats-Unis de 1934⁴⁶. Il y a inclus la section 230, en tant que partie du *Communication Decency Act (CDA)*⁴⁷.

Le CDA consacre un principe d'immunité en matière de responsabilité civile et pénale pour les intermédiaires de l'Internet. On parle de principe d'immunité, car il déroge au régime classique de la responsabilité éditoriale, lequel prévoit que les acteurs dans le domaine de la presse écrite, de la radio et de la télévision suivent des principes de déontologie journalistique et exercent un contrôle sur les contenus qu'ils publient afin de s'assurer de ne pas publier des contenus qui seraient illégaux et engageraient la responsabilité de ces organismes⁴⁸. À l'inverse, le CDA prévoit que les « sites web et autres services en ligne ne sont pas responsables du contenu des tiers »⁴⁹. La justification de cette dérogation au régime de responsabilité éditoriale a été motivée par des considérations économiques. Le Congrès américain a en effet considéré que le régime de la responsabilité éditoriale était trop restrictif et qu'il risquait de compromettre la croissance des nouveaux acteurs de l'internet⁵⁰. Ce régime alternatif au profit des « intermédiaires en ligne » implique donc que leur responsabilité tant civile que pénale ne peut pas être engagée, même dans les cas où ceux-ci ont connaissance de la présence de contenus illégaux et se refusent à les retirer⁵¹.

Dans la foulée, s'inspirant de la conception américaine, l'Union européenne a repris l'idée américaine d'immunité avec l'adoption, le 8 juin 2000, de la directive sur le commerce électronique (directive e-commerce)⁵². Toutefois, le législateur européen s'est légèrement distancié du régime américain en optant pas pour une irresponsabilité absolue. En effet, la directive e-commerce prévoit que les entreprises peuvent être tenues pour responsables des contenus de tiers si elles ne retirent pas les contenus illicites lorsqu'elles sont informées de

⁴⁶ Communications Act of 1934. Consulté le 15 novembre 2022. Disponible sous : <https://transition.fcc.gov/Reports/1934new.pdf>

⁴⁷ US Congress., Congressional Research Service. (2024). Section 230: An Overview. Consulté le 15 janvier 2024. Disponible sous : <https://crsreports.congress.gov/product/pdf/R/R46751>

⁴⁸ Ibid.

⁴⁹ Goldman, E. (2020). An Overview of the United States' Section 230 Internet Immunity. In Frosio, G. (ed.) *The Oxford Handbook of Online Intermediary Liability*. Oxford University Press.

⁵⁰ Cohen, *idem* n38.

⁵¹ Kuczerawy, A. (2015). Intermediary liability & freedom of expression: Recent developments in the EU notice & Action initiative. *Computer Law & Security Review*, 31(1), 46-56. <https://doi.org/10.1016/j.clsr.2014.11.004>

⁵² Yannopoulos, G. (2017). *The Immunity of Internet Intermediaries Reconsidered?*. In Taddeo, M., Floridi, L. (eds), *The Responsibilities of Online Service Providers*. Springer, 43-59.

leur existence (principe du « *notice and action* »⁵³). Autrement dit, c'est uniquement dans le cas où elles n'ont pas connaissance de l'existence de ces contenus sur leurs services que leur responsabilité ne peut pas être engagée⁵⁴.

Cela étant, depuis la fin du 20^{ème} siècle et la mise en place du régime de responsabilité décrit ci-dessus, les entreprises en ligne ont connu une croissance fulgurante. Le développement de leurs activités s'est accompagné de défis croissants, notamment en matière de prolifération de contenus illégaux ou pouvant être dommageables pour les utilisateurs et utilisatrices⁵⁵. Pour faire face à ces défis, les entreprises se sont volontairement dotées d'un mécanisme d'auto-régulation : la modération des contenus⁵⁶. Ce mécanisme organisationnel désigne une pratique qui consiste à opérer un tri des contenus postés sur leurs plateformes. Cette pratique est formalisée dans une série de textes et mise en œuvre à l'aide de différents outils. Parmi les textes figurent notamment les conditions générales, qui détaillent les conditions selon lesquelles les entreprises décident de retirer ou de restreindre l'accessibilité de certains contenus⁵⁷. La modération des contenus contribue au confort des utilisateurs et utilisatrices des plateformes : s'ils sont confrontés à des contenus désagréables ou qui les choquent, il est probable qu'ils passent moins de temps sur les réseaux sociaux, générant moins de revenus pour les entreprises⁵⁸. Du point de vue des outils, les entreprises en ligne ont progressivement eu de plus en plus recours à des processus algorithmiques, au point où la modération des contenus se fait désormais principalement de façon automatique, par exemple sur la base de détection basée sur le « *hash matching* » ou sur des processus prédictifs de classification⁵⁹. Ainsi, dès le début de leurs activités, les entreprises de réseaux sociaux ont commencé à trier les contenus postés par les utilisatrices et utilisateurs, tant à des fins commerciales qu'à des fins d'auto-régulation.

La modération des contenus ne se déroule toutefois pas en vase clos. Parallèlement à son développement et sa mise en place par les entreprises de réseaux sociaux, des initiatives ont émergé afin de développer des standards non contraignants sur comment la modération des contenus pourrait s'organiser. Ces initiatives visent à encourager les entreprises de réseaux sociaux à rendre plus transparente la façon dont ils modèrent les contenus et de mieux protéger les utilisateurs. Parmi ces différentes initiatives, certaines se veulent globales, comme

⁵³ Streel, A., Husovec, M. (2020). *The e-commerce Directive as the cornerstone of the Internal Market – assessment and options for reforms*. Disponible sous : https://www.europarl.europa.eu/RegData/etudes/STUD/2020/648797/IPOL_STU%282020%29648797_EN.pdf

⁵⁴ Kuczerawy, *idem* n50.

⁵⁵ Maarouf, A., Pröllochs, N., Feurriegel, S. *Idem* n45.

⁵⁶ Gillespie, T. (2018). *Custodians of the Internet*. New Haven/London: Yale University Press.

⁵⁷ Klonick, K. (2018). The New Governors. *Harvard Law Review*, 131(6), 1598-1670.

⁵⁸ Gillespie, *idem* n56.

⁵⁹ Gorwa, R., Binns, R., Katzenbach, C. (2020). Algorithmic content moderation: technical and political challenges in the automation of platform governance. *Big Data & Society*, 7(1). <https://doi.org/10.1177/2053951719897945>. Le « *hash matching* » désigne une pratique qui consiste à attribuer un « label » à un certain contenu et à comparer ce « label » avec une base de données de contenus illégaux. Les contenus postés qui contiennent le même label que celui présent dans la base de données sont retirés.

les Principes de Manille sur la responsabilité des intermédiaires⁶⁰ ou les Principes de Santa Clara sur la transparence et la responsabilité dans la modération des contenus⁶¹. À côté de ces initiatives globales, l'Union Européenne a développé une réponse plus régionale en développant plusieurs codes de conduite pour les entreprises de réseaux sociaux. Ainsi, en matière de discours de haine particulièrement, l'UE a adopté en 2016 le Code de conduite contre les discours haineux en ligne⁶². L'UE a également mis sur pied un Code de bonnes pratiques en matière de désinformation en 2022⁶³.

3. Approches, données et méthode d'analyse

Afin de mettre en lumière les modèles réglementaires en matière de lutte contre les discours de haine, cette étude s'inscrit dans une double approche et procède à une analyse qualitative de différents documents.

Ainsi, la *première approche* est à la croisée entre droit international des droits de l'homme et les études sur la régulation (« *regulation studies* »). Elle est mobilisée dans le cadre du chapitre 4 dédié à l'étude des documents produits par des instances spécialisées dans la gouvernance des droits humains, à savoir le rapporteur spécial en matière de liberté de pensée et d'expression et la Commission européenne pour la lutte contre le racisme du Conseil de l'Europe. Comme on l'a vu dans la section précédente, la diffusion de messages de haine a historiquement été la préoccupation des juristes face aux expériences traumatiques de la seconde guerre mondiale. Le développement de l'environnement numérique a offert un nouveau medium de diffusion des discours de haine, lequel a également vocation à être pris en compte par les différents comités et mandataires internationaux en matière de droits humains. Nous nous focalisons sur ces deux instances, car nous estimons qu'en matière de circulation des idées, elles assurent un rôle d'*incubateur* du fait de leurs fonctions de réceptacle des observations et analyses réalisées au sein de la société civile et d'autres intermédiaires étatiques, un rôle de synthèse et enfin un rôle de diffusion d'un certain type de réponse à large échelle et spécialement à l'attention des intermédiaires nationaux⁶⁴. Ainsi, ces instances ont la

⁶⁰ Manila Principles on Intermediary Liability. Consulté le 14 octobre 2023. Disponible sous : <https://manilaprinciples.org/principles.html>

⁶¹ The Santa Clara Principles on Transparency and Accountability in Content Moderation. Consulté le 14 octobre 2023. Disponible sous : <https://santaclaraprinciples.org/>

⁶² Commission Européenne. (2016). The EU Code of conduct on countering illegal hate speech online. Consulté le 17 octobre 2023. Disponible sous : https://commission.europa.eu/strategy-and-policy/policies/justice-and-fundamental-rights/combating-discrimination/racism-and-xenophobia/eu-code-conduct-countering-illegal-hate-speech-online_en?prefLang=fr

⁶³ Commission Européenne. (2022). The EU Code of Practice on Disinformation. Consulté le 17 octobre 2023. Disponible sous : <https://digital-strategy.ec.europa.eu/fr/policies/code-practice-disinformation>

⁶⁴ Subeid, S., Wheatley, S., Mukherjee, A., Ntute Ngane, S. (2011). The Role of the Special Rapporteurs of the United Nations Human Rights Council in the Development and Promotion of International Human Rights Norms. *The International Journal of Human Rights*, 15(2), 155-161. <https://doi.org/10.1080/13642987.2011.537463>

capacité de contribuer à la mise à l'agenda des gouvernements nationaux de certaines problématiques⁶⁵.

Les études sur la régulation ont également montré que la mise en œuvre du droit international des droits de l'homme constitue un phénomène plus complexe qui dépasse la simple ratification ou transposition *top-down* d'une norme internationale en droit interne telle qu'elle est classiquement décrite en droit international public. Elle est le fruit d'un enchevêtrement d'influences impliquant une diversité d'acteurs qui sont qualifiés d'intermédiaires de la régulation⁶⁶. Dans cette perspective, la transposition du droit international n'est pas seulement question d'un rapport entre l'organisation internationale (le régulateur) et les Etats (la cible de la régulation), mais implique l'intervention d'une diversité d'acteurs (les intermédiaires). Toutefois, à la différence de recherches en sciences politiques, nous ne chercherons pas à reconstruire la chaîne d'acteurs et encore moins à établir un lien de causalité entre le régulateur, les différents intermédiaires et la cible de la régulation. Portant davantage notre attention sur la norme, nous nous intéressons au contenu véhiculé par ces intermédiaires de la régulation en aval de l'adoption de la réglementation et examinerons si ce contenu figure dans les réglementations nationales et de l'Union européenne⁶⁷.

La *seconde approche* est alors celle du droit comparé. Elle est appliquée dans le cadre du chapitre 5 dédié à l'analyse des différentes législations qui ont été adoptées pour réglementer les entreprises de réseaux sociaux. L'approche comparative peut viser plusieurs finalités dans le domaine des études juridiques : il peut permettre d'améliorer un système juridique à partir des enseignements tirés de bonnes et/ou mauvaises expériences, ou peut viser une dimension théorique en contribuant à la connaissance des systèmes juridiques et à la conceptualisation de notions juridiques. Pour cela, Il s'agit d'opérer un travail de comparaison des différences et des similitudes. Le travail de comparaison doit toutefois pouvoir s'appuyer sur un *tertium comparationis*, qui sert de point de comparaison. Ce *tertium comparationis* peut résulter d'une construction théorique ou d'une analyse empirique d'un cas spécifique, lequel est alors pris comme modèle. En l'occurrence, nous postulons que le travail des intermédiaires de la régulation contribue au cadrage réglementaire et, dans une finalité analytique, nous permet de dégager un *tertium comparationis*. Toutefois, considérant la dimension dynamique qu'impliquent des réponses régulatrices comparées qui se sont échelonnées dans le temps, nous postulons

⁶⁵ Majone, G. (2006). *Agenda Setting*. Moran, M., Rein, M., Goodin, R.E. (eds). *The Oxford Handbook of Public Policy*. Oxford University Press.

⁶⁶ Les études sur la régulation proposent le modèle du RIT. Ce modèle est fondé sur l'observation qu'une réglementation n'est pas formulée et mise en œuvre d'un rapport bilatéral *top-down* entre le régulateur (regulator (R)) et la cible de la réglementation (Target (T)), mais qu'elle implique généralement des intermédiaires (I). Abbott, K. W., Levi-Faur, D., & Snidal, D. (2017). *Introducing Regulatory Intermediaries*. *The ANNALS of the American Academy of Political and Social Science*, 670(1), 6-13. <https://doi.org/10.1177/0002716217695519>.

⁶⁷ Les intermédiaires jouent divers rôles tout au long du cycle politique. Ainsi, ils peuvent intervenir en amont du processus réglementaire, mais aussi en avant lors de la mise en œuvre de la réglementation. Ils contribuent alors à l'interprétation des normes réglementaires, ou encore dans l'évaluation dans la mise en œuvre. Voir Abbott, K. W., Levi-Faur, D., & Snidal, D. (2017). *Idem* n66.

que le cadrage par les intermédiaires de la régulation n'exclut pas une évolution de la réponse réglementaire.

Sur le plan de la collecte des documents, nous avons sélectionné les rapports produits par les rapporteurs spéciaux des Nations Unies. En l'occurrence, en 1993, la Commission des droits de l'homme des Nations Unies a institué un rapporteur spécial sur la promotion et la protection du droit à la liberté d'opinion et d'expression. Depuis 1993, les rapporteurs spéciaux des Nations unies sur la promotion et la protection du droit à la liberté d'opinion et d'expression ont publié chaque année un rapport annuel, lequel était parfois complété par d'autres rapports thématiques⁶⁸. Parmi tous ces rapports, nous avons retenu 15 rapports, au motif qu'ils mentionnaient au moins une fois la notion de discours de haine. Etant donné la richesse de ces documents et la complexité de l'analyse (synchronique et diachronique), nous n'avons pas intégré ici les travaux de la commission pour lutter contre le racisme et l'intolérance (ECRI) du Conseil de l'Europe, lesquels permettent de retrouver la même approche que celle développée dans le cadre des rapports des rapporteurs spéciaux onusiens.

Au niveau national, notre recherche dans la littérature sur la réglementation des réseaux sociaux et de veille de l'actualité législative européenne, nous ont conduit à sélectionner cinq législations adoptées par l'Allemagne, la France, l'Autriche, le Royaume-Uni et l'Union européenne. Pour chacune de ces législations, nous avons collecté les textes législatifs ainsi que les documents préparatoires publiés par les autorités politiques et administratives dans les cinq systèmes politico-juridiques. Ces documents ont été sélectionnés sur les sites officiels des différents journaux officiels ainsi que des sites internet des gouvernements et des parlements nationaux.

⁶⁸ La fonction de rapporteur spécial constitue un mandat d'une durée de trois ans, qui est attribué par l'organe des Nations Unies compétent en matière de droits humains à une personnalité du monde scientifique, laquelle est chargée de rédiger des rapports thématiques. Le premier mandat relatif à la liberté d'expression est attribué en 1993. À la suite du remplacement de la Commission sur les droits de l'homme par le Conseil des droits de l'homme en 2008, le mandat relatif à la liberté d'opinion et d'expression a été prolongé le 28 mars 2008 (résolution du Conseil des droits de l'homme 7/36, https://ap.ohchr.org/documents/F/HRC/resolutions/A_HRC_RES_7_36.pdf)

4. Conceptualisation du discours de haine et sa régulation à la lumière du droit international des droits humains

Cette section propose un premier niveau d'analyse à partir de l'étude des rapports du rapporteur spécial des Nations Unies en matière de liberté d'opinion et d'expression. L'analyse de ces rapports permet de mettre en exergue trois phases au cours desquels va être précisé les contours des discours en ligne problématique parmi lesquels figurent le discours de haine et la réponse régulatoire proposée aux Etats pour résoudre ce problème. Ces trois phases sont les années 1990 et 2000 marquées par le libre accès à Internet (4.1.) ; les années 2010 dominées par une cristallisation des tensions sociales, politiques, économiques et culturelles, lesquelles se traduisent par un renforcement de la problématique générale des discours de haine (4.2) et la troisième période, à partir des années 2010, qui est marquée par la reconnaissance de l'importance du potentiel rôle des entreprises en ligne dans la lutte contre le discours de haine (4.3.).

4.1. Les années 1990 et 2000 : l'Etat régulateur de la liberté d'expression et l'absence de définition du discours de haine

Le premier rapport qui aborde la question des nouvelles technologies de l'information et la liberté d'expression date de 1998. Le rapport fait état d'une vision positive des développements technologiques. Les nouvelles technologies y sont décrites comme fondamentalement démocratiques⁶⁹. Plusieurs autres rapports suivent et viennent souligner cet aspect, décrivant les nouvelles technologies de l'information comme un moyen permettant de réaliser la pluralité des voix et leur intégration dans le débat public⁷⁰. À ce titre, il est noté que « les nouvelles technologies (ont) ouvert de nouvelles possibilités d'expression, de diffusion des idées et de transfert de l'information (...) ». L'internet est décrit comme « un moyen de communication unique » car il est « mondial, décentralisé, interactif et surtout indépendant de toute infrastructure, abolissant de ce fait les frontières nationales »⁷¹. À cet égard, il doit devenir « l'un des instruments les plus efficaces pour combattre l'intolérance en ouvrant la voie à des messages de respect mutuel qui peuvent circuler librement à travers le monde et en encourageant la concertation pour lutter et mettre un terme à des phénomènes

⁶⁹ Rapport E/CN.4/1998/40 de Monsieur Abid Hussain, Rapporteur spécial sur la promotion et la protection du droit à la liberté d'opinion et d'expression, 28 janvier 1998, p. 14, §45. Disponible sous <https://documents-dds-ny.un.org/doc/UNDOC/GEN/G98/103/13/PDF/G9810313.pdf?OpenElement>

⁷⁰ Rapport E/CN.4/2000/63 de Monsieur Abid Hussain, Rapporteur spécial sur la promotion et la protection du droit à la liberté d'opinion et d'expression, 18 janvier 2000, p.19, §54. Disponible sous <https://documents-dds-ny.un.org/doc/UNDOC/GEN/G00/102/60/PDF/G0010260.pdf?OpenElement> ; Rapport E/CN.4/2001/64 de Monsieur Abid Hussain, Rapporteur spécial sur la promotion et la protection du droit à la liberté d'opinion et d'expression, 13 février 2001, p.20, §61. Disponible sous : <https://documents-dds-ny.un.org/doc/UNDOC/GEN/G01/111/24/PDF/G0111124.pdf?OpenElement>.

⁷¹ Rapport E/CN.4/2000/63, *idem* n70, p. 21, §55.

tels que les propos haineux et le racisme »⁷². Internet est donc décrit comme un instrument de promotion à l'échelle globale pour la diffusion d'une culture de paix.

Le rapport permet d'identifier plusieurs sujets de préoccupations. Le premier concerne l'action des États qui devraient s'abstenir de réglementer de façon excessive cet environnement : « les mesures prises par les États pour soumettre l'utilisation de ces techniques [de l'Internet et des autres technologies de l'information] à une réglementation excessive en faisant valoir que l'accès doit en être limité, réglementé ou refusé afin de préserver la texture morale et l'identité culturelle des sociétés, nient le pouvoir de raison et méconnaissent la capacité de résistance des individus et de la société »⁷³. Le rapport critique aussi les gouvernements dont certains « sont plus soucieux de contrôler et réglementer ces technologies que de développer les réseaux existants, de relever les capacités techniques pour atteindre des régions insuffisamment desservies ou pas desservies du tout, (...) »⁷⁴. Face à ce problème, le rapporteur plaide pour que ces technologies soient examinées sur la base des mêmes normes internationales que les autres moyens de communication et qu'il ne soit appliqué aucune mesure de nature à restreindre l'accès à la liberté d'expression et d'information. Dès lors, en cas de doute, il estime que la décision prise devrait être en faveur de la libre expression et de la libre circulation de l'information. À ce titre, l'idée de symétrie entre la vie hors ligne et en ligne est exprimée : l'expression « en ligne » devrait se confirmer aux normes internationales et bénéficier de la même protection que les autres formes d'expression⁷⁵.

La seconde préoccupation concerne l'utilisation des technologies à des fins problématiques sur le plan juridique et social : « il est également incontestable que l'utilisation de ces nouvelles sources d'information par certains groupes et certains individus a suscité de sérieuses préoccupations qui concernent plus spécialement, par exemple, les propos racistes et haineux, les incitations à la violence, la pornographie (notamment la pédophilie et le tourisme sexuel), la protection de la vie privée et de la réputation et les valeurs culturelles et sociétales »⁷⁶. La question de l'impact des technologies de l'information et des communications sur la liberté d'expression et sur la promotion de comportements et propos répréhensibles est partagé par la Commission des droits de l'homme, dans sa résolution 1998/42, qui invite le rapporteur « à étudier les avantages et les défis que présentent les nouvelles technologies des télécommunications, y compris l'Internet, pour l'exercice du droit à la liberté d'opinion et d'expression, qui comprend le droit de rechercher, de recevoir et de répandre des

⁷² Rapport E/CN.4/1999/64 de Monsieur Abid Hussain, Rapporteur spécial sur la promotion et la protection du droit à la liberté d'opinion et d'expression, 29 janvier 1999, p. I I, §32. Disponible sous <https://documents-dds-ny.un.org/doc/UNDOC/GEN/G99/107/67/PDF/G9910767.pdf?OpenElement>

⁷³ Rapport E/CN.4/1999/64, *idem* n72, p. I I, §30.

⁷⁴ Rapport E/CN.4/2000/63, *idem* n70, p.20, §54.

⁷⁵ Rapport E/CN.4/1998/40, *idem* n69, p.33, §116.

⁷⁶ Rapport E/CN.4/1998/40, *idem* n69, p. 10, §33.

informations » en tenant compte des travaux entrepris par le comité pour l'élimination de la discrimination raciale⁷⁷.

Le rapport de 1999 confirme cette perception ambivalente des technologies de l'information et des communications du point de vue des droits humains. Ainsi, il est dit que « l'Internet ne devrait pas être une « zone sans foi ni loi »⁷⁸ et qu'il donne lieu à toute une série de problèmes auxquels les gouvernements essaient de répondre : « la plupart des gouvernements (qui) ont essayé de résoudre les problèmes posés par la diffusion et la prolifération de la pornographie, et de documents réalisés par des groupes néonazis et/ou d'autres "groupes animés par la haine" »⁷⁹. Toutefois, ces problèmes ne doivent pas pour autant occulter l'intérêt pour les nouvelles technologies : « nous ne devrions pas nous préoccuper de façon excessive des risques associés aux nouvelles technologies grâce auxquelles les marginaux, les sans-grades, peuvent se faire entendre »⁸⁰.

Les années 2000 sont celles de la commercialisation des smartphones et des attentats du 11 septembre qui conduisent à une recrudescence de la surveillance étatique. Ainsi, en 2001, le rapport fait mention de la préoccupation du rapporteur face à l'afflux d'informations « faisant état d'une réglementation excessive du contenu de l'Internet et d'une surveillance en ligne ». Il souligne la persistance du problème de « la diffusion de matériel illicite sur Internet (pornographie à caractère pédophile, propagande néonazie et appels à la haine) (qui) constitue certes un sujet de préoccupation légitime mais toute restriction imposée à l'Internet devrait s'inscrire rigoureusement dans le cadre des restrictions du paragraphe 3 de l'article 19 du Pacte »⁸¹. Autrement dit, la question d'internet continue d'être envisagée à travers le prisme de la liberté d'expression. Dans ce contexte, le rapporteur appelle à la fixation de standards internationaux pour garantir la liberté d'expression en ligne. En effet, il est également observé que « des gouvernements, des organisations et des particuliers s'intéressaient de plus en plus à la question de savoir s'il était possible ou non de concevoir un système de réglementation applicable d'une manière générale à l'Internet. Une telle démarche s'expliquait, au moins en partie, par les craintes relatives à la confidentialité, à la diffusion de matériel pornographique, aux discours racistes et haineux et à l'utilisation des technologies pour promouvoir par exemple le tourisme à caractère sexuel impliquant des enfants (...) »⁸².

Le rapport de 2005 présente encore cette double dimension. D'un côté, « le principal défi pour l'avenir sera de garantir la liberté d'opinion et d'expression sur l'Internet et dans les

⁷⁷ Nations Unies, Commission des droits humains. (1998). Résolution 1998/42, Document E/CN.4/1998/42. Disponible sous : <http://hrlibrary.umn.edu/UN/1998/Res042.html>

⁷⁸ Rapport E/CN.4/1999/64, *idem* n72, p.13, §36

⁷⁹ Rapport E/CN.4/2000/63, *idem* n74, p.21, §56.

⁸⁰ Rapport E/CN.4/1999/64, *idem* n72, p.13, §36

⁸¹ Rapport E/CN.4/2002/75 de Monsieur Abid Hussain, Rapporteur spécial sur la promotion et la protection du droit à la liberté d'opinion et d'expression, 30 janvier 2002, p. 23, §93. Disponible sous : https://ap.ohchr.org/documents/alldocs.aspx?doc_id=2900

⁸² Rapport E/CN.4/2001/64, *idem* n72, p.19, §57.

autres moyens de communication »⁸³. De l'autre côté, le rapporteur indique qu'il est « particulièrement inquiet de l'usage croissant de formes d'expression à connotation discriminatoire à l'égard de groupes ethniques et d'autres groupes vulnérables. Les discours haineux et autres formes d'expressions semblables peuvent contribuer considérablement à la détérioration d'une stabilité précaire et mener un pays à l'affrontement armé »⁸⁴. Autrement dit, si internet doit être un lieu propice à la liberté d'expression, certaines activités promues ou propos tenus constituent aussi un problème social et politique.

Il faut attendre 2007, l'année de la commercialisation de l'iPhone, pour que la question de la gouvernance d'internet soit évoquée. Pour le rapporteur spécial, il faut mettre sur pied la création d'une organisation à vocation mondiale qui placerait Internet sous le giron des droits humains. Internet n'est plus seulement vu comme un espace politique, mais également comme un espace économique. Il estime alors qu'il faut « garantir un monde numérique où les initiatives commerciales coexistent avec les projets sociaux et culturels et évitent aux consommateurs de payer une taxe pour tout service disponible ». Le rapporteur spécial note toutefois que « malheureusement, les publicitaires se sont jusque-là montrés très agressifs sur le marché numérique et qu'ils mettent au point des technologies ciblées sur les cyberconsommateurs, et notamment les jeunes, lors de leurs explorations du cyberspace. En outre, la collecte intrusive de données et les pratiques de détection et pistage restreignent le droit à la vie privée et partant la liberté d'expression des utilisateurs réguliers de l'internet »⁸⁵. Le problème d'internet n'est donc plus seulement une question entre l'Etat et les utilisateurs ou entre les utilisateurs entre eux, il y a également une troisième catégorie d'acteurs : les entreprises qui font du commerce avec les données des utilisateurs.

Les rapports produits au cours de ces deux décennies montrent ainsi que la liberté d'expression constitue bien le cadrage des rapports des rapporteurs spéciaux, mais qu'est aussi évoqué accessoirement la dimension économique d'internet avec l'arrivée des entreprises poursuivant des buts commerciaux. Si internet est décrit comme un moyen de promotion de la liberté, il est également perçu comme une menace pour la protection des individus eu égard aux développements de discours de haine et aux activités problématiques que sont la promotion de la violence ou encore la pédopornographie. Dans sa résolution 7/36 de 2009, le Conseil des droits de l'homme charge alors le rapporteur spécial de faire rapport « sur les cas dans lesquels l'atteinte au droit à la liberté d'expression constitue un acte de discrimination raciale ou religieuse, en tenant compte du paragraphe 3 de l'article 19 et de l'article 20 du Pacte international relatif aux droits civils et politiques, ainsi que la recommandation générale n°15 du Comité pour l'élimination de la discrimination raciale, aux termes de laquelle l'interdiction de diffusion de toute idée fondée sur la supériorité ou la haine

⁸³ Rapport E/CN.4/2005/64 de Monsieur Ambeyi Ligabo, Rapporteur spécial sur la promotion et la protection du droit à la liberté d'opinion et d'expression, 15 février 2005, p.14, §57. Disponible sous https://ap.ohchr.org/documents/alldocs.aspx?doc_id=11160

⁸⁴ *Ibid.*, p. 14, §58.

⁸⁵ Rapport A/HRC/4/27 de Monsieur Ambeyi Ligabo, Rapporteur spécial sur la promotion et la protection du droit à la liberté d'opinion et d'expression, 2 janvier 2007, p.11, §41. Disponible sous : <https://documents-dds-ny.un.org/doc/UNDOC/GEN/G07/101/82/PDF/G0710182.pdf?OpenElement>.

raciale est incompatible avec le droit à la liberté d'opinion ou d'expression »⁸⁶. À cet égard, le comité des droits de l'homme établit un lien entre liberté d'expression, interdiction des appels à la haine et discrimination raciale et religieuse. De son côté, le rapporteur spécial indique que « le principal défi [de la tension entre liberté d'expression et actes de discrimination raciale ou religieuse] consiste à fixer des seuils à ne pas dépasser »⁸⁷, précisant ainsi déjà la direction à suivre.

4.2. Au cours de la première moitié des années 2010 : l'émergence des entreprises comme intermédiaires de la régulation et la conception restrictive du discours de haine à l'aune du droit pénal international

Dans la foulée de la résolution du Conseil des droits de l'homme⁸⁸, le rapporteur spécial consacre son rapport de 2011 à la liberté d'expression sur Internet⁸⁹. Il identifie quatre types d'expression exceptionnelles que les États sont tenus d'interdire. Il s'agit de (I) la pornographie mettant en scène des enfants⁹⁰, de (II) l'incitation directe et publique à commettre le génocide⁹¹, de (III) l'apologie de la haine nationale, raciale ou religieuse qui constitue une incitation à la discrimination, à l'hostilité ou à la violence⁹² et de (IV) l'incitation au terrorisme depuis la résolution 1624 de 2005 du Conseil de sécurité des Nations unies⁹³. Ce faisant, le

⁸⁶ Conseil des droits de l'homme, Résolution 7/36. *Idem* n°68 ; Conseil des droits de l'homme, Assemblée générale, Onzième session, document A/HRC/11/4, 30 avril 2009. Pp. 4-5, §4. Disponible sous : <https://documents-dds-ny.un.org/doc/UNDOC/GEN/G09/130/33/PDF/G0913033.pdf?OpenElement>.

⁸⁷ Conseil des droits de l'homme, A/HRC/11/4, *idem* n°86, p. 12, §39. Il faut ici signaler qu'il est mentionné qu'« une interprétation plus large de ces limites, comme celle qui a été récemment proposée dans des enceintes internationales, ne serait pas conforme aux instruments internationaux en vigueur et compromettrait en définitive la pleine jouissance des droits de l'homme ».

⁸⁸ *Ibid.*

⁸⁹ Rapport A/66/290 de Monsieur Franck La Rue, Rapporteur spécial sur la promotion et la protection du droit à la liberté d'opinion et d'expression, 10 août 2011. Disponible sous <https://documents-dds-ny.un.org/doc/UNDOC/GEN/N11/449/79/PDF/N1144979.pdf?OpenElement>.

⁹⁰ Article 3, paragraphe 1 (c) du Protocole facultatif à la convention relative aux droits de l'enfant concernant la vente d'enfants, la prostitution des enfants et la pornographie mettant en scène des enfants : les États doivent veiller à ce que leur droit pénal couvre « le fait de produire, de distribuer, de diffuser, d'importer, d'exporter, d'offrir, de vendre ou de détenir (...) des matériels pornographiques mettant en scène des enfants ». Disponible sous : <https://www.ohchr.org/fr/instruments-mechanisms/instruments/optional-protocol-convention-rights-child-sale-children-child>.

⁹¹ Article III (c) de la Convention pour prévention et la répression du crime de génocide qui stipule que « l'incitation directe et publique à commettre un génocide » sera punie. Disponible sous : <https://www.ohchr.org/fr/instruments-mechanisms/instruments/convention-prevention-and-punishment-crime-genocide>.

⁹² Voir article 20 du Pacte pour les droits civils et politiques. Disponible sous : <https://www.ohchr.org/fr/instruments-mechanisms/instruments/international-covenant-civil-and-political-rights>.

⁹³ Conseil de sécurité, résolution S/RES/1624 (2005). Disponible sous : <https://documents-dds-ny.un.org/doc/UNDOC/GEN/N05/510/53/PDF/N0551053.pdf?OpenElement>. L'incitation au terrorisme constitue une quatrième forme d'incitation proscrite par le droit international depuis cette résolution du Conseil de sécurité, dans laquelle celui-ci a appelé les États à « interdire par la loi l'incitation à commettre des actes de terrorisme » et à « empêcher toute incitation à commettre de tels actes ». Le droit international ne fournit pas

rapporteur spécial appelle les Etats à mettre en œuvre l'interdiction de l'apologie de la haine au sens de l'article 20 du Pacte sur les droits civils et politique, tout en les reliant à l'appel à la discrimination, à l'hostilité ou à la violence. Le rapporteur ne s'arrête pas là. Pour lui, « il importe d'établir une nette distinction entre les trois modes d'expressions suivants :

- a) ceux qui constituent une infraction au regard du droit international et sont passibles de poursuites pénales,
- b) ceux qui ne sont passibles de poursuites pénales mais devraient faire l'objet de restrictions et de poursuites au civil ; et
- c) ceux qui ne sont passibles ni de sanctions pénales ni de sanctions civiles mais qui sont néanmoins inquiétants dans la mesure où ils n'obéissent pas aux principes de tolérance, de civilité et de respect d'autrui ».⁹⁴

Cette catégorisation des propos montre que la question de ce qui est dit sur internet dépasse le problème des discours de haine, lesquels constituent une des infractions du droit pénal international, pour aussi constituer des atteintes à la personnalité ou encore des propos qui ne présentent pas un degré d'intensité justifiant une action juridique, mais qui n'en sont pas moins blessants pour les individus. Pour cette dernière situation, le rapporteur estime que « dans le cas des expressions qui ne donnent pas lieu à des sanctions pénales ou civiles mais suscitent malgré tout une préoccupation sur le plan de la civilité et du respect d'autrui, il faudrait s'efforcer de s'attaquer aux causes profondes du problème, y compris l'intolérance, le racisme et le sectarisme, en appliquant des stratégies axées sur la prévention »⁹⁵. Ainsi, pour lutter contre les comportements qui font état d'un manque de civilité et de respect d'autrui, il préconise de renforcer l'éducation et la communication.

En outre, concernant la tension entre liberté d'expression et interdiction des propos de haine, le rapporteur précise que : « ce qui distingue les actes visés à l'article 20 d'autres actes qui

de définition du terrorisme. Le rapporteur spécial pour la liberté d'expression et d'opinion a proposé de considérer que « constitue une infraction le fait de diffuser ou de mettre un message à disposition du public par tout autre moyen, délibérément et illégalement, avec l'intention d'inciter à la commission d'une infraction terroriste, lorsqu'un tel comportement, qu'il préconise expressément ou non la commission d'infractions terroristes, crée un danger qu'une ou plusieurs de ces infractions soient commises » (voir rapport A/HRC/16/51 de Monsieur Martin Scheinin, Rapporteur spécial sur la promotion et la protection des droits de l'homme et des libertés fondamentales dans la lutte antiterroriste, 22 décembre 2010, §32. Disponible sous <https://documents-dds-ny.un.org/doc/UNDOC/GEN/G10/178/99/PDF/G1017899.pdf?OpenElement>). Cette formulation remplit deux conditions, à savoir : « a) l'intention que ce message incite à la commission d'un acte terroriste, et b) l'infraction doit comporter un risque réel que l'acte préconisé par l'incitation sera commis » (Rapport A/66/290, *idem* n89, p. 12, §33).

⁹⁴ Rapport A/66/290, *idem* n89, p.8, §18. Cette distinction en trois catégories trouve son origine dans une étude préparée aux fins de la réunion d'experts régionaux sur l'article 20, organisée à Vienne par le Haut-Commissariat, les 8 et 9 avril 2010. Bukovska, B., Callamard, A., Parrmer, S. (2010). Towards an interpretation of article 20 of the ICCPR: Thresholds for the prohibition of incitement to hatred. Disponible sous : www.ohchr.org/Documents/Issues/Expression/ICCPR/Vienna/CRP7Callamard.pdf.

⁹⁵ Rapport A/66/290, *idem* n89, p.15, §40 ; ce point se retrouve également dans les recommandations relatives à l'accès au contenu d'internet (pp. 24-25, §83-84).

peuvent également être soumis à une restriction conformément au paragraphe 3 de l'article 19, c'est que pour les premiers, le Pacte indique la réponse précise attendue de l'État : leur interdiction par la loi. Ce n'est que dans cette mesure que l'article 20 peut être considéré comme une « *lex specialis* » à l'égard de l'article 19 »⁹⁶. En d'autres termes : « les articles 19 et 20 sont compatibles l'un avec l'autre et se complètent. Les actes visés à l'article 20 tombent sous le coup des restrictions énoncées au paragraphe 3 de l'article 19. De ce fait, une limite qui est justifiée par l'invocation de l'article 20 doit également être conforme au paragraphe 3 de l'article 19 ».⁹⁷

Enfin, le rapport de 2011 révèle un fléchissement dans la conception du rôle de l'Etat sur Internet. Ainsi, à propos de la lutte contre le terrorisme, le rapporteur estime que les États, plutôt que de restreindre l'accès à internet comme ce fut noté dans des rapports précédents, devraient tirer avantage de cette infrastructure et des acteurs qui permettent son utilisation : « les moyens dont on dispose pour empêcher la diffusion de contenus considérés comme constituant une incitation au terrorisme sont souvent insuffisants ou inefficaces, voire les deux, et il peut être par conséquent plus utile de concevoir des stratégies qui s'appuient sur Internet plutôt que de chercher à lutter contre, y compris en diffusant rapidement des contre-communications pour lutter contre les messages extrémistes qui constituent une incitation au terrorisme »⁹⁸. Le rapport évoque ainsi le cas des blocages des contenus qui est alors décrit comme la méthode la plus couramment employée pour restreindre les types d'expression interdits sur Internet. Il demande toutefois aux États de donner des précisions complètes sur la nécessité et les motifs justifiant le blocage d'un site Web particulier et que « les critères utilisés pour déterminer les contenus devant être bloqués doivent être définis par une autorité judiciaire compétente ou par un organe indépendant de toute pression politique ou commerciale ou de toute autre influence injustifiée, afin d'éviter que le blocage ne soit utilisé comme moyen de censure »⁹⁹. Cet exemple montre ainsi que l'Etat n'est plus seulement perçu dans une position de surplomb, mais qu'il doit également être un acteur à l'intérieur de l'infrastructure.

Le rapport de 2012 est dédié à la question spécifique du « discours haineux » et à l'incitation à la haine. Le rapporteur spécial souligne la difficulté persistante à trouver des solutions permettant de concilier la nécessité de protéger et promouvoir le droit à la liberté d'opinion et d'expression, d'une part, et de lutter contre la discrimination et l'incitation à la haine, d'autre part¹⁰⁰. Il réitère son appel pour que seuls les cas d'une gravité extrême soient considérés

⁹⁶ Rapport A/66/290, *idem* n89, p.10, §27.

⁹⁷ Nations Unies, Comité des droits de l'homme. *Idem* n21.

⁹⁸ Rapport A/66/290, *idem* n89, p. 13, §35.

⁹⁹ Rapport A/66/290, *idem* n89, p. 14, §38, citant le Rapport A/HRC/17/27 de Monsieur Franck La Rue, Rapporteur spécial sur la promotion et la protection du droit à la liberté d'opinion et d'expression, 16 mai 2011, p. 21, §70. Disponible sous :

<https://undocs.org/Home/Mobile?FinalSymbol=A%2FHRC%2F17%2F27&Language=E&DeviceType=Desktop&LangRequested=False>

¹⁰⁰ Rapport A/67/357 de Monsieur Franck La Rue, Rapporteur spécial sur la promotion et la protection du droit à la liberté d'opinion et d'expression, 7 septembre 2012, p. 4, §3. Disponible sous

comme des infractions pénales. Pour cela, il « encourage les États à fixer des seuils élevés et rigoureux, en prenant en compte les éléments ci-après : gravité, intention, teneur, portée, probabilité de nuire, imminence et contexte. Un tel examen doit être effectué au cas par cas, en considérant le contexte »¹⁰¹. Autrement dit, le rapporteur spécial appelle à l'adoption de normes pénales suffisamment précises pour justifier l'atteinte à la liberté. Par ailleurs, il estime que « s'agissant de la diffusion du discours haineux en ligne, les États devraient être habilités à requérir l'élimination des contenus incriminés seulement sur ordre de la justice et que *les intermédiaires ne devraient jamais être tenus responsables des contenus* dont ils ne sont pas les auteurs. Le droit de s'exprimer en ligne sous couvert d'anonymat doit également être pleinement garanti »¹⁰².

Pour les autres propos qui ne relèvent pas du champ de l'infraction pénale, le rapporteur spécial recommande aussi aux États d'adopter des textes sur le plan du droit civil qui prévoieraient divers recours de procédure ou des recours quant au fond pour restaurer la réputation, empêcher toute récidive et accorder une indemnité financière. Il indique encore que « des formations devraient être proposées aux magistrats afin qu'ils appréhendent avec clarté et logique les formes de discours haineux et les critères qui les définissent en droit international »¹⁰³. Des « possibilités d'éducation permanente en matière de prescriptions nationales et internationales, dont les critères permettant de définir l'incitation, devraient être offertes en grand nombre aux juristes et aux agents de la force publique »¹⁰⁴.

Le rapport de 2013 revient sur la question du rôle de l'Etat évoqué dans le rapport de 2011. Il est consacré à la surveillance des communications par les États¹⁰⁵. Le rapporteur y observe que les Etats déploient des services pour contrôler ce qui est diffusé sur internet. Ainsi, il note qu'« outre les technologies qui facilitent le filtrage, nombre d'États organisent un filtrage manuel sur Internet, en mettant en place des forces de police et des inspecteurs en ligne qui surveillent matériellement le contenu des sites WEB, les réseaux sociaux, les blogs et autres formes de médias »¹⁰⁶. Le rapporteur indique aussi que « dans certains États, des « forces de cyberpolice » sont chargées d'inspecter et de contrôler l'Internet, de rechercher des sites Web et les nœuds critiques qu'ils peuvent contenir (en particulier dans les forums de

<https://undocs.org/Home/Mobile?FinalSymbol=A%2F67%2F357&Language=E&DeviceType=Desktop&LangRequested=False>

¹⁰¹ *Ibid.*, p. 24, §79.

¹⁰² Rapport A/67/357, *idem* n°100, p. 25, §87.

¹⁰³ *Ibid.*, p. 24, §82.

¹⁰⁴ *Ibid.*

¹⁰⁵ Rapport A/HRC/23/40 de Monsieur Franck La Rue, Rapporteur spécial sur la promotion et la protection du droit à la liberté d'opinion et d'expression, 17 avril 2013, p. 14, §46. Disponible sous : <https://undocs.org/Home/Mobile?FinalSymbol=A%2FHRC%2F23%2F40&Language=E&DeviceType=Desktop&LangRequested=False>.

¹⁰⁶ *Ibid.*, p. 14, §46.

discussion en ligne), en vue de bloquer ou de fermer les sites Web dont le gouvernement désapprouve le contenu, notamment les critiques à l'égard du pouvoir en place »¹⁰⁷.

En outre, il constate que « cette surveillance est confiée à des intermédiaires privés tels que des moteurs de recherche et des plateformes de réseaux sociaux, grâce à des lois qui élargissent la responsabilité des contenus prohibés, de l'intervenant original à l'ensemble des intermédiaires »¹⁰⁸. Autrement dit, la surveillance d'internet n'est pas seulement le fait des autorités publiques, mais elle intègre également les entreprises qui offrent des services de moteur de recherche et de réseaux sociaux.

Enfin, il observe un délitement des garanties de procédures. Ce délitement est observé au niveau des autorités étatiques qui ne présentent pas toujours des garanties d'indépendance : « bien que la surveillance des communications doive traditionnellement être autorisée par le pouvoir judiciaire, cette règle est de plus en plus souvent allégée ou supprimée. Dans certains pays, l'interception des communications peut être autorisée par un ministre du gouvernement, son délégué ou une commission. Au Royaume-Uni par exemple, elle est autorisée par le secrétaire d'État (Art. 5, loi de 2000 sur la réglementation des pouvoirs d'enquête) »¹⁰⁹. Le délitement des garanties de procédure tient aussi du fait que les États s'appuient aussi sur des acteurs privés pour contrôler internet : « (l)es obligations des États en matière de droits de l'homme exigent non seulement qu'ils respectent et encouragent les droits à la liberté d'expression et au respect de la vie privée, mais qu'ils protègent les individus des violations des droits fondamentaux perpétrés par les entreprises. En outre, les États devraient exercer un contrôle adéquat, de manière à remplir leurs obligations internationales en matière de droits de l'homme lorsqu'ils recourent aux services des entreprises ou prévoient de le faire, si cela peut avoir une incidence sur l'exercice des droits de l'homme »¹¹⁰. Face à une telle situation, le rapporteur considère que « la législation doit stipuler que la surveillance des communications par l'État ne doit intervenir que dans les circonstances les plus exceptionnelles et exclusivement sous le contrôle d'une autorité judiciaire indépendante. Des garanties doivent être clairement énoncées dans la loi en ce qui concerne la nature, l'étendue et la durée des mesures possibles, les motifs requis pour les ordonner, les autorités compétentes pour les autoriser, les mettre en œuvre et les superviser, et le type de recours prévu par le droit interne »¹¹¹.

Au terme de l'analyse des rapports publiés dans la première partie des années 2010, il faut retenir que les rapporteurs spéciaux ont œuvré à la clarification des discours en ligne problématiques – ceux qui relèvent du droit pénal international, ceux qui doivent faire l'objet de restrictions au niveau civil et ceux qui, sans relever ni d'une infraction pénale ou d'une procédure civile, n'en restent pas moins inquiétants. Du point de vue des appels à la haine, ils ont insisté sur le besoin de préciser les hypothèses de l'interdiction. Les rapports ont

¹⁰⁷ *Ibid.*

¹⁰⁸ Rapport A/HRC/23/40, *idem* n° 05, p. 14, §46.

¹⁰⁹ *Ibid.*, p. 15, §54.

¹¹⁰ *Ibid.*, p. 22, §76

¹¹¹ *Ibid.*, p. 23, §81.

également montré un élargissement de la palette des réponses réglementaires avec l'idée qu'il ne fallait pas se limiter à une réponse légale, mais intégrer aussi des mesures incitatives comme des programmes d'éducation et de sensibilisation. Ils ont encore montré un changement dans le rôle de l'Etat qui impose aussi aux intermédiaires en ligne de contribuer à la régulation des discours en ligne.

4.3. Au cours de la seconde moitié des années 2010 : la montée en puissance des entreprises privées comme intermédiaires de la régulation et l'élargissement de la conception du discours de haine

Les rapports publiés au cours de la seconde moitié des années 2010 se caractérisent par le rôle accru des entreprises privées dans le domaine de la régulation des propos en ligne.

Ainsi, en 2016, le rapporteur spécial livre un état des lieux de la situation d'internet. Il décrit l'importance structurelle des acteurs privés à l'ère numérique : « (...) le secteur privé occupe un rôle de plus en plus considérable ; il est en effet devenu le fer de lance de la plus grande expansion de l'accès à l'information de l'histoire. Les grands forums de médias sociaux où s'expriment les usagers appartiennent à des sociétés privées. Les principales plateformes, qui regroupent et indexent les connaissances mondiales, et qui élaborent des algorithmes qui déterminent les informations qui apparaissent en ligne, sont issues d'initiatives privées. Les infrastructures de technologie mobile qui permettent à des milliards de personnes de communiquer et de consulter Internet dépendent d'investissement privés, de services de maintenance privés et, en définitive, des décisions d'actionnaires du secteur privé »¹¹².

Plus encore, le rapport révèle que les entreprises sont vues comme des intermédiaires susceptibles de restreindre aussi la liberté d'expression : « les politiques menées concernant les intermédiaires et la réglementation de leur activité peuvent avoir un effet considérable sur la liberté d'expression. Si les conditions d'utilisation des intermédiaires constituent la source principale de réglementation, les choix en matière de conception et d'ingénierie peuvent également avoir un effet sur la livraison des contenus »¹¹³. À cet égard, le rapporteur note que les règles privées des entreprises peuvent également jouer ce rôle de restriction : « les conditions d'utilisation, que les particuliers doivent en général accepter pour pouvoir accéder à une plateforme, contiennent souvent des restrictions sur le partage des contenus. Ces restrictions varient en fonction des lois et réglementations locales et contiennent des interdictions récurrentes, parmi lesquelles celles relative au harcèlement, aux discours de haine, à l'incitation aux activités délictueuses, à la violence gratuite et aux menaces directes. Les conditions d'utilisations sont souvent formulées de manière tellement générale qu'il peut être difficile de savoir à l'avance avec une certitude raisonnable quel contenu peut faire l'objet

¹¹² Rapport A/HRC/32/38 de Monsieur David Kaye, Rapporteur spécial sur la promotion et la protection du droit à la liberté d'opinion et d'expression, 11 mai 2016, p. 3, §1. Disponible sous : <https://undocs.org/Home/Mobile?FinalSymbol=A%2FHRC%2F32%2F38&Language=E&DeviceType=Desktop&LangRequested=False>.

¹¹³ *Ibid.*, p. 16, §51.

d'une restriction »¹¹⁴. Le rapporteur observe également que dans ce rapport entre l'intermédiaire et l'utilisateur, les États peuvent aussi s'immiscer. Il considère ainsi que « les États utilisent de plus en plus fréquemment les conditions d'utilisation pour retirer les contenus qu'ils jugent douteux. (...) le service chargé de la lutte contre l'utilisation d'Internet à des fins manœuvre du Royaume-Uni, par exemple, se consacre à retirer les contenus en ligne de « nature extrémiste, violente ou terroriste », notamment en recourant à « des mécanismes de signalement des contenus des sites web, qui lui permettent de déterminer que tel ou tel contenu constitue une violation des « conditions d'utilisation) du site ». De telles pratiques peuvent laisser penser que les États pourraient s'appuyer sur les conditions d'utilisation du secteur privé pour contourner les droits de l'homme ou les lois nationales relatives aux restrictions de contenu »¹¹⁵. Le rapport de 2016 met ainsi bien en évidence la relation Etat (*regulator*), utilisateurs (*targets*) et entreprises (*intermediaries*) particulièrement complexe dès lors que l'Etat peut aussi intervenir dans le rapport direct entre les utilisateurs et les entreprises.

En 2017, le rapporteur spécial met en évidence le pouvoir des États sur certains acteurs clés d'Internet : « les États sont particulièrement bien placés pour faire pression sur les fournisseurs d'accès à internet qu'ils tiennent pour responsables des contenus affichés par les utilisateurs sur leurs réseaux ou encore faire pression sur eux pour qu'ils restreignent ces contenus en vertu de lois destinées à réprimer des infractions allant du discours haineux à la diffamation, en passant par la cybercriminalité et le lèse-majesté. Or, cette responsabilité des intermédiaires incite fortement à la censure, les fournisseurs pouvant estimer qu'il est plus risqué de contester la loi que de limiter les contenus à ce point strictement que l'expression légitime et licite s'en trouve restreinte »¹¹⁶. Ainsi, le rapporteur souligne l'instrumentalisation des lois pénales qui est faite par certains États pour contraindre les entreprises de réseaux sociaux pour réguler les propos en ligne. Il prône alors de travailler non seulement avec les États, mais également avec les entreprises privées : « le fait qu'à l'ère numérique, l'action des autorités et celle des entreprises se superposent à certains égards reste une idée relativement nouvelle pour bon nombre d'États. Il serait utile, tant à l'échelon national qu'au niveau international, d'élaborer des plans d'action nationaux consacrés à la question des entreprises et des droits de l'homme et d'y définir les méthodes permettant à tous les acteurs du numérique de mesurer l'incidence de leurs activités sur le respect des droits de l'homme et de prendre les mesures pour y faire face »¹¹⁷. Autrement dit, il s'agit pour le rapporteur d'aider les entreprises à se conformer au droit des droits humains dans leur activité de régulation de l'espace numérique.

¹¹⁴ *Ibid.*, p. 16, §52.

¹¹⁵ Rapport A/HRC/32/38, *idem* n° 12, pp. 16-17, §53.

¹¹⁶ Rapport A/HRC/35/22 de Monsieur David Kaye, Rapporteur spécial sur la promotion et la protection du droit à la liberté d'opinion et d'expression, 30 mars 2017, p.16, §49. Disponible sous : <https://www.undocs.org/Home/Mobile?FinalSymbol=A%2FHRC%2F35%2F22&Language=E&DeviceType=Desktop&LangRequested=False>.

¹¹⁷ *Ibid.*, p. 24, §81.

Le rapport de 2018 est spécialement dédié à la réglementation du contenu en ligne. Le rapporteur note que « les États cherchent à façonner l'environnement dans lequel les entreprises modèrent les contenus, tandis que les entreprises fondent l'accès des utilisateurs à leurs plateformes sur l'acceptation des conditions d'utilisation, qui régissent ce qui peut être exprimé et la manière dont cela peut être exprimé »¹¹⁸. Il constate aussi que les États exigent des entreprises qu'elles imposent des restrictions concernant les contenus manifestement illégaux, mais que certains vont plus loin en adoptant des lois restrictives formulées en termes généraux sur l'« extrémisme », le blasphème, les fausses informations¹¹⁹. Le rapporteur fait aussi écho aux initiatives privées comme les principes de Manille de 2014 sur la responsabilité des entreprises de réseaux sociaux, qui ont défini des principes essentiels permettant de guider les autorités publiques qui souhaiteraient régir la responsabilité de ces entreprises¹²⁰. Enfin, en lien avec les rapports entre entreprises et utilisateurs, le rapporteur souligne que les premières exigent des seconds qu'ils respectent les conditions d'utilisation et les normes communautaires qui régissent l'expression sur leurs plateformes¹²¹. Les entreprises déploient également des équipes de « confiance et sécurité », incluant des conseillers juridiques, des gestionnaires de politiques publiques et de produits ainsi que des cadres supérieurs¹²².

Évoquant clairement la question du discours de haine, le rapporteur note que les politiques des entreprises en matière de haine, de harcèlement et de mauvais traitements n'indiquent pas clairement ce qui constitue une infraction¹²³. Autrement dit, la catégorisation des discours de haine sur la base du droit pénal, civil et des normes sociales pensée dans le cadre de la régulation par l'État n'est pas adaptée dans le contexte de son utilisation par des entreprises privées. Il observe en effet un manque de précision qui a entraîné des effets négatifs pour les minorités et le renforcement des positions dominantes. Il ajoute que « les lois nationales ne sont pas adaptées aux entreprises qui souhaitent s'appuyer sur des normes communes applicables à leur communauté d'utilisateurs, dont la diversité est importante du point de vue

¹¹⁸ Rapport A/HRC/38/35 de Monsieur David Kaye, Rapporteur spécial sur la promotion et la protection du droit à la liberté d'opinion et d'expression, 6 avril 2018, p. 6, §12. Disponible sous : <https://undocs.org/Home/Mobile?FinalSymbol=A%2FHRC%2F38%2F35&Language=E&DeviceType=Desktop&LangRequested=False>.

¹¹⁹ Rapport A/HRC/38/35, *idem* n° 18, pp. 6-7, §13.

¹²⁰ Rapport A/HRC/38/35, *idem* n° 18, p. 7, §14. Le rapporteur discute également brièvement l'exemple allemand, estimant que « les règles sur la responsabilité comportent des risques pour la liberté d'expression car elles font peser une pression importante sur les entreprises, de sorte que celles-ci peuvent supprimer des contenus licites pour éviter, de manière générale, que leur responsabilité soit engagée. Elles impliquent également la délégation de fonctions de réglementation à des acteurs privés qui ne disposent pas des outils de base pour l'établissement des responsabilités. Les demandes de suppressions rapides et automatiques risquent de donner lieu à de nouvelles formes de restrictions préalables qui menacent déjà la créativité dans le contexte des droits d'auteur. Les questions complexes de fait et de droit devraient en général être tranchées par des institutions publiques et non par des acteurs privés dont les procédures peuvent être incompatibles avec les règles d'une procédure régulière et dont les motifs sont essentiellement d'ordre économique » ; *ibid.*, p. 8, §17.

¹²¹ *Ibid.*, p. 11, §24.

¹²² *Ibid.*

¹²³ *Ibid.*, p. 11, §26.

géographique et culturel »¹²⁴. Il considère que « les normes privées, qui varient en fonction du modèle économique de l'entreprise et reposent sur des déclarations vagues concernant l'existence d'une communauté d'intérêts, ont créé des environnements instables, imprévisibles et peu sûrs pour les utilisateurs, dans lesquels les pouvoirs publics exercent une surveillance accrue »¹²⁵. Tout en restant attaché aux principes du droit des droits de l'homme, le rapporteur spécial estime que « les entreprises devraient toutefois énoncer les fondements de ces restrictions et démontrer la nécessité et la proportionnalité de toute mesure prise à l'égard de contenu (comme les suppressions ou les suspensions de compte) »¹²⁶.

Dans ce contexte, le rapporteur indique plusieurs pistes dans une logique « approche par défaut fondée sur les droits de l'homme ». Il s'agit d'abord de mettre en place un système de responsabilité (« *accountability* »). À ce titre, il considère que les normes de droit international des droits de l'homme offrent une alternative « permettant de demander aux États ainsi qu'aux entreprises de rendre des comptes aux utilisateurs indépendamment des frontières nationales »¹²⁷. Ensuite, les normes des entreprises devraient respecter le principe de sécurité juridique. Ainsi, le rapporteur considère que toute réglementation doit s'inscrire dans le paradigme de la liberté et de la protection de la vie privée : « elles devraient faire en sorte que toute ambiguïté dans la loi soit interprétée en faveur de la liberté d'expression et du respect du droit à la vie privée et des autres droits de l'homme »¹²⁸. Très concrètement, il considère que « les entreprises devraient faire directement figurer dans leurs conditions d'utilisation et dans leurs « normes applicables à la communauté » les principes pertinents du droit des droits de l'homme selon lesquels les mesures ayant des incidences sur les contenus doivent être conformes aux principes de légalité, de nécessité et de légitimité par lesquels les États sont liés lorsqu'ils réglementent la liberté d'expression »¹²⁹. À ce titre, il critique le fait que « les règles des entreprises sont généralement dépourvues de la clarté et de la précision qui permettrait aux utilisateurs de prédire avec un degré de certitude raisonnable quel type de contenu leur ferait franchir la ligne rouge »¹³⁰. À son sens, « les entreprises ne devraient pas se borner à décrire plus en détail les règles controversées et les règles propres à un contexte particulier ; elles devraient aussi fournir des données et des exemples permettant de se faire une idée des éléments pris en considération dans la détermination d'une violation, de sa gravité et des mesures en conséquence »¹³¹. Enfin, dans le cas spécifique du discours de haine, le rapporteur estime qu'« expliquer la façon dont certains cas ont été réglés peut permettre aux utilisateurs de mieux comprendre comment les entreprises abordent la question de la distinction délicate entre les contenus offensants et l'incitation à la haine, ou comment des éléments tels que l'intention de l'auteur des propos ou le risque de violence sont appréciés

¹²⁴ *Ibid.*, pp. 16-17, §41.

¹²⁵ *Ibid.*

¹²⁶ Rapport A/HRC/38/35, *idem* n° 18., p.12, §28.

¹²⁷ *Ibid.*, p. 17, §41.

¹²⁸ *Ibid.*, p. 19, §50.

¹²⁹ *Ibid.*, p. 18, §45.

¹³⁰ *Ibid.*, p. 18, §46.

¹³¹ *Ibid.*, p. 18, §47.

lorsqu'il s'agit de contenus diffusés en ligne ». ¹³² Il demande aussi qu'il soit prévu que les entreprises « (expliquent) dans quelles circonstances les entreprises appliquent des restrictions moins intrusives (avertissements, limites d'âge ou démonétisation) » ¹³³. Autrement dit, le rapporteur préconise une approche fondée sur la transparence pour les entreprises.

Le rapport de 2019 porte spécifiquement sur le discours de haine en ligne. Il indique que plusieurs États ont adopté ou envisagent d'adopter des lois obligeant les entreprises de réseaux sociaux à supprimer les discours « manifestement contraire à la loi » dans un délai donné, généralement dans les 24 heures, voire rien qu'en une heure, ou alors à éliminer, dans un délai plus long, les contenus interdits par la loi » ¹³⁴. Évoquant le cas spécifique de la législation allemande, le rapporteur considère que « même si la loi « NetzDG » doit être comprise comme un effort entrepris de bonne foi pour répondre aux inquiétudes générales que suscitent les discours haineux en ligne et leurs conséquences hors ligne, l'absence de définition des termes clés qui y sont utilisés fait qu'il est hasardeux d'affirmer que les obligations découlant de ses dispositions sont conformes au droit international relatif aux droits de l'homme » ¹³⁵. Il observe encore que « rares sont les États qui ont fait participer leurs autorités judiciaires à l'évaluation des plateformes diffusant des discours haineux interdits par leur droit interne ; or, les États ne devraient permettre l'attribution de la responsabilité que sur ordre des tribunaux indépendants et avec la possibilité, pour l'intermédiaire ou toute autre partie intéressée (comme l'utilisateur concerné), de faire appel de la décision en question. Les pouvoirs publics font de plus en plus pression sur les entreprises pour qu'elles jouent un rôle d'arbitre face aux discours haineux » ¹³⁶. Ainsi, le rapporteur souligne l'importance du rôle que font peser les États sur les entreprises en ligne et sur l'absence d'un régime de garanties de procédures pour les utilisateurs.

Le rapporteur estime que « les États devraient, en fait, adopter des lois et des politiques qui amènent les entreprises à protéger la liberté d'expression et à lutter contre les formes de discours haineux légalement restreintes, en exigeant la transparence requise pour permettre au public d'exercer un contrôle en veillant à l'application du droit interne par des autorités judiciaires indépendantes (...) » ¹³⁷. Autrement dit, les États devraient réguler les intermédiaires en ligne dans leur rôle de régulateur des contenus en ligne. Cette régulation de l'intermédiaire doit également répondre à certaines exigences. Ainsi, l'intervention publique devrait viser à ce que les entreprises adoptent une attitude plus responsable et plus protectrice de la cohésion de nos sociétés. Dans cette philosophie, l'intervention de l'État doit alors « 1) respecter la diversité des modèles de réseaux sociaux qui forment un ensemble particulièrement

¹³² *Ibid.*

¹³³ *Ibid.*

¹³⁴ Rapport A/74/486 de Monsieur David Kaye, Rapporteur spécial sur la promotion et la protection du droit à la liberté d'opinion et d'expression, 9 octobre 2019, p.15, §32. Disponible sous : <https://undocs.org/Home/Mobile?FinalSymbol=A%2F74%2F486&Language=E&DeviceType=Desktop&LangRequested=False>.

¹³⁵ *Ibid.*, p. 15, §32.

¹³⁶ *Ibid.*, pp. 15-16, §33.

¹³⁷ *Ibid.*, p. 16, §35.

hétéroclite, 2) faire preuve de transparence, notamment en y associant la société civile, 3) viser un objectif d'intervention minimum conformément aux principes de nécessité et de proportionnalité et 4) s'en remettre aux juridictions pour la qualification de la licéité des contenus pris individuellement »¹³⁸.

Dans ses recommandations à destination des Etats, le rapporteur établit alors en 2019 une feuille de route indiquant aux Etats ce qu'il préconise, à savoir :

- Définir les formes d'expression interdites au titre du paragraphe 2 de l'article 20 du Pacte I et de l'article 4 de la CERD ;
- S'abstenir d'ériger en infractions pénales l'emploi de ces expressions, sauf dans les cas les plus graves comme celui de l'appel à la haine nationale, raciale ou religieuse constitutifs d'une incitation à la discrimination, à l'hostilité ou à la violence et relayer les interprétations du droit des droits de l'homme ;
- Revoir les textes régissant la responsabilité des intermédiaires, ou en adopter de nouveaux, pour les mettre en totale conformité avec les normes relatives aux droits de l'homme et s'abstenir d'exiger des entreprises qu'elles restreignent la liberté d'expression de leurs utilisateurs si les Etats ne sont pas eux-mêmes en droit d'imposer directement ces restrictions par voie législative ;
- Mettre en place des mécanismes judiciaires indépendants ;
- Astreindre les entreprises à donner publiquement une définition détaillée de ce qu'elles considèrent comme des discours haineux, à décrire en pareils termes la manière dont elles appliquent leurs règles relatives à ces discours et à garder trace, dans des bases de données, des mesures qu'elles prennent contre quiconque tient de tels propos¹³⁹.

*

Au terme de ce chapitre, il ressort que l'analyse des rapports du rapporteur à la liberté d'opinion et d'expression a permis de montrer que dans le cadre de l'environnement en ligne, une interprétation élargie a été donnée à la notion de propos haineux pour privilégier l'expression discours de haine. En outre, la régulation des propos tenus en ligne n'est plus appréhendée à travers le prisme Etat-individu, mais dans la relation triangulaire entre Etat-entreprises en ligne-utilisateurs. Dans ce nouveau schéma, la régulation des propos en ligne a été déferée aux intermédiaires en ligne, les entreprises en ligne, c'est-à-dire des acteurs privés. Autrement dit, la question des discours de haine a été doublement redéfinie : du point de vue de son contenu et du point de vue de l'organe de régulation.

¹³⁸ *Ibid.*, p. 17, §37. Le Rapporteur spécial cite un rapport français : Rapport de mission « Créer un cadre français de responsabilisation des réseaux sociaux : agir en France avec une ambition européenne », mai 2019. Disponible sous : https://www.economie.gouv.fr/files/files/2019/Mission_Regulation_des_reseaux_sociaux-Rapport_public.pdf?v=1557755659

¹³⁹ Rapport A/74/486, *idem* n° 34, pp. 25-26, §57.

5. Analyse des réglementations applicables aux entreprises en ligne

Dans cette section, il s'agit d'examiner si les législations nationales et européenne ont connu une évolution similaire et si leur dispositif juridique fait écho aux recommandations du rapporteur spécial des nations unies. Pour cela, cinq réglementations seront analysées en suivant la chronologie de leur adoption : les législations allemande, française, autrichienne, européenne et britannique. Cette analyse se basera sur une grille d'analyse développée à partir des éléments mis en évidence dans la section précédente que nous synthétisons ci-dessous. Avant cela, chaque législation sera replacée dans son contexte politique.

Grille analytique :

Critères d'analyse	Description
<i>Champ d'application</i>	Qui est visé par la réglementation ? Quels types de contenu sont visés ? Quels sont les comportements et discours qui entrent dans le champ d'application du texte de loi ?
<i>Règles organisationnelles dans le rapport entre autorités et entreprises</i>	Quelles sont les mesures qui sont imposées aux entreprises de réseaux sociaux dans leurs processus d'identification, d'évaluation et de retrait des discours de haine ?
<i>Règles de procédure dans le rapport entre entreprises et utilisateurs</i>	Quelles sont les règles qui organisent le rapport entre les utilisateurs et les entreprises en lien avec les discours de haine sur les plateformes ?
<i>Autorités spéciales</i>	Qui sont les acteurs externes appelés à contribuer à la façon dont les entreprises de réseaux sociaux identifient, évaluent et retirent les discours de haine ? Qui sont les acteurs externes qui appuient et/ou contrôlent la mise en œuvre des obligations législatives ?

5.1. Allemagne : *Netzwerkdurchsetzungsgesetz (NetzDG)*

L'Allemagne est le premier État membre de l'Union européenne à s'être doté d'une loi régissant spécifiquement les obligations des entreprises de réseaux sociaux en matière de discours de haine avec la *Netzwerkdurchsetzungsgesetz (NetzDG)*¹⁴⁰.

¹⁴⁰ *Netzwerkdurchsetzungsgesetz*. Disponible sous <https://www.gesetze-im-internet.de/netzdg/BJNR335210017.html>

5.1.1. Contexte

Le texte législatif puise son origine dans une proposition parlementaire datant du 16 mai 2017. Le rapport accompagnant la proposition de loi indique que la proposition législative vise à répondre aux nouveaux défis posés par le développement d'Internet, au motif que celui-ci a généré « un changement massif des discours de la société sur la toile et en particulier sur les réseaux sociaux, où la culture du débat est souvent agressive, blessante et souvent haineuse »¹⁴¹. Il est également dit que pour lutter contre ce phénomène, « le cadre pénal existant en Allemagne est clair et suffisant ». Toutefois, en raison de la quantité des contenus partagés, de leur distribution immédiate et de l'anonymat de nombreux auteurs, sa mise en œuvre rencontre des difficultés et n'est pas satisfaisante, puisque les individus victimes de discours haineux en ligne sont insuffisamment protégés¹⁴². C'est donc en vue de palier à un problème de mise en œuvre que l'adoption d'une nouvelle loi se justifie.

La proposition de loi du 16 mai 2017 a été discutée au Bundestag le 19 mai 2017¹⁴³. À la suite de cette première discussion, le gouvernement allemand a rendu un avis sur le texte¹⁴⁴, et a chargé la commission parlementaire *Recht und Verbraucherschutz* de rédiger un rapport intégrant ses propositions. Ce rapport a été rendu le 28 juin 2017, synthétisant les discussions et recommandant l'adoption du texte de loi¹⁴⁵. Une seconde discussion au Bundestag a eu lieu le 30 juin 2017, aboutissant à l'adoption du texte par le Bundestag¹⁴⁶. Le 7 juillet 2017, le Bundesrat a suivi la recommandation d'adoption qui lui avait été faite par le Bundestag et a également adopté le texte¹⁴⁷. La loi constituée de 6 articles est entrée en vigueur le 7 septembre 2017.

5.1.2. Champ d'application

La NetzDG a pour champ d'application personnel les entreprises de réseaux sociaux (lesquelles sont comprises comme « des plateformes internet qui permettent, en vue d'obtenir un profit, l'échange de contenus entre utilisateurs ou envers le public »¹⁴⁸) qui ont plus de deux

¹⁴¹ Deutscher Bundestag – Gesetzentwurf der Fraktionen der CDU/CSU und SPG. Drucksache 18/12356, p.1. Disponible sous : <https://dserver.bundestag.de/btd/18/123/1812356.pdf>, consulté le 10 février 2023.

¹⁴² *Ibid.*, p.11.

¹⁴³ Deutscher Bundestag – Stenografischer Bericht. 18. Wahlperiode – 235. Sitzung. Berlin, Freitag, den 19. Mai 2017. Disponible sous : <https://dserver.bundestag.de/btp/18/18235.pdf>, consulté le 11 février 2023.

¹⁴⁴ Gesetzentwurf der Bundesregierung – Drucksache 18/12727. Disponible sous : <https://dserver.bundestag.de/btd/18/127/1812727.pdf>. Consulté le 11 février 2023.

¹⁴⁵ Deutscher Bundestag – Beschlussempfehlung und Bericht des Ausschusses für Recht und Verbraucherschutz. Drucksache 18/13013. Disponible sous : <https://dserver.bundestag.de/btd/18/130/1813013.pdf>. Consulté le 11 février 2023.

¹⁴⁶ Deutscher Bundestag – Stenografischer Bericht. 18. Wahlperiode – 235. Sitzung. Berlin, Freitag, den 30. Juni 2017. Disponible sous : <https://dserver.bundestag.de/btp/18/18244.pdf>. Consulté le 11 février 2023.

¹⁴⁷ Deutscher Bundesrat – Beschluss des Bundesrates – Gesetz zur Verbesserung der Rechtsdurchsetzung in sozialen Netzwerken. Drucksache 536/17. Disponible sous : <https://dserver.bundestag.de/brd/2017/0536-17.pdf>. Consulté le 11 février 2023.

¹⁴⁸ *Netzwerkdurchsetzungsgesetz*, *idem* n° 40, §1(1).

millions d'utilisateurs et d'utilisatrices inscrits sur leurs services en Allemagne. La loi englobe ainsi les principaux réseaux sociaux. Elle ne s'applique en revanche pas aux services d'audio- et de télédiffusion (§1(1)).

La législation impose à ces entreprises d'agir contre les « contenus contraires au droit » (« *Rechtswidrige Inhalte* »). Ces « contenus contraires au droit » sont les comportements réprimés dans les articles 86, 86a, 89a, 91, 100a, 111, 126, 129 bis à 129b, 130, 131, 140, 166, 184b, 185 à 187, 189, 201a, 241 et 269 du code pénal allemand¹⁴⁹ (« *Strafgesetzbuch* »¹⁵⁰). En l'occurrence, les comportements nécessitant une intervention des entreprises de réseaux sociaux sont :

- §86 - Diffusion de matériel de propagande d'organisations anticonstitutionnelles et terroristes
- §86a - Utilisation de signes distinctifs d'organisations anticonstitutionnelles et terroristes
- §89a - Préparation d'un acte de violence grave menaçant la sûreté de l'État
- §91 - Instructions en vue de commettre un acte de violence grave menaçant la sûreté de l'État
- §100a - Falsifications trahissant la nation
- §111 - Provocation publique à commettre des infractions
- §126 - Perturbation de la paix publique par la menace d'infractions pénales
- §129 - Formation d'organisations criminelles
- §129a - Formation de groupes terroristes
- §129b - Organisations criminelles et terroristes à l'étranger ; confiscation
- §130 - Incitation à la haine
- §131 - Représentation de la violence
- §140 - Récompense et approbation de délits
- §166 - Injures à l'encontre de confessions, de sociétés religieuses et d'associations de conviction
- §184b - Diffusion, acquisition et détention de contenus pédopornographiques
- §185 - Insulte
- §186 - Diffamation
- §187 - Calomnie
- §189 - Dénigrement de la mémoire des défunts
- §201a - Atteinte à la vie privée et aux droits de la personnalité par l'enregistrement d'images
- §241 - Menace
- §269 - Falsification de données à caractère probant

Ce faisant, le législateur allemand n'évoque pas directement la notion de discours de haine, mais inclut les comportements qui entrent dans le champ des comportements proscrits par l'article 20 du Pacte sur les droits civils et politiques des Nations Unies, certaines autres dispositions du droit pénal international (en lien avec des activités à des fins de terrorisme et de pédopornographie notamment) ainsi que d'autres dispositions pénales comme les atteintes à l'intégrité physique et psychique et à l'ordre démocratique.

4.1.3. Règles organisationnelles applicables à l'activité des entreprises de réseaux sociaux

Les entreprises de réseaux sociaux doivent désigner un mandataire ayant domicile en Allemagne pour garantir la communication et la collaboration avec les organes étatiques chargés de la mise en œuvre et du contrôle de la loi.

¹⁴⁹ *Netzwerkdurchsetzungsgesetz*, idem n° 40, §1(3).

¹⁵⁰ *Strafgesetzbuch*. Disponible sous : <https://www.gesetze-im-internet.de/stgb/>

De plus, les entreprises sont tenues de publier un rapport en langue allemande qui explique comment les obligations législatives sont mises en œuvre¹⁵¹. Ce rapport doit contenir les informations suivantes :

- 1 – des explications générales sur les efforts déployés par le fournisseur de réseau social pour empêcher les actes punissables sur les plateformes ;
- 2 – la nature, les grandes lignes de fonctionnement et la portée des procédures utilisées, le cas échéant, pour la reconnaissance automatique des contenus ;
- 3 – une présentation des mécanismes de transmission des plaintes, une exposition des critères de décision pour le retrait et le blocage des contenus illicites et une présentation de la procédure d'examen ;
- 4 – le nombre de plaintes reçues au cours de la période de référence ;
- 5 – l'organisation, la dotation en personnel, les compétences professionnelles et linguistiques des unités de travail chargées de traiter les plaintes ;
- 6 – l'appartenance à des organisations sectorielles ;
- 7 – le nombre de plaintes pour lesquelles un organisme externe a été consulté afin de préparer la décision ;
- 8 – le nombre de plaintes ayant abouti à la suppression ou au blocage du contenu incriminé au cours de la période de référence ;
- 9 – le nombre de plaintes relatives à des contenus illicites qui ont été retirées ou bloquées dans les 24 heures, les 48 heures, la semaine ou ultérieurement ;
- 10 – les mesures prises pour informer le plaignant, ainsi que l'utilisateur pour lequel le contenu contesté a été stocké, de la décision prise concernant la plainte ;
- 11 – le nombre de contre-propositions reçues durant la période de référence ;
- 12 – le nombre de contre-propositions reçues durant la période de référence pour lesquelles il a été renoncé à faire un examen ;
- 13 – des informations indiquant si et dans quelle mesure les milieux scientifiques et de la recherche ont eu accès aux informations du fournisseur au cœur de la période de référence, notamment en lien avec les contenus illicites retirés ou bloqués ;
- 14 – toute autre mesure prise par le fournisseur pour protéger et aider les personnes concernées par des contenus illicites ;
- 15 – un résumé sous forme de tableau comparant le nombre total de plaintes reçues concernant des contenus illicites, le pourcentage de contenus retirés ou bloqués à la suite de ces plaintes, le nombre de contre-propositions reçues, ainsi que le pourcentage de décisions modifiées à la suite de ces contre-propositions ;
- 16 – une explication des dispositions des conditions générales du fournisseur concernant l'admissibilité de la diffusion de contenus sur le réseau social que le fournisseur utilise pour les contrats avec les consommateurs.

Le rapport doit être rendu public tous les 6 mois, par le biais d'une publication sur le site internet de l'intermédiaire lui-même ainsi qu'au journal officiel (« *Bundesanzeiger* »).

4.1.4. Règles de procédure applicables dans le rapport entre entreprises et utilisateurs

Le législateur allemand a prévu le déroulement d'une procédure particulière en lien avec les contenus contraires au droit.

Tout d'abord, il a institué une procédure de signalement permettant aux utilisateurs d'annoncer les contenus contraires au droit. Tout signalement impose une obligation d'agir à charge des entreprises de réseaux sociaux, ainsi qu'une obligation de communiquer la décision prise. Le législateur opère une distinction entre « un contenu **manifestement** contraire au droit » et « un contenu contraire au droit ». Les entreprises ont l'obligation de retirer les contenus manifestement contraires au droit ou de bloquer l'accès à ces contenus dans un délai

¹⁵¹ *Netzwerkdurchsetzungsgesetz*, idem n° 40, §2.

de 24 heures. Dans le cas où le contenu signalé constitue un « contenu contraire au droit » (c'est-à-dire que son illicéité n'est pas manifeste mais doit faire l'objet d'un examen plus approfondi), les entreprises de réseaux sociaux doivent alors le retirer dans un délai de 7 jours. Pour permettre le fonctionnement d'une telle procédure, les entreprises doivent prendre immédiatement connaissance des signalements et en contrôler le contenu (§3, paragraphe 2, alinéa 1).

Ensuite, le législateur a prévu une procédure de recours contre les décisions prises à la suite d'un signalement. L'article 3b de la loi¹⁵² prévoit en effet la possibilité pour les utilisateurs de soumettre des « contre-propositions » (« *Gegenvorstellungen* »). Cette possibilité est ouverte aux utilisateurs dont les contenus ont été retirés et qui souhaitent s'opposer à ces retraits ainsi qu'aux utilisateurs qui ont soumis un signalement aux entreprises et qui sont en désaccord avec l'issue donnée à leur signalement (en principe, une décision de ne pas retirer un contenu). Le droit de recours implique de motiver la contre-proposition et de la déposer dans un délai de deux semaines à partir du moment où le contenu illicite a été retiré ou qu'une décision de non-retrait a été prononcée. Cette procédure de contre-proposition requiert que les entreprises mettent en place des mesures organisationnelles. Les entreprises ont ainsi l'obligation de mettre en place un formulaire de contact en ligne leur permettant de réceptionner directement ces contre-propositions. Dans le cas où une contre-proposition est soumise, les entreprises sont obligées de réexaminer la décision initiale. Elles peuvent décider d'accepter la contre-proposition et, le cas échéant, soit rétablir un contenu qui aurait été indûment retiré (lorsque c'est un utilisateur dont le contenu a été retiré qui a soumis une contre-proposition), soit retirer un contenu qui ne l'aurait pas été (lorsque c'est un utilisateur qui avait soumis un signalement qui a soumis une contre-proposition).

Enfin, la NetzDG prévoit que des **instances de médiation** (« *privatrechtlich organisierte Einrichtungen als Schlichtungsstellen* ») peuvent être instituées. Leur mission est d'offrir un service de médiation dans le cadre du règlement (extrajudiciaire) des litiges entre les utilisateurs et les entreprises de réseaux sociaux¹⁵³. Les utilisatrices et utilisateurs peuvent saisir cette institution au terme de la procédure de contre-procédure ou lorsque les entreprises ont fait appel à une entité externe d'auto-régulation réglementée (voir point 4.1.5.) pour les aider à prendre une décision.

4.1.5. Autorités spéciales

Le législateur allemand a prévu l'intervention de plusieurs organismes dans le processus de modération des contenus. Ces organismes interviennent à l'égard des entreprises en qualité de conseil, de recours ou de surveillance ou peuvent être saisies par les utilisateurs.

Ainsi, dans le cadre de la procédure d'appréciation de l'illicéité du contenu, les entreprises de réseaux sociaux peuvent requérir le conseil d'une instance indépendante. Cette entité est désignée comme étant une « **entité de l'auto-régulation réglementée** » (« *Einrichtung der*

¹⁵² *Netzwerkdurchsetzungsgesetz, idem n° 40, §3b.*

¹⁵³ *Netzwerkdurchsetzungsgesetz, idem n° 40, §3c, par. 1.*

Regulierten Selbstregulierung »)¹⁵⁴. Elle peut être mise en place par les réseaux sociaux afin de fournir des services d'audit à ceux-ci et les appuyer dans la prise de décision quant aux retraits des contenus illicites. Les auditeurs engagés par cette entité doivent être indépendants et doivent respecter un règlement qui régit l'étendue et le déroulement de l'examen des contenus.

Dans le cas de recours contre les décisions rendues par les entreprises en ligne, il y a également les **instances de médiation** déjà mentionnées.

Enfin, le respect de la NetzDG relève de la compétence du **ministère fédéral de la justice** (« *Bundesamt für Justiz* »). Ce ministère a la compétence de prendre des mesures à l'encontre des intermédiaires en ligne, notamment des amendes en cas de non-respect de la loi. En tant qu'autorité administrative en charge de l'application de la loi, il a la compétence d'attribuer leur statut aux entités d'autorégulation réglementée ainsi qu'aux organes de médiation. Le **bureau fédéral de la police criminelle** (« *Bundeskriminalamt* ») constitue l'autorité de poursuite pour les infractions commises sur les plateformes¹⁵⁵.

5.2. France : Loi en vue de lutter contre les discours de haine sur les réseaux sociaux (Loi Avia)

La France est le deuxième État membre de l'Union européenne à adopter une législation en vue de lutter contre les discours de haine sur les réseaux sociaux, avec la *Loi Avia*.

5. 2.1. Contexte

C'est au cours des années 2017/2018 que la question de la lutte contre les discours de haine devient un enjeu dont se saisissent les institutions publiques en France. Le gouvernement adopte un plan pour la période 2018-2020 de lutte contre le racisme et l'antisémitisme¹⁵⁶. Parallèlement, la députée Laetitia Avia est chargée en mars 2018 de formuler des mesures concrètes pour lutter contre les délits de discours raciste et antisémites sur les réseaux sociaux. Dans son rapport, elle détaille quelles nouvelles règles pourraient être adoptées et souligne comment l'adoption de la NetzDG en Allemagne a commencé à faire bouger les choses au niveau de l'Union Européenne relativement à la question de l'encadrement des

¹⁵⁴ *Netzwerkdurchsetzungsgesetz*, *idem* n° 40, §3, par. 6.

¹⁵⁵ Il doit notamment être informé par les entreprises en ligne dans certaines circonstances (les infractions mettant en danger la sûreté de l'État ou les infractions liées à la pédocriminalité) en vue de permettre la poursuite d'infractions pénales, en vertu de l'article 3a, paragraphe 2.

¹⁵⁶ Délégation interministérielle à la lutte contre le racisme, l'antisémitisme et la haine anti-LGBT. (2018). Plan national de lutte contre le racisme et l'antisémitisme 2018-2020. Disponible sous : <https://www.gouvernement.fr/plan-national-de-lutte-contre-le-racisme-et-l-antisemitisme-2018-2020>

entreprises de réseaux sociaux¹⁵⁷. À la suite de ce rapport, la députée Laetitia Avia dépose un projet de loi réglementant les obligations de ces entreprises¹⁵⁸.

La procédure législative s'est déroulée du 19 mars 2019 au 25 juin 2020, date où le texte de loi a été publié au journal officiel, et a donné lieu à un intense débat parlementaire entre les deux chambres. La procédure a débuté à l'Assemblée nationale où le texte de loi a été examiné en commission des lois et a donné lieu à un premier rapport n°2062 du 19 juin 2019¹⁵⁹. Le texte a été adopté le 9 juillet 2019 en séance publique par l'Assemblée nationale¹⁶⁰. Il a ensuite été transmis au Sénat le même jour. La commission des lois a été saisie au fond et deux autres commissions ont été saisies pour avis. Au cours des travaux de la commission des lois, des amendements ont été déposés et un rapport les résumant a été rendu le 11 décembre 2019¹⁶¹. Sur la base de ce rapport, le texte a été modifié avant d'être adopté par le Sénat le 17 décembre 2019¹⁶². Le Sénat ayant amendé le texte transmis par l'Assemblée nationale, une commission mixte paritaire a été saisie de la proposition de loi le 8 janvier 2020¹⁶³, en vue de régler les désaccords entre les deux chambres. Un rapport conjoint a été produit par la commission mixte paritaire à cette même date¹⁶⁴. Le texte a été renvoyé à l'Assemblée nationale pour une nouvelle lecture, qui a donné lieu au rapport du 15 janvier 2020¹⁶⁵. Le texte a été adopté par l'Assemblée nationale le 22 janvier 2020¹⁶⁶. À la suite du vote de l'Assemblée nationale, le texte a été renvoyé au Sénat le même jour. La commissions des lois a proposé de nouveaux amendements, conduisant à un second rapport rendu le 5 février 2020¹⁶⁷. Le texte a été discuté en séance publique, et plusieurs amendements ont à nouveau été déposés.

¹⁵⁷ La députée relève notamment que si la France venait à suivre l'Allemagne en adoptant une loi, la « convergence franco-allemande constituerait une base robuste pour faire évoluer le droit européen dans le même sens ». Rapport du 20 septembre 2018 - Renforcer la lutte contre le racisme et l'antisémitisme sur Internet, p.5. Disponible sous : https://medias.vie-publique.fr/data_storage_s3/rapport/pdf/184000622.pdf.

¹⁵⁸ Proposition de loi n° 1785 visant à lutter contre la haine sur internet. Disponible sous : https://www.assemblee-nationale.fr/dyn/15/textes/115b1785_proposition-loi.

¹⁵⁹ Assemblée nationale, Rapport n°2062, 19 juin 2019. Disponible sous : https://www.assemblee-nationale.fr/dyn/15/rapports/cion_lois/115b2062_rapport-fond#

¹⁶⁰ Assemblée nationale, Analyse du scrutin n°2039, 9 juillet 2019. Disponible sous : [https://www2.assemblee-nationale.fr/scrutins/detail/\(legislature\)/15/\(num\)/2039](https://www2.assemblee-nationale.fr/scrutins/detail/(legislature)/15/(num)/2039)

¹⁶¹ Sénat, Commission des lois, Rapport n°197 (2019-2020), 11 décembre 2019. Disponible sous : <https://www.senat.fr/rap/119-197/119-197.html>

¹⁶² Sénat, Texte n°36 rectifié, 17 décembre 2019. Disponible sous : <https://www.senat.fr/leg/tas19-036.html>

¹⁶³ Commission mixte paritaire, Rapport sur la proposition de loi visant à lutter contre les contenus haineux sur internet, 8 janvier 2020. Disponible sous : <https://www.senat.fr/compte-rendu-commissions/20200106/cmp.html#toc3>

¹⁶⁴ Commission mixte paritaire, Rapport n°239, 8 janvier 2020. Disponible sous : <https://www.senat.fr/rap/119-239/119-239.html>

¹⁶⁵ Assemblée nationale, Rapport n°2583, 15 janvier 2020. Disponible sous : https://www.assemblee-nationale.fr/dyn/15/rapports/cion_lois/115b2583_rapport-fond#

¹⁶⁶ Assemblée nationale, Proposition de loi n°388, 22 janvier 2020. Disponible sous : https://www.assemblee-nationale.fr/dyn/15/textes/115t0388_texte-adopte-seance#

¹⁶⁷ Sénat, Rapport n°299, 5 février 2020. Disponible sous : <https://www.senat.fr/rap/119-299/119-299.html>

Le texte modifié a été voté le 26 février 2020¹⁶⁸. Après la discussion et le vote en seconde lecture au Sénat, le texte a été renvoyé à l'Assemblée nationale le 28 février 2020. Celle-ci a adopté définitivement le texte le 13 mai 2020¹⁶⁹.

À l'issue de l'adoption définitive du texte de loi, des sénateurs ont saisi le Conseil constitutionnel, le 18 mai 2020. En date du 18 juin 2020, celui-ci a partiellement invalidé le texte législatif. Le Conseil constitutionnel a estimé¹⁷⁰ que le texte de loi, en obligeant les entreprises de réseaux sociaux à supprimer des contenus en ligne dans des délais très courts (parfois dans l'heure suivant leur signalement), constituait une atteinte non tolérable à la liberté d'expression des individus. Ainsi, selon le Conseil constitutionnel, malgré la bonne volonté du législateur de vouloir faire cesser la prolifération de contenus illégaux en ligne, la solution retenue porte « atteinte à la liberté d'expression et de communication » d'une façon qui « n'est pas adaptée, nécessaire et proportionnée au but poursuivi »¹⁷¹. Finalement, c'est donc une version partielle du texte législatif qui est entrée en vigueur le 25 juin 2020¹⁷².

Bien que le texte législatif ait été partiellement amputé de son dispositif, l'analyse de la version adoptée par les deux chambres, c'est-à-dire avant l'invalidation par le Conseil Constitutionnel, permet de cerner la réponse régulatoire qui avait été adoptée pour lutter contre le discours de haine.

5.2.2. Champ d'application

La « Loi Avia » s'applique aux « opérateurs de plateforme en ligne (...) qui proposent un service de communication au public en ligne reposant sur la mise en relation de plusieurs parties en vue du partage de contenus publics et dont l'activité sur le territoire français dépasse des seuils déterminés par décret »¹⁷³.

Comme dans le cas allemand, le texte ne définit pas ce qui constitue les discours ou contenus haineux. Il vise « tout contenu contrevenant manifestement aux dispositions mentionnées aux cinquièmes, septièmes et huitièmes alinéas de l'article 24, à l'article 24bis et aux troisièmes et quatrièmes alinéas de l'article 33 de la loi du 29 juillet 1881 sur la liberté de la presse, aux articles 222-33, 227-23 et 421-2-5 du code pénal ainsi que, lorsque l'infraction porte sur un contenu à caractère pornographique, à l'article 227-24 du même code »¹⁷⁴. Le législateur français vise ainsi:

¹⁶⁸ Sénat, Texte n°64, 26 février 2020. Disponible sous : <https://www.senat.fr/leg/tas19-064.html>

¹⁶⁹ Assemblée nationale, Texte n°419, 13 mai 2020. Disponible sous : https://www.assemblee-nationale.fr/dyn/15/textes/15t0419_texte-adopte-seance#

¹⁷⁰ Conseil constitutionnel, Décision n°2020-801, 18 juin 2020. Disponible sous : <https://www.conseil-constitutionnel.fr/decision/2020/2020801DC.htm>

¹⁷¹ *Ibid.*, p.4.

¹⁷² Loi n°2020-766 du 24 juin 2020 visant à lutter contre les contenus haineux sur Internet. Disponible sous : <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000042031970>

¹⁷³ Assemblée nationale, Texte n°419, *idem* n°169, p.2.

¹⁷⁴ *Ibid.*, p.3.

- La complicité par le discours (selon les termes de l'article 23 de la loi sur la liberté de la presse¹⁷⁵) à l'apologie des crimes de guerre, des crimes contre l'humanité, des crimes de réduction en esclavage et des crimes de collaboration avec l'ennemi (Art. 24, al. 5 de la loi sur la liberté de la presse¹⁷⁶) ;
- la provocation à la discrimination, à la haine ou à la violence à l'égard d'une personne ou d'un groupe de personnes à raison de leur origine ou de leur appartenance ou de leur non-appartenance à une ethnie, une nation, une race ou une religion déterminée (Art. 24, al. 7 de la loi sur la liberté de la presse) ;
- la provocation à la haine ou à la violence à l'égard d'une personne ou d'un groupe de personnes à raison de leur sexe, de leur orientation sexuelle ou identité de genre ou de leur handicap (Art. 24 al. 8 de la loi sur la liberté de la presse) ;
- l'injure (Art. 33 de la loi sur la liberté de la presse) ;
- le harcèlement sexuel (Article 222-33 du code pénal¹⁷⁷) ;
- la pédopornographie (Article 227-23 du code pénal) ;
- l'apologie du terrorisme (Article 421-2-5 du code pénal) ;
- la diffusion de contenu à caractère terroriste ou pédopornographique (Article 227-24 du code pénal).

Les propos susceptibles de faire l'objet d'une restriction sont donc liés aux infractions de droit pénal international en lien avec le génocide, la pédopornographie, le terrorisme et l'incitation à la haine. À cela sont ajoutés les enjeux de harcèlement sexuel et de l'appel à la discrimination. Sur ce plan, le législateur français va ainsi plus loin que le législateur allemand.

5.2.3. Règles organisationnelles applicables à l'activité des entreprises de réseaux sociaux

Le législateur français requiert que les entreprises de réseaux sociaux désignent « une personne physique située sur le territoire français exerçant les fonctions d'interlocuteur référent chargé de recevoir les demandes de l'autorité judiciaire en application de la loi » (art. 9 al 5). Il s'agit ainsi pour le législateur de garantir que les autorités disposent d'un interlocuteur établi sur le territoire de la République, permettant de faciliter les interactions avec les utilisateurs ainsi qu'avec les autorités de surveillance.

La loi prévoit également que les entreprises de réseaux sociaux doivent formuler leurs conditions générales « en termes précis, aisément compréhensibles, objectifs et non discriminatoires » (art. 5 al 10). Ces entreprises doivent aussi « rendre compte des moyens humains et technologiques qu'ils mettent en œuvre et des procédures qu'ils adoptent pour se conformer à leurs obligations » (art. 5 al. 6), c'est-à-dire rendre un rapport détaillant ces moyens et leur mise en œuvre.

¹⁷⁵ L'article précédent est l'article 23 de la loi sur la liberté de presse qui puni « comme complices d'une action qualifiée de crime ou de délit ceux qui, soit par des discours, cris ou menaces proférés dans des lieux ou réunions publics, soit par des écrits, imprimés, dessins, gravures, peintures, emblèmes, images ou tout autre support de l'écrit, de la parole ou de l'image venus ou distribués, mis en vente ou exposés dans des lieux ou réunions publics, soit par des placards ou des affiches exposés dans des lieux ou réunions publics, soit par tout moyen de communication au public par voie électronique, auront directement provoqué l'auteur ou les auteurs à commettre ladite action, si la provocation a été suivie d'effet ».

¹⁷⁶ Loi du 29 juillet 1881 sur la liberté de la presse. Disponible sous : <https://www.legifrance.gouv.fr/loda/id/LEGISCTA000006089701>

¹⁷⁷ Code pénal français. Disponible sous : https://www.legifrance.gouv.fr/codes/texte_lc/LEGITEXT000006070719/2020-12-16

5.2.4. Règles de procédure applicables dans le rapport entre entreprises et utilisateurs

Les entreprises de réseaux sociaux doivent mettre en place « un dispositif de notification uniforme directement accessible et facile d'utilisation permettant à toute personne de notifier un contenu illicite dans la langue d'utilisation du service » (art. 4 al. 4).

Dans le cadre de cette procédure de notification, les entreprises de réseaux sociaux doivent prendre position par rapport aux signalements, et sont tenus d'informer les utilisateurs en cas de retrait de leur contenu de la décision de retrait ainsi que des raisons qui l'ont motivées (art. 7 al. 2). Le législateur français impose donc une obligation de motivation des décisions.

La loi prévoit une procédure de recours interne pour les cas où les entreprises rendent inaccessibles un contenu qui leur a été notifié (art. 5 al. 5).

5.2.5. Autorités spéciales

Plusieurs autorités sont amenées par la loi à intervenir dans le processus de contrôle de l'activité des entreprises de réseaux sociaux. Le Conseil supérieur de l'audiovisuel (CSA) est désigné comme l'autorité chargée de surveiller la mise en œuvre des dispositions légales par les entreprises (art. 16). Il s'agit donc d'une extension des compétences de l'autorité administrative indépendante qui régule les radios, les télévisions et les services de vidéo à la demande¹⁷⁸. Un « Observatoire de la haine en ligne » dont la mission est « d'assurer le suivi et l'analyse de l'évolution des contenus mentionnés à l'article 1^{er} de la loi » (art. 16) doit également être mis en place au sein du CSA pour être spécifiquement dédié à la surveillance du phénomène de discours de haine. Finalement, la loi Avia requière également que le gouvernement désigne ultérieurement, par décret, un tribunal judiciaire spécialisé dans la poursuite des infractions pénales qui font l'objet d'une plainte adressée par voie électronique (art. 10).

5.3. Autriche : *Kommunikationsplattformengesetz (Ko-PI)*

La république d'Autriche constitue le troisième Etat membre de l'Union européenne à avoir adopté une loi réglementant les entreprises de réseaux sociaux. Ici encore, l'objectif de la Ko-PI a été de clarifier le régime de responsabilités de ces entreprises quant à ce que l'Etat attend

¹⁷⁸ Le Conseil supérieur de l'audiovisuel (CSA) a été institué par la loi n°89-25 du 17 janvier 1989 (disponible sous : [Loi n° 89-25 du 17 janvier 1989 modifiant la loi n° 86-1067 du 30 septembre 1986 relative à la liberté de communication - Légifrance \(legifrance.gouv.fr\)](#)), en modification de la loi Léotard du 30 septembre 1986 relative à la liberté de communication (disponible sous : [Loi n° 86-1067 du 30 septembre 1986 relative à la liberté de communication \(Loi Léotard\) - Légifrance \(legifrance.gouv.fr\)](#)). Le CSA agit au nom de l'Etat de français, qui lui a délégué sa compétence de régulation pour le secteur audiovisuel. Le CSA a ainsi la compétence de s'assurer du respect et de la protection des droits et libertés individuels dans le secteur audiovisuel, ainsi que d'assurer la bonne régulation économique et technologique de ce marché.

d'eux en matière de lutte contre les contenus haineux et de permettre aux utilisateurs de disposer d'un mécanisme de plainte efficace¹⁷⁹.

5.3.1. Contexte

La Ko-PI a été adoptée durant l'année 2020, dans le contexte de la pandémie de Covid-19. Cette loi s'inscrit dans un paquet législatif plus large intitulé « *Hass im Netz* »¹⁸⁰, qui prévoit une réactualisation du code pénal et civil autrichien face aux nouveaux développements liés au digital. Parmi les mesures du paquet « *Hass im Netz* », la loi sur les plateformes de communication est l'élément central.

Le projet a été déposé par le gouvernement autrichien le 9 septembre 2020¹⁸¹. Le 18 novembre 2020, le Conseil national autrichien a pris connaissance du projet de loi et a décidé de le transférer pour consultation au comité constitutionnel (« *Verfassungsausschuss* »), qui est une commission parlementaire chargée d'examiner les textes de loi pouvant avoir un impact sur la Constitution autrichienne. Ce dernier a rendu son rapport le 25 novembre 2020¹⁸². Sur la base du projet de loi et du rapport du comité constitutionnel, le conseil national a adopté le texte lors de sa séance du 10 décembre 2020¹⁸³. Le projet adopté a ensuite été transmis au Bundesrat autrichien, lequel l'a à son tour adopté lors de sa séance du 17 décembre 2020¹⁸⁴. Le texte législatif a été publié au journal officiel le 23 décembre 2020 et est entrée en vigueur le 1^{er} avril 2021¹⁸⁵. La mise en œuvre de la Ko-PI à la suite de son entrée en vigueur a donné lieu à une question préjudicielle auprès de la Cour de Justice de l'Union Européenne (CJUE), dont nous détaillons les implications à la fin du rapport¹⁸⁶. Indépendamment de la décision de la CJUE et de ses implications pour le texte de loi, l'analyse du dispositif permet de cerner quelles sont les mesures que le législateur autrichien a envisagées.

¹⁷⁹ Bundesministerium Justiz, Regierungsvorlage, Erläuterung zum Gesetzestext. Disponible sous : [fname_848587.pdf \(parlament.gv.at\)](https://www.parlament.gv.at/XXVII/I/509/fname_848587.pdf)

¹⁸⁰ Bundesministerium Justiz, Hass im Netz. Disponible sous : <https://www.bmj.gv.at/themen/Fokusthemen/gewalt-im-netz.html>

¹⁸¹ Ministerialentwurf betreffend Bundesgesetz, mit dem ein Bundesgesetz über Massnahmen zum Schutz der Nutzer auf Kommunikationsplattformen erlassen wird, 9 septembre 2020. Disponible sous : <https://www.parlament.gv.at/gegenstand/XXVII/ME/49>

¹⁸² Bericht des Verfassungsausschusses, 25 novembre 2020. Disponible sous : https://www.parlament.gv.at/dokument/XXVII/I/509/fname_851399.pdf

¹⁸³ Beschluss des Nationalrates vom 10. Dezember 2020. Disponible sous : <https://www.parlament.gv.at/gegenstand/XXVII/BNR/154>

¹⁸⁴ Beschluss im Bundesrat 154/BNR. Disponible sous : <https://www.parlament.gv.at/gegenstand/XXVII/BNR/154>

¹⁸⁵ Kommunikationsplattformengesetz. Disponible sous : <https://www.ris.bka.gv.at/GeltendeFassung.wxe?Abfrage=Bundesnormen&Gesetzesnummer=2001415&FassungVom=2023-07-09>

¹⁸⁶ Cour Européenne de Justice, Google Ireland Limited, Meta platforms Ireland Limited, Tik Tok Technology Limited contre Kommunikationsbehörde Austria (KommAustria), Arrêt du 9 novembre 2023, Affaire C-376-22. Disponible sous : <https://curia.europa.eu/juris/documents.jsf?num=C-376/22>

5.3.2. Champ d'application

Comme dans les cas précédents, la loi s'applique aux « fournisseurs de services nationaux et étrangers qui, dans un but lucratif, mettent en place des plateformes de communication ». Elle concerne toutefois les fournisseurs qui ont plus de 100'000 utilisateurs.

La loi autrichienne ne définit pas ce que recouvre la notion de discours de haine et elle instaure une obligation similaire à celle de la NetzDG, à savoir retirer les contenus « contraires au droit ». Elle renvoie à une liste de dispositions du code pénal autrichien¹⁸⁷:

- §105 Contrainte
- §107 Menace dangereuse
- §107a Persécution persistante
- §107c harcèlement persistant par voie de télécommunications
- §115 insultes
- §120a prise d'images non autorisée
- §144 extorsion
- §188 dégradation de signes religieux
- §207a représentation pornographique de mineurs
- §208a initiation de contacts sexuels avec des mineurs
- §278b association terroriste
- §278f incitation à la commission d'une infraction terroriste
- §282a encouragement à des actes terroristes et apologie d'actes terroristes
- §283 incitation à la haine

Parmi ces dispositions figure la disposition du §283 sur l'incitation à la haine, laquelle vise : « Celui qui (...) incite à la violence ou à la haine à l'égard d'une église ou d'une société religieuse, ou d'un autre groupe de personnes défini par les critères existants ou manquants de la race, de la couleur de peau, de la langue, de la religion ou des convictions, de la nationalité, de l'ascendance ou de l'origine nationale ou ethnique, du sexe, du handicap, de l'âge ou de l'orientation sexuelle (...) ». Elle vise aussi certaines infractions en lien avec le terrorisme et avec la pédopornographie ainsi qu'en lien avec la liberté de religion, l'atteinte à l'image et le harcèlement.

5.3.3. Règles organisationnelles applicables aux activités des entreprises de réseaux sociaux

Comme dans les cas allemand et français, les entreprises doivent désigner une personne établie sur le territoire autrichien en tant que point de contact pour les communications avec les autorités administratives et les tribunaux (art. 5).

Également comme dans les cas allemands et français, les entreprises de réseaux sociaux doivent annuellement rédiger un rapport sur leurs activités de modération de contenu (art. 4) :

- 1- Des explications générales sur les efforts déployés par l'intermédiaire pour empêcher la présence de contenus illicites sur la plateforme ;
- 2- Une présentation de la conception et du caractère convivial du processus de signalement qui a été mis en place, ainsi qu'une présentation des critères de décision utilisés pour la suppression ou le blocage des contenus illicites ;

¹⁸⁷ Strafgesetzbuch. Disponible sous : <https://www.jusline.at/gesetz/stgb>

- 3- Une présentation du nombre de notifications de contenus illicites présumés reçues au cours de la période de référence ;
- 4- Un aperçu du nombre de notifications qui ont conduit à la suppression ou au blocage de contenu ;
- 5- Un aperçu du nombre, du contenu et du résultat des procédures de vérification (voir point ci-dessous) ;
- 6- Une présentation de l'organisation des personnes chargées du traitement des notifications ;
- 7- Un aperçu des délais entre la réception d'un signalement et le début de l'examen et la suppression ou le blocage d'un contenu illicite ;
- 8- Un aperçu du nombre et du type de cas dans lesquels le prestataire de services a renoncé à mettre en œuvre une procédure de notification.

5.3.4. Règles de procédure applicables aux rapports entre entreprises et utilisateurs

Le législateur autrichien exige que les entreprises de réseaux sociaux mettent en place une procédure pour que les utilisateurs puissent signaler les contenus illicites présents sur leurs plateformes. La possibilité de signalement doit être « facile(s) à trouver, disponible(s) en permanence et facile(s) à utiliser » (art. 3, al. 1 et 2).

Suite à un signalement, l'entreprise doit prendre une décision relativement au contenu et doit notifier la décision et sa motivation aux utilisateurs (art. 3 al. 2). En cas de contenu manifestement contraire au droit, le contenu doit être retiré de la plateforme sous 24 heures. Dans le cas où l'illicéité n'est pas manifeste, l'intermédiaire en ligne dispose d'un délai de 7 jours. Pour le législateur, est manifestement contraire au droit le contenu dont l'illicéité est « déjà évidente pour un profane en matière juridique » (art. 3 al. 3).

De plus, les utilisateurs dont le contenu a fait l'objet d'une décision de retrait peuvent faire une demande de reconsidération (art. 3 al. 4), c'est-à-dire faire un recours contre la décision, et ils doivent être informés du résultat de cette demande.

Enfin, l'article 7 prévoit la mise en place d'un bureau des plaintes auprès duquel les utilisateurs peuvent adresser des plaintes en cas d'insatisfaction relativement aux procédures de signalement et de vérification

5.3.5. Autorités spéciales

Comme dans le cas français, le législateur européen confie la mission de surveillance de l'application de la loi à l'autorité administrative de surveillance et de communication (« *Kommunikationsbehörde Austria* », KommAustria¹⁸⁸) (art. 8). La fonction de bureau des plaintes est confiée à la Société autrichienne de régulation de la radio et des télécommunications (« *Rundfunk und Telekom Regulierungs-GmbH (RTR)* »¹⁸⁹), qui est une sous-entité de KommAustria.

¹⁸⁸ KommAustria est l'autorité de surveillance en matière audiovisuelle. Cette autorité a été mise en place par la loi du 30 mars 2001, la « *Bundesgesetz über die Einrichtung einer Kommunikationsbehörde Austria* » (disponible sous : [2001_32_1.pdf \(bka.gv.at\)](#)). KommAustria assure notamment le respect des lois en matière audiovisuelles et garantit un marché et une compétition saine entre les différents acteurs présents sur le marché.

¹⁸⁹ La RTR a également été instituée par la loi du 30 mars 2001 (*ibid.*). Parmi ses compétences, la RTR joue notamment un rôle de médiation entre les différents acteurs présents sur le marché audiovisuel et les consommateurs.

5.4. Union européenne : Règlement sur les services numériques (DSA)

5.4.1. Contexte

Le règlement européen sur les services numériques (Digital Services Act - DSA) tire son origine dans la nécessité d'adapter la directive e-Commerce (le cadre législatif prévalant jusqu'alors sur les activités des intermédiaires en ligne, incluant les entreprises de réseaux sociaux) qui date de l'an 2000¹⁹⁰ et dans le développement des réponses nationales précédemment étudiées. L'objectif du DSA est d'opérer une réforme du cadre juridique de l'UE en matière de commerce électronique « tout en maintenant les principes essentiels de son régime de responsabilité, l'interdiction des obligations générales de surveillance et la clause relative au marché intérieur, considérée comme toujours valable »¹⁹¹. Dans cette perspective, les nouvelles règles doivent permettre la réalisation d'un environnement numérique compétitif en Europe, tout en apportant de la clarté juridique aux plateformes et aux utilisateurs. Pour les entreprises de réseaux sociaux, cela passe par une définition claire de leurs obligations « de diligence raisonnable » et l'introduction de la possibilité pour les utilisateurs de contester les décisions relatives à des retraits de contenus notamment. Par cette harmonisation à l'échelle de l'union, la loi doit apporter « plus d'équité, de transparence et de responsabilité dans les processus de modération de contenu » pour les utilisateurs¹⁹².

Le texte de loi émane d'une proposition de la Commission européenne rendue publique le 15 décembre 2020¹⁹³. Une procédure de consultation s'est déroulée durant l'année 2021, et le Conseil de l'Union Européenne a travaillé sur le texte sur la base de ces retours et des consultations publiques. Le texte a été amendé par le parlement européen les 19 et 20 janvier 2022¹⁹⁴, pour être vu en première lecture par ce même parlement le 5 juillet 2022¹⁹⁵. Sur la base de cette première lecture, l'acte a été adopté par le Conseil de l'Union Européenne le 4 octobre 2022¹⁹⁶. Il a été publié au journal officiel le 27 octobre 2022 et est entré en vigueur le 17 février 2024¹⁹⁷.

¹⁹⁰ Directive sur les services électroniques. Disponible sous : [Directive - 2000/31 - EN - e-commerce directive - EUR-Lex \(europa.eu\)](#)

¹⁹¹ Commission Européenne, Proposition de loi, 15 décembre 2020, p.3. Disponible sous : <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020PC0825>

¹⁹² *Ibid.*

¹⁹³ *Ibid.*

¹⁹⁴ Parlement européen, Verbatim, débats du 19 janvier 2022, p. 103. Disponible sous https://www.europarl.europa.eu/doceo/document/CRE-9-2022-01-19_EN.pdf. Consulté le 12 juillet 2023.

¹⁹⁵ Parlement européen, Textes adoptés, 5 juillet 2022. Disponible sous : https://www.europarl.europa.eu/doceo/document/TA-9-2022-0269_FR.html

¹⁹⁶ Conseil Européen, Communiqué de presse du 4 octobre 2022. Disponible sous : <https://www.consilium.europa.eu/fr/press/press-releases/2022/10/04/dsa-council-gives-final-approval-to-the-protection-of-users-rights-online/>

¹⁹⁷ Digital Services Act. Disponible sous : https://eur-lex.europa.eu/legal-content/EN/TXT/?toc=OJ%3A%3A2022%3A277%3ATOC&uri=uriserv%3AOJ.L_2022.277.01.0001.01.ENG

5.4.2. Champ d'application

Le DSA s'applique aux « fournisseurs de services intermédiaires », lesquels sont définis comme les services de la société de l'information qui proposent (i) un service de transport d'information sur un réseau de communication, (ii) un service de « mise en cache » impliquant le stockage d'information en vue de leur transmission, (iii) un service d'hébergement consistant à stocker des informations (Art. 3, let. g). En outre, dans la même logique que l'Autriche, parmi les différents fournisseurs de services intermédiaires, le DSA introduit un régime différentiel lié à la taille du fournisseur : les « très grandes plateformes en ligne et très grands moteurs de recherche en ligne » et les autres plateformes. Les très grandes plateformes sont celles qui ont « un nombre mensuel moyen de destinataires actifs du service dans l'Union égal ou supérieur à 45 millions » (Art. 33 al. 1).¹⁹⁸

Le DSA s'applique aux contenus juridiquement problématiques qu'il qualifie de « contenus illicites », par lesquels sont compris « toute information qui, en soi ou par rapport à une activité, y compris la vente de produits ou la fourniture de services, n'est pas conforme au droit de l'Union ou au droit d'un État membre qui est conforme au droit de l'Union, quel que soit l'objet précis ou la nature précise de ce droit » (Article 3, lettre h). Autrement dit, ce qui relève du comportement illicite découle soit du droit national des États membres soit du droit de l'Union.

5.4.3. Règles organisationnelles relatives aux activités des entreprises de réseaux sociaux

Le DSA requiert que les entreprises de réseaux sociaux désignent un point de contact auxquels les utilisateurs peuvent s'adresser directement (art. 11 et 12) ainsi qu'un représentant légal au sein de l'Union Européenne (article 13 DSA) auquel pourront être adressés les communications d'ordre judiciaire. Il s'agit donc d'une obligation organisationnelle plus précise que ce qui avait été prévu dans les législations nationales.

Les intermédiaires Internet doivent également rédiger leurs conditions générales d'une manière qui informe les utilisateurs sur les « éventuelles restrictions qu'ils imposent en ce qui concerne l'utilisation de leur service », et s'assurer que ces conditions soient rédigées dans un langage « clair, simple, intelligible, aisément abordable et dépourvu d'ambiguïté, et soient mis à disposition du public dans un format facilement accessible et lisible par une machine » (art. 14).

Le législateur européen exige aussi que les entreprises rédigent des rapports de transparence (art. 15 et 24). Ces rapports doivent être rendus publics sur une base annuelle. Ils doivent contenir des informations sur « les activités de modération des contenus auxquelles (les entreprises) se sont livrés au cours de la période concernée ». Ils doivent inclure le nombre

¹⁹⁸ Le 25 avril 2023, la Commission européenne a arrêté la liste de ces très grandes plateformes. Il s'agit de Alibaba, Aliexpress, Amazon Store, Apple AppStore, Booking.com, Facebook, Google Maps, Google Play, Google Shopping, Instagram, LinkedIn, Pinterest, Snapchat, TikTok, Twitter, Wikipedia, Youtube, Zalando. European Commission (2023) DSA : Very large online platforms and search engines. Disponible sous <https://digital-strategy.ec.europa.eu/en/policies/dsa-vlops>. Consulté le 25 avril 2023.

d'injonctions reçues des autorités des États membres, classées par type de contenu illicite concerné, ainsi qu'indiquer s'ils ont eu recours à « des moyens automatisés à des fins de modération des contenus ».

De plus, les « très grandes plateformes en ligne » (art. 33 DSA) doivent procéder à des évaluations des risques (art. 34 DSA), notamment relativement au risque systémique de diffusion de contenus illicites par l'intermédiaire de leurs services, se doter d'un mécanisme de réaction aux crises (art. 36 DSA), et se soumettre à des audits indépendants (art. 37 DSA).

Elles doivent aussi garantir – sur demande – un accès à leurs données au coordinateur pour les services numériques de l'État membre dans lequel ils sont établis (art. 40 DSA) ainsi qu'aux chercheurs agréés qui en font la demande motivée en vue de « procéder à des recherches contribuant à la détection, au recensement et à la compréhension des risques systémiques, (...) ainsi qu'à l'évaluation du caractère adéquat, de l'efficacité et des effets des mesures d'atténuation des risques ».

5.4.4. Règles de procédure applicables aux rapports entre entreprises et utilisateurs

Les entreprises de réseaux sociaux doivent mettre en place des « mécanismes de notification et d'action » (art. 16). Ces mécanismes doivent permettre « à tout particulier ou à toute entité de leur signaler la présence au sein de leur service d'éléments d'information spécifiques que le particulier ou l'entité considère comme du contenu illicite ».

Dans le cas où une entreprise prend une décision – par exemple en retirant ou en rendant moins visible un contenu ou en suspendant un compte - de sa propre initiative ou à la suite d'une notification, il est tenu de motiver sa décision (art. 17).

Les entreprises de réseaux sociaux doivent également mettre en place un « système interne de traitement des réclamations » (art. 20). Ce système doit permettre aux « destinataires de leur service », c'est-à-dire les utilisateurs, de faire une réclamation contre une décision de retrait, de suspension de compte ou de suspension de la capacité de monétiser les contenus.

Outre l'obligation d'offrir une procédure de réclamation interne, le DSA prescrit un système de « règlement extrajudiciaire des litiges » (art. 21). Cette possibilité doit être ouverte pour les utilisateurs destinataires d'une décision pouvant faire une réclamation au sens de l'article 20 (voir ci-dessus) mais qui ne souhaitent pas utiliser la procédure de réclamation interne mise en place par l'entreprise.

5.4.5. Autorités spéciales

Comme dans le cas allemand, le législateur européen a prévu une diversité d'autorités qui exercent des activités de conseil, de recours à l'égard des décisions prises par les entreprises, de surveillance de ces entreprises ou encore de coordination.

Ainsi, le DSA institue la fonction de « *trusted flaggers* » (art. 22). Ces entités sont des organismes qui bénéficient d'un statut privilégié pour soumettre des notifications aux entreprises de réseaux sociaux. Ce statut de signaleur de confiance est attribué par les

coordinateurs pour les services numériques (voir ci-dessous), en vertu de l'article 22, alinéa 4.

En outre, la possibilité pour les utilisateurs de faire appel à un règlement extrajudiciaire des litiges exige de déterminer des instances de règlement extrajudiciaire. En l'occurrence, l'article 21 alinéa 3 du DSA détaille les conditions auxquelles des entités peuvent être certifiées en tant qu'organes de règlement extrajudiciaire des litiges.

Les autorités publiques (judiciaire ou administrative) nationales disposent du pouvoir d'adresser une injonction d'agir aux entreprises en ligne en présence d'un contenu illicite sur sa plateforme (art. 9). Dans un tel cas, l'autorité compétente doit inclure « une référence à la base juridique au titre du droit de l'Union ou du droit national » sur laquelle repose l'injonction. L'injonction doit également détailler les motifs qui expliquent « pourquoi les informations constituent un contenu illicite, en référence à une ou plusieurs dispositions spécifiques du droit de l'Union ou du droit national conforme au droit de l'Union » (art. 9, al. 2. Let. A, chiffre ii).

La surveillance du respect du DSA est exercée par la Commission européenne. Elle peut engager des procédures à l'égard de la conduite de « fournisseur de très grandes plateformes en ligne » (article 66 DSA), requérir des informations des entreprises (article 67 DSA), effectuer des inspections (article 69 DSA), prononcer aussi des sanctions, notamment en infligeant des amendes aux intermédiaires (article 74 DSA). Dans le cas de sanction administrative, et conformément à l'article 261 du TFUE, un recours est garanti auprès de la Cour de Justice de l'Union européenne (art. 81 DSA).

Le règlement européen impose à chaque État membre de désigner une autorité compétente pour qu'elle agisse en tant que « coordinateur pour les services numériques » (art. 49). Ce coordinateur est responsable « de toutes les questions en lien avec la surveillance et l'exécution » du DSA. Il assure la coordination au niveau national des questions liées au DSA et dispose d'un certain pouvoir à l'égard des intermédiaires en ligne (article 51 DSA).

Enfin, le DSA institue un « Comité européen des services numériques » chargés de la surveillance « des fournisseurs de services intermédiaires » (art. 61). Ce comité assure aussi une fonction de conseil aux coordinateurs de services numériques des États membres.

5.5. Royaume Uni : Online Safety Act

5.5.1. Contexte

Les discussions concernant l'adoption d'un régime réglementaire propre aux entreprises de réseaux sociaux ont débuté en mai 2018, lorsque le gouvernement britannique a annoncé son intention de publier un *white paper* contenant ses projets de législation en matière de sécurité en ligne¹⁹⁹. La publication de ce papier est survenue en avril 2019 sous le titre d'*Online Harms*

¹⁹⁹ Government response to the Internet Safety Strategy Green Paper, May 2018. Disponible sous : https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/708873/Government_Response_to_the_Internet_Safety_Strategy_Green_Paper_-_Final.pdf

*White Paper*²⁰⁰. Dans ce *white paper*, le gouvernement identifie la prolifération des contenus « illégaux et inacceptables » en ligne comme un problème important au Royaume-Uni. Cette prolifération fait courir des risques tant à la sécurité nationale (en raison de la propagande terroriste) qu'à la protection des enfants (en raison de l'usage d'internet par les délinquants sexuels), et peut concourir à saper les valeurs démocratiques (en raison de la prolifération de la désinformation et des abus anonymes en ligne qui conduisent les utilisatrices et utilisateurs à se restreindre dans ce qu'ils vont exprimer en ligne, phénomène décrit sous le nom de « *chilling effect* »)²⁰¹. En réponse à ces problèmes, le gouvernement rappelle que des réponses réglementaires existent déjà, mais que celles-ci ne vont pas assez loin. Ainsi, avec le *white paper*, le gouvernement britannique énonce son intention d'établir un cadre réglementaire général pour lutter contre les préjudices en ligne (« *online harms* »).

Le *Online Harms White Paper* a donné lieu à une consultation qui s'est terminée en juillet 2019²⁰². En réponse à cette consultation, le gouvernement britannique a annoncé en décembre 2020 qu'un *Online Safety Bill* serait introduit²⁰³. En conséquence, un premier draft de l'*Online Safety Bill* a été publié en mai 2021²⁰⁴, à propos duquel un comité mixte composé de membres de la House of Commons et de la House of Lords ont rédigé un rapport contenant des propositions d'améliorations. Ce rapport, publié le 14 décembre 2021²⁰⁵, rappelle en introduction que l'approche d'auto-régulation pour les intermédiaires Internet s'est avérée être un échec (« *the self-regulation of online services has failed* »). Selon le rapport, les plus grandes entreprises de réseaux sociaux ont pendant trop longtemps pu se présenter en tant qu'acteurs neutres, qui ne seraient pas responsables pour les contenus qui sont créés et partagés par les utilisatrices et utilisateurs. En conséquence, le rapport soutient la mise en place d'une loi qui rendrait les plus grandes entreprises responsables du design et de la mise en œuvre de leurs produits et services, notamment lorsqu'ils sont basés sur des fonctionnements algorithmiques. Le rapport appelle également à la mise à jour du code pénal britannique pour qu'il inclue de nouvelles infractions « pour les communications préjudiciables, menaçantes et sciemment fausses », pour le « cyber-flashing » (l'envoi de photos non consenties à des tiers, principalement à caractère sexuel), pour « la promotion de l'auto-mutilation » et pour les

²⁰⁰ Government Online Harms White Paper, April 2019. Disponible sous : https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/973939/Online_Harms_White_Paper_V2.pdf

²⁰¹ *Ibid.*, p. 5.

²⁰² Department for Digital, Culture, Media & Sport and Home Office, Consultation on the Online Harms White Paper, April 2019. Disponible sous : <https://www.gov.uk/government/consultations/online-harms-white-paper>

²⁰³ Department for Digital, Culture, Media & Sport and Home Office, Online Harms White Paper: Full government response to the consultation, December 2020. Disponible sous : <https://www.gov.uk/government/consultations/online-harms-white-paper/outcome/online-harms-white-paper-full-government-response>

²⁰⁴ Draft Online Safety Bill, May 2021. Disponible sous : https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/985033/Draft_Online_Safety_Bill_Bookmarked.pdf

²⁰⁵ Joint Committee on the Draft Online Safety Bill – Report of Session 2021-22, December 2021. Disponible sous : <https://committees.parliament.uk/publications/8206/documents/84092/default/>

« discours de haine fondés sur le sexe, le genre ou le handicap »²⁰⁶. Les entreprises de réseaux sociaux devraient en conséquence avoir une obligation d'atténuer les risques de réalisation de ces infractions en ligne, et être tenus pour responsables dans les cas où ils échoueraient à atténuer ces risques. Cette obligation doit être détaillée dans des codes de conduite à caractère contraignant pour les entreprises. Le gouvernement a rendu un document de réponse au rapport du comité mixte le 17 mars 2022²⁰⁷. Celui-ci détaille comment le gouvernement intègre ses commentaires. Principalement, le rapport explique qu'il intègre les recommandations du comité mixte en matière de clarté des obligations qui s'appliquent aux fournisseurs de services, mais ne souhaite pas déléguer des pouvoirs plus étendus au régulateur que proposé dans la première version du draft. Le rapport indique de plus ne pas souhaiter suivre la recommandation du comité de rendre les codes de conduite contraignants, souhaitant rester dans une approche basée sur la promotion de l'innovation²⁰⁸.

La procédure législative du *Online Safety Bill* s'est déroulée en plusieurs phases, sur deux années parlementaires²⁰⁹. La première phase a débuté le 17 mars 2022, durant la session parlementaire 2021-2022, avec une première lecture du texte²¹⁰ devant la *House of Commons*, suivi d'une seconde lecture²¹¹ toujours devant la *House of Commons* le 19 avril 2022. Le 19 avril 2022, la *House of Commons* a convenu d'une « *carry-over motion* », motion permettant de transférer l'examen d'un texte de loi sur la session parlementaire suivante²¹². La seconde phase s'est déroulée durant la session parlementaire 2022-2023. Le texte a été lu trois fois devant la *House of Commons*, les premières et deuxième lectures se déroulant le 11 mai 2022 et la troisième lecture le 17 janvier 2023. C'est à la fin de cette troisième lecture que le texte a été adopté par la *House of Commons* et transmis à la *House of Lords*. Trois lectures ont eu lieu devant la *House of Lords*, les 18 janvier²¹³, 1^{er} février²¹⁴ et 6 septembre²¹⁵ 2023. Le projet de loi

²⁰⁶ Joint Committee on the Draft Online Safety Bill – Report of Session 2021-22, *idem* n206, p. 3.

²⁰⁷ Government Response to the Report of the Joint Committee on the Draft Online Safety Bill, March 2022.

Disponible sous :

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1061446/E0272_1600_Gov_Resp_to_Online_Safety_Bill_Accessible_v1.0.pdf

²⁰⁸ *Ibid.*, p. 12.

²⁰⁹ Online Safety Act 2023 – Parliamentary Bills. Disponible sous :

<https://bills.parliament.uk/bills/3137/publications>

²¹⁰ House of Commons, Online Safety Act, 1st reading. Disponible sous :

<https://bills.parliament.uk/bills/3137/stages/16299>

²¹¹ House of Commons, Online Safety Act, 2nd reading. Disponible sous :

<https://bills.parliament.uk/bills/3137/stages/16300>

²¹² UK Parliament, Online Safety Act 2023, stages. Disponible sous :

<https://bills.parliament.uk/bills/3137/stages/16385>

²¹³ House of Lords, Online Safety Act, 1st reading. Disponible sous :

<https://bills.parliament.uk/bills/3137/stages/17303>

²¹⁴ House of Lords, Online Safety Act, 2nd reading. Disponible sous :

<https://bills.parliament.uk/bills/3137/stages/17304>

²¹⁵ House of Lords, Online Safety Act, 3rd reading. Disponible sous :

<https://bills.parliament.uk/bills/3137/stages/17870>

a finalement reçu l'assentiment royal le 26 octobre 2023²¹⁶ et a ainsi été formellement adopté au terme d'un long processus²¹⁷. Le texte de loi est entré en vigueur le 10 janvier 2024²¹⁸.

5.5.2. Champ d'application

L'*Online Safety Act* vise l'activité des entreprises de réseaux sociaux, qu'il comprend dans ce qu'il nomme les « services d'utilisateurs à utilisateurs ». Ces « services d'utilisateurs à utilisateurs » sont définis comme des « services internet qui permettent aux utilisateurs de créer ou de téléchargement directement leurs propres contenus via le service, et où d'autres utilisateurs peuvent voir ces contenus » (Article 3).

Le législateur britannique a en outre indiqué que les « services d'utilisateurs à utilisateurs » qui sont soumis à la loi doivent être liés au Royaume-Uni. Ce lien existe lorsque le service a un nombre significatif d'utilisatrices et d'utilisateurs, lorsque le Royaume-Uni est un des marchés de destination du service, ou lorsque le service peut être utilisé par des individus situés au Royaume-Uni et qu'un risque de préjudice existe en raison des contenus postés sur le service (Article 3, Sous-sections 2, 5 et 6).

Comme les législations précédentes, l'*Online Safety Act* ne fait pas référence à la notion de discours de haine. Il fait référence aux « contenus et activités illégaux » (« illegal content and activity »). Toutefois, à la différence des autres législations nationales, il n'énumère pas les dispositions en vigueur dans le droit pénal ou dans des législations spéciales. Il opte pour un renvoi général indiquant qu'est illicite ce qui constitue une infraction selon le droit britannique (art. 59). Faisant écho aux règles du droit pénal international, il différencie les infractions dites prioritaires, qui comprennent les infractions en matière de terrorisme et de pédopornographie, et les autres infractions. En outre, le législateur britannique vise aussi les contenus utiles pour la démocratie (contenu d'importance démocratique et « contenu provenant d'éditeurs de presse »).

5.5.3. Règles organisationnelles relatives à l'activité des entreprises de réseaux sociaux

L'*Online Safety Act* introduit une obligation générale de diligence pour les entreprises actives dans les services en ligne. Cette obligation de diligence comprend l'obligation de garantir la sûreté en ligne des utilisatrices et utilisateurs et l'obligation de s'attaquer aux contenus préjudiciables publiés sur leurs services²¹⁹. Cela implique de prendre des mesures de mitigation et de gestion des risques de préjudices (« risks of online harms ») qui pourraient résulter des « contenus et activités illégales » ainsi que des « contenus et activités qui peuvent être préjudiciables aux enfants » (*Online Safety Act, introduction*). L'analyse de risques se décline sur plusieurs niveaux. Elle implique d'évaluer le risque en matière de publication de contenus

²¹⁶ Royal Assent Online Safety Act. Disponible sous : <https://bills.parliament.uk/bills/3137/stages/17955>

²¹⁷ Neudert, L.-M. (2023). Regulatory capacity capture: The United Kingdom's online safety regime. *Internet Policy Review*, 12(4). <https://doi.org/10.14763/2023.4.1730>

²¹⁸ UK Statutory Instruments, Online Safety Act, Commencement n°3. Disponible sous : [The Online Safety Act 2023 \(Commencement No. 3\) Regulations 2024 \(legislation.gov.uk\)](https://www.legislation.gov.uk/ukstatutoryinstruments/2023/12/1/online-safety-act-commencement-no-3/regulations-2024)

²¹⁹ *Ibid.*, p. 42.

illégaux (art.9). Les entreprises doivent également évaluer la capacité de leurs propres services à contribuer à la propagation des contenus illégaux. L'analyse de risque doit également intégrer une évaluation du risque de préjudice que la présence de contenus illégaux peut causer aux utilisateurs. Face à ce risque de préjudice, les entreprises de réseaux sociaux doivent prendre des mesures techniques et internes pour prévenir que les utilisateurs ne soient exposés à des contenus illégaux lors de l'utilisation de leurs services. Ces mesures concernent les fonctionnalités et les algorithmes de la plateforme et les conditions d'utilisation.

Aussi, l'article 17 introduit une obligation pour les entreprises de protéger plus particulièrement les contenus « d'importance démocratique ». Cette obligation implique pour les entreprises de déterminer dans leurs conditions d'utilisation quels sont les contenus qui vont bénéficier de cette protection spéciale, qu'il s'agisse de contenu provenant d'éditeurs de presse mais également de simples utilisateurs. Le contenu provenant d'éditeurs de presse bénéficie également d'une protection générale introduite par l'article 18 (et donc pas seulement lorsque ce contenu consiste également en un contenu d'importance démocratique).

Le législateur britannique impose aux entreprises de rendre public, une fois par année, un rapport de transparence. Ce rapport doit détailler les mesures prises pour respecter les différentes obligations introduites par la loi (art 77).

5.5.4. Règles de procédure applicables aux rapports entre entreprises et utilisateurs

Comme dans les cas précédemment examinés, l'*Online Safety Act* impose une obligation pour les entreprises de réseaux sociaux de mettre en place une procédure de signalement pour les contenus illégaux, mais également pour les contenus non illégaux qui sont susceptibles de porter préjudice à des enfants.

Cette procédure peut être utilisée par les utilisateurs et les « personnes affectées », c'est-à-dire une personne qui, sans être utilisateur d'un service ou d'une plateforme, est concernée par un contenu publié en ligne car elle en est le sujet, ou que son enfant mineur en est le sujet.

Les utilisateurs qui ont vu un de leur contenu restreint ou retiré, ou dont le compte a été suspendu, doivent pouvoir formuler une plainte à l'encontre de la décision de l'entreprise (art. 21). L'entreprise doit réexaminer sa décision en conséquence.

L'article 22 rappelle que lorsque les entreprises de réseaux sociaux prennent des mesures à l'encontre du contenu des utilisateurs, ceux-ci doivent tenir compte de leur droit à la liberté d'expression.

5.5.5. Autorités spéciales

Le législateur britannique charge l'office des communications britannique (« Office of communications », OFCOM) de mettre en œuvre et de contrôler la bonne application de la

législation (article 91)²²⁰. Il est également chargé de réaliser des codes de conduite formulant des recommandations sur la procédure de traitement des contenus illégaux par les intermédiaires (article 41).

5.6. Synthèse : mise en œuvre et tableau récapitulatif

Au terme de cette analyse et bien que la question de l'évaluation de la légalité des différentes réglementations et de leur évolution n'entre pas directement dans le champ de cette étude, il faut noter que les réglementations allemandes, françaises et autrichiennes ont été sujettes à évolutions et critiques. Il convient de les évoquer succinctement avant de présenter un tableau récapitulatif de l'ensemble de l'analyse.

Ainsi, la législation allemande a notamment été critiquée par le rapporteur spécial, celui-ci estimant qu'elle risquait de conduire à des retraits excessifs de contenus de la part des entreprises de réseaux sociaux qui préféreraient retirer préventivement des contenus que de voir leur responsabilité engagée²²¹. De plus, trois ans après l'adoption de la loi, le gouvernement fédéral a proposé, le 1^{er} avril 2020, une adaptation de la loi, sous le titre « NetzDGÄnG »²²². Cette nouvelle loi est entrée en vigueur le 28 juin 2021 et a renforcé les droits des utilisateurs sur les réseaux sociaux, qui bénéficient de la possibilité de donner leur avis lorsqu'ils sont en désaccord avec la décision d'une entreprise de retirer leur contenu. En outre, la mise en œuvre de la NetzDG a conduit au développement d'une jurisprudence relative à la responsabilité des entreprises de réseaux sociaux. Le *Bundesgerichtshof* a notamment jugé que les entreprises de réseaux sociaux avaient le droit de fixer, dans leurs conditions générales, des règles qui vont au-delà de ce que prescrit le droit pénal allemand²²³. Dans le cas français, bien que partiellement censurée par le Conseil constitutionnel, la loi Avia a été mise en œuvre sur le plan institutionnel avec un décalage de plusieurs années. Ainsi, l'autorité chargée par la loi d'enquêter sur « les délits de harcèlement sexuel ou moral (...), dès lors que les faits sont commis par l'utilisation d'un service de communication au public en ligne ou par le biais d'un support numérique » a été mise sur pied en juin 2021²²⁴. Elle a pris forme avec la création d'un « pôle national de lutte contre la haine en ligne » au sein du parquet

²²⁰ L'OFCOM, au Royaume-Uni, est l'entité chargée de la régulation du secteur des communications. Elle a été instituée par le Communication Act de 2003 (disponible sous : [The Online Safety Act 2023 \(Commencement No. 3\) Regulations 2024 \(legislation.gov.uk\)](https://legislation.gov.uk/ukpga/2003/3/section/3)).

²²¹ Rapport A/74/486, *idem* n° 34, p. 15, §32.

²²² Disponible sous : [NetzDGÄnG Gesetz zur Änderung des Netzwerkdurchsetzungsgesetzes \(buzer.de\)](https://www.buzer.de/NetzdG)

²²³ Parmi les nombreux arrêts rendus par la Cour fédérale, on peut notamment citer : « BGH, 29 juillet 2021, III ZR 179/20, disponible sous [BGH, 29.07.2021 - III ZR 179/20 - dejure.org](https://www.dejure.org/lit/j/bgh/2021/179_20) ; BGH, 29 juillet 2021, III ZR 192/20, disponible sous [BGH, 29.07.2021 - III ZR 192/20 - dejure.org](https://www.dejure.org/lit/j/bgh/2021/192_20), ou de la Cour constitutionnelle (« Bundesverfassungsgericht ») qui traite de la relation d'égalité entre privés et des conséquences de suspension de comptes lors de périodes électorales (BVerfG, 22 mai 2019, I BvQ 42/19. Disponible sous : [BVerfG, Beschluss der 2. Kammer des Ersten Senats vom 22. Mai 2019 - I BvQ 42/19 - Rn. \(I - 25\), http://www.bverfg.de/e/qlk20190522_1bvq004219.html \(bundesverfassungsgericht.de\)](https://www.bverfg.de/e/qlk20190522_1bvq004219.html))

²²⁴ Ministre de la justice, Circulaire relative à la lutte contre la haine en ligne, 24 novembre 2020. Disponible sous : [JUSD2032620C.pdf \(justice.gouv.fr\)](https://www.justice.gouv.fr/IMG/pdf/JUSD2032620C.pdf).

de Paris ²²⁵. Ce pôle a pour mission de centraliser le traitement des affaires complexes et « apporte son expertise à l'ensemble des juridictions, (pouvant) se saisir de toute affaire relevant de la lutte contre la haine en ligne »²²⁶. Le Conseil supérieur de l'audiovisuel (CSA), en tant qu'autorité de surveillance, a quant à lui disparu à l'occasion de la création de l'ARCOM, auquel il a été intégré²²⁷. Cette autorité accueille l'« observatoire de la haine en ligne » créé sur la base de la loi Avia depuis juillet 2020, dont la mission est « d'analyser et quantifier les contenus relatifs à la haine en ligne » et d'œuvrer à « améliorer la compréhension du phénomène », en servant d'intermédiaire entre les différents acteurs publics et privés concernés²²⁸. Quant à la législation autrichienne, elle a été remise en cause à la suite d'un recours introduit par Meta, Tiktok et Google auprès de la plus haute juridiction administrative contre la décision de *KommAustria* les désignant en tant qu'entreprises soumises aux obligations de la Ko-PI. Cette procédure a donné lieu à une question préjudicielle auprès de la Cour de Justice de l'UE (CJUE) du 10 juin 2022²²⁹, concernant la compatibilité de la Ko-PI avec la directive e-commerce encore pleinement en vigueur à ce moment²³⁰. Dans son arrêt rendu le 9 novembre 2023²³¹, la CJUE a rappelé que sous le régime de la directive e-commerce, les Etats membres ne pouvaient pas adopter des mesures à caractère **général et abstrait s'imposant** à des « services de la société de l'information » (les entreprises de réseaux sociaux notamment) si ces services ne sont pas domiciliés dans leur pays. Meta, Tiktok et Google n'étant pas domiciliés en Autriche mais bien en Irlande, l'Autriche ne peut en principe donc pas les soumettre à des mesures car celles-ci risquent de mettre en danger le bon fonctionnement du marché unique. L'Autriche, n'ayant pas rempli les conditions permettant une exception à ce principe (pour déroger au principe selon lequel seul l'état dans lequel les services sont domiciliés peut adopter des mesures s'imposant aux entreprises de réseaux sociaux, il faut viser un service spécifique et concret, et justifier des mesures qui visent à garantir « l'ordre public, la protection de la santé publique, la sécurité publique ou la

²²⁵ Légipresse. (2020). Mise en œuvre du « Parquet numérique » prévu par la loi. Edition du 27 novembre 2020. Disponible sous : [Mise en œuvre du « Parquet numérique » prévu par la loi Avia - Communications électroniques \(legipresse.com\)](#) ; également Ministère de la justice, Lutte contre la haine en ligne, 28 juin 2021. Disponible sous : [Lutte contre la haine en ligne | Ministère de la justice](#)

²²⁶ Gouvernement français, Délégation interministérielle à la lutte contre le racisme, l'antisémitisme et la haine anti-LGBT. Disponible sous : [Lutter contre la haine en ligne | DILCRAH](#)

²²⁷ L'ARCOM est née de la fusion du Conseil supérieur de l'audiovisuel (CSA) et de la Haute Autorité pour la diffusion des œuvres et la protection des droits sur Internet (Hadopi), sur la base du décret n°2021-1853 du 27 décembre 2021 (disponible sous [Décret n° 2021-1853 du 27 décembre 2021 relatif aux compétences et à l'organisation de l'Autorité de régulation de la communication audiovisuelle et numérique en matière de protection du droit d'auteur et des droits voisins - Légifrance \(legifrance.gouv.fr\)](#)). Ses prérogatives incluent la régulation technique et économique du marché audiovisuel mais également la protection des droits d'auteurs.

²²⁸ ARCOM, Observatoire de la haine en ligne : analyser pour mieux lutter. Disponible sous : [Observatoire de la haine en ligne : analyser pour mieux lutter | Arcom](#).

²²⁹ CJUE, Affaire C-376-22, Demande de décision préjudicielle. Disponible sous : [CURIA - Fiche détaillée \(europa.eu\)](#)

²³⁰ Certains articles de la directive e-commerce ont cessé de faire effet avec l'entrée en vigueur du DSA ; toutefois, au moment de la question préjudicielle auprès de la CJUE, le DSA n'avait pas encore été adopté et n'était pas encore en vigueur.

²³¹ CJUE, Affaire C-376-22, Arrêt. Disponible sous : [CURIA - Fiche détaillée \(europa.eu\)](#)

protection des consommateurs »), n'était ainsi pas habilité à prendre de telles mesures. La CJUE a donc mis en évidence le défaut de conformité de la législation autrichienne au motif que celle-ci « empiète sur la compétence de l'Etat membre d'origine » des entreprises.

L'entrée en vigueur du DSA le 16 novembre 2022 a aussi eu une incidence directe sur les législations nationales. Ainsi, les autorités allemandes ont annoncé l'abrogation de la NetzDG, considérant que le DSA vise à atteindre les mêmes objectifs que ceux poursuivis par la NetzDG²³². En plus du DSA, qui est un règlement de l'Union Européenne, les Etats membres doivent adopter une loi nationale qui règle l'application de la loi et qui désigne notamment l'entité à l'échelle nationale chargée de jouer le rôle de coordinateur pour les services numériques. En principe, ces lois nationales doivent entrer en vigueur parallèlement à l'entrée en vigueur du DSA, c'est-à-dire au 17 février 2024. Toutefois, le gouvernement allemand a annoncé qu'il avait pris du retard sur ce processus et qu'il n'arrivera pas à être prêt pour le 17 février 2024²³³. La NetzDG reste ainsi en vigueur tant que la loi désignant le coordinateur pour les services numériques n'aura pas été adoptée. Du côté autrichien, et eu égard aussi à l'arrêt de la CJUE, le gouvernement autrichien a annoncé que la question de la conformité de la Ko-PI était rendue caduque par l'entrée en vigueur du DSA²³⁴. En effet, la loi d'application du DSA – « *Koordinator-für-digitale-Dienste-Gesetz* » (KDD-G)²³⁵ – a été adoptée et est entrée en vigueur le 17 février 2024. Elle institue le coordinateur national des services numériques conformément aux dispositions du DSA et prévoit l'abrogation de la Ko-PI à dater de l'entrée en vigueur de la KDD-G²³⁶.

Enfin, à des fins de synthèse et de comparaison, nous résumons notre analyse dans le tableau ci-dessous.

²³² Deutscher Bundestag, Antwort der Bundesregierung, Digital Services Act, Drucksache 20/2308. 17 juin 2022. Disponible sous : [Drucksache 20/2308 \(bundestag.de\)](https://www.bundestag.de/Drucksache20/2308)

²³³ Euronews, Platform rules to apply this month as member states lag on oversight personnel, 7 février 2024. Disponible sous : [Platform rules to apply this month as member states lag on oversight personnel | Euronews](https://www.euronews.com/fr/actualites/digital/2024/02/07/platform-rules-to-apply-this-month-as-member-states-lag-on-oversight-personnel)

²³⁴ Koordinator-für-digitale-Dienste-Gesetz, §10 alinéa 2. Disponible sous : [BGBl_2023_I_182.pdfsig \(bka.gv.at\)](https://www.bka.gv.at/BGBl_2023_I_182.pdf).

²³⁵ Disponible sous : [BGBl_2023_I_182.pdfsig \(bka.gv.at\)](https://www.bka.gv.at/BGBl_2023_I_182.pdf)

²³⁶ KommAustria, Kommunikationsplattformen, News. Disponible sous : [Kommunikationsplattformen | RTR](https://komm.austria.at/kommunikationsplattformen)

		Allemagne	France	Autriche	Union Européenne	Royaume-Uni
Champ d'application	<i>Sur le plan personnel</i>	« Plateformes internet qui permettent, en vue d'obtenir un profit, l'échange de contenus entre utilisateurs ou envers le public »	« Opérateurs de plateforme en ligne qui proposent un service de communication au public en ligne reposant sur la mise en relation de plusieurs parties en vue du partage de contenus publics »	« Fournisseurs de services nationaux et étrangers qui proposent des plateformes de communication dans un but lucratif »	« Fournisseurs de services intermédiaires proposés aux destinataires du service dont le lieu d'établissement est situé dans l'Union ou qui sont situés dans l'Union, quel que soit le lieu d'établissement des fournisseurs de ces services intermédiaires »	« Services internet qui permettent aux utilisateurs de créer ou de téléchargement directement leurs propres contenus via le service, et où d'autres utilisateurs peuvent voir ces contenus »
	<i>Sur le plan matériel</i>	Contenus « contraires au droit »	Contenus qui correspondent à certaines dispositions de lois	Contenus « contraires au droit »	« Contenus et activités illicites »	« Contenus et activités illicites »
		Renvoi direct aux dispositions du code pénal	Renvoi direct aux dispositions de la loi sur la liberté de la presse et code pénal français	Renvoi direct à des dispositions du code pénal	Renvoi indirect au droit national des États membre et au droit de l'Union	Renvoi indirect au droit pénal britannique
Mesures organisationnelles dans le rapport entre autorités et entreprises	<i>Point de contact</i>	Mandataire	Mandataire	Mandataire	Représentant légal et point de contact pour les utilisateurs	-
	<i>Rapport annuel</i>	Publication de rapport sur la mise en œuvre des obligations	Publication de rapport sur la mise en œuvre des obligations	Publication de rapport sur la mise en œuvre des obligations	Publication de rapport sur la mise en œuvre des obligations	Publication de rapport sur la mise en œuvre des obligations

	Accessibilité des conditions d'utilisation	-	Formuler les conditions générales « en termes précis, aisément compréhensibles, objectifs et non discriminatoires »	-	Expliquer les restrictions imposées aux utilisateurs dans les conditions générales / Usage d'un langage clair et accessible	Inclure des dispositions « claires et accessibles » informant les utilisateurs de leur droit de déposer une plainte pour rupture de contrat si leur contenu est retiré en violation des conditions générales
	Accès aux données	-	-	-	Pour les coordinateurs de services numériques des Etats membres et pour les chercheurs agréés, pour mener des études sur les risques systémiques	-
	Analyse des risques et audits externes	-	-	-	Analyse (interne) de risque et audits (externes) pour les très grandes plateformes	Analyse de risque en matière de contenus illégaux
Règles de procédure dans le rapport entre entreprises et utilisateurs	Procédure	Procédure de signalement des contenus pour les utilisateurs Obligation d'expliquer les motifs de retrait et d'informer de la possibilité de soumettre une « contre-proposition »	Procédure de notification pour « toute personne » Obligation d'informer les utilisateurs de la voie de recours interne et des motifs de la décision de retrait	Procédure de signalement des contenus pour les utilisateurs Obligation d'informer les utilisateurs dont les contenus ont été retirés « des voies de recours dont ils disposent » et des motifs de la décision	Procédure de « notification et action » pour que « tout particulier ou toute entité » puisse signaler des contenus ou activités illégales Obligation de fournir un « exposé des motifs » dans les cas où une restriction est imposée	Procédure de plainte pour les utilisateurs qui constatent des contenus illégaux ou qui se sont vu retirer leurs contenus ainsi que pour « les personnes affectées »

	<i>Recours interne</i>	Possibilité de soumettre des « contre-propositions » pour réexamen de la décision	Possibilité de recours interne contre les retraits de contenus	Possibilité de soumettre des « contre-propositions » pour réexamen de la décision	Mise en place d'un « système interne de traitement des réclamations » permettant aux utilisateurs de demander aux entreprises de vérifier les décisions de retraits une seconde fois	-
	<i>Recours externe</i>	-	-	-	Mise en place d'un « système extrajudiciaire de résolution des litiges »	-
Autorités spéciales	<i>Fonction conseil</i>	Entités « d'autorégulation réglementée » pouvant aider la prise de décision quant au caractère illicite	-	-	« Signaleurs de confiance », qui bénéficient d'un statut privilégié pour signaler des contenus	L'office des communications peut établir des codes de conduite
	<i>Fonction d'autorité de recours contre les décisions des entreprises</i>	Organes de médiation pour médier d'éventuelles procédures entre les utilisateurs et les entreprises	-	Création d'un bureau des plaintes auprès de la Société autrichienne de régulation de la radio et des télécommunications	Des entités externes peuvent être certifiées en tant qu'organes de règlement extrajudiciaire des litiges	À déterminer par l'office des communications ultérieurement

	<i>Fonction de surveillance</i>	Ministère fédérale de la Justice contrôle l'application de la loi Le Bureau fédéral de la police criminelle doit être informé en vue de la poursuite pénale	Le Conseil supérieur de l'audiovisuel contrôle l'application de la loi (<i>remplacé ultérieurement par l'ARCOM</i>) Création d'un « observatoire de la haine en ligne » Création d'une juridiction spécialisée	L'autorité de communication contrôle l'application de la loi	La Commission Européenne surveille l'application de la loi en commun avec les Etats membres Les autorités administratives et judiciaires des États membres peuvent adresser directement des injonctions d'agir Chaque État membre désigne un « coordinateur pour les services numériques »	L'office des communications contrôle l'application de la loi
--	---------------------------------	---	--	---	---	---

6. Conclusions et remarques prospectives

Au terme de cette étude sur les modèles réglementaires de lutte contre les discours de haine en ligne, deux éléments méritent d'être mis en exergue concernant la notion de discours de haine et les différents modèles réglementaires. Ils nous permettront alors de formuler quelques remarques prospectives.

Le premier concerne la notion de discours de haine. La définition de cette notion est longtemps restée incertaine²³⁷. En effet, pour des systèmes politiques fondés sur la valeur de la liberté comme les démocraties libérales, la définition et la juridicisation du discours de haine s'avère particulièrement délicate pour celles et ceux qui veulent restreindre autant que possible la censure. Dans le contexte de cette étude sur les modèles réglementaires de la lutte contre le discours de haine, cette notion nécessitait clarification. En l'occurrence, l'étude des travaux des organes conventionnels de droit international a permis de montrer que la notion peut être rattachée à deux dispositions internationales : l'article 20 du Pacte sur les droits civils et politiques et l'article 4 de la Convention sur l'élimination de la discrimination raciale. Du point de vue du droit onusien, elle est ainsi intimement liée à la question de la violence – atteinte physique ou psychique – à l'égard des personnes et à la question de la discrimination raciale, religieuse et nationale alors que le Conseil de l'Europe dispose d'une conception plus étendue de ce qui relève des propos haineux. En outre, ayant opté pour une lecture dynamique du droit international, l'étude des rapports produits par les rapporteurs spéciaux a permis de montrer que la notion de discours de haine en ligne s'inscrit dans un cadre plus large de contenus jugés problématiques. Il y a d'une part les propos qui sont susceptibles de conduire, inciter ou encore contribuer à la discrimination pour les motifs précédemment évoqués, mais aussi les contenus qui sont illicites au motif qu'ils constituent des infractions du point de vue du droit pénal international : la pédopornographie, les activités de terrorisme et celles liées au génocide. Outre ces quatre infractions du point de vue du droit pénal international, l'étude des législations nationales qui règlent la question de la responsabilité des entreprises en ligne montrent que d'autres contenus peuvent encore relever d'infractions en fonction de sensibilités nationales. Enfin, les rapports des rapporteurs spéciaux ont permis de mettre en évidence que les services offerts par les entreprises en ligne conduisent encore à un troisième type de propos problématiques, à savoir ceux qui sans relever du champ pénal sont pourtant blessants ou peuvent poser une forme de préjudice (« *harmful content* »). Cet élargissement de la perspective montre ainsi que le discours de haine est une notion aux contours juridiques encore mouvants, mais qui est en tout susceptible de couvrir des propos qui vont au-delà d'une vision strictement ancrée dans les dispositions conventionnelles onusiennes.

Le deuxième enseignement concerne les modèles réglementaires qui ressortent de cette étude. En l'occurrence, les perspectives diachroniques et synchroniques appliquées dans le cadre de cette étude ont permis de mettre au jour une double distinction. La première distinction oppose le modèle initial (américain) au modèle contemporain (européen). Le modèle initial

²³⁷ Gelber, K. (2019). Differentiating hate speech: a systemic discrimination approach. *Critical Review of International Social and Political Philosophy*, 24(4). <https://doi.org/10.1080/13698230.2019.1576006>

puise sa source dans l’approche américaine qui a été adoptée lors des balbutiements de l’Internet accessible à tous avec l’idée que l’Etat (le régulateur) exonérait les entreprises (les cibles de la régulation) de (quasiment) toute responsabilité pour les contenus diffusés par le biais de leurs services. Ce modèle est celui qui a largement inspiré la réponse européenne dans le cadre de la directive e-commerce. Par la suite, le modèle contemporain qui s’est développé à travers les différentes législations européennes dépasse la logique classique du régulateur/cible et fixe deux types d’obligations à destination des entreprises de réseaux sociaux. Le premier type d’obligations reste ancré dans la logique régulateur/cible et concerne l’activité de ces entreprises actives sur le territoire national ou européen. Celles-ci doivent notamment avoir un point de contact et publier un rapport de transparence annuel. Le second type d’obligations impose aux entreprises de jouer le rôle d’intermédiaires entre les usagers. Elles sont dorénavant contraintes d’exécuter leur activité de modération en ligne en respectant une série de garanties de procédure. Cette évolution de la réglementation fait directement écho au contenu des rapports des rapporteurs spéciaux qui appellent progressivement à un changement de perspective tant dans l’implication des Etats dans la lutte contre les contenus problématiques que dans celui des entreprises en ligne que les Etats doivent utiliser comme intermédiaires de la régulation.

	<i>Modèle réglementaire initial (d’origine américaine)</i>	<i>Modèle réglementaire contemporain (d’origine européenne)</i>
À partir des années 1990	Section 230 (USA, 1996) / Directive e-commerce (UE, 2001)	
À partir des années 2017		Version 1.0 : NetzDG (Allemagne, 2017) Loi Avia (France, 2020) Ko-PI (Autriche, 2021)
		Version 2.0 DSA (UE, 2024) Online Safety Bill (UK, 2024)

L’analyse des réponses réglementaires adoptées dans le cadre du modèle contemporain européen permet d’opérer une seconde distinction entre ce que l’on peut désigner comme les réglementations 1.0, qui regroupent les trois législations nationales adoptées avant la réglementation européenne, et les réglementations 2.0, qui incluent la réglementation européenne et la réglementation britannique. Entre ces deux générations de loi, on retrouve des points de continuité en ce sens que toutes exigent que les entreprises de réseaux sociaux jouent le rôle d’intermédiaire dans l’exercice de la liberté en ligne des utilisateurs. Les différentes lois requièrent le respect de garanties de procédure (obligation de motivation des décisions, recours interne, recours externe), lesquelles font ainsi écho à l’approche des droits fondamentaux promues dans les rapports des rapporteurs spéciaux. Les deux générations

dépassent aussi une logique de gouvernance entre l'Etat, l'intermédiaire et les utilisateurs. Sous réserve du cas français, elles font intervenir d'autres acteurs qui exercent des fonctions de conseil, soutien ou instances de recours à l'amiable. En termes de différence, les réglementations 2.0 élargissent substantiellement le champ des obligations à charge des entreprises en ligne. Elles ne doivent plus seulement disposer d'un point de contact, rédiger un rapport annuel ou disposer de conditions d'utilisations accessibles pour les utilisateurs, mais doivent aussi prendre des mesures d'évaluation des risques et se soumettre à des audits externes. Ces réglementations développent ainsi une conception plus systémique des obligations à charge des entreprises de réseaux sociaux et de la façon dont celles-ci modèrent les contenus, s'éloignant ainsi de l'approche basée avant tout sur le traitement de cas individuels comme on le retrouve dans les réglementations 1.0.

Enfin, force est de constater que les différentes réglementations européennes ont redessiné le régime initial de responsabilité des entreprises de réseaux sociaux. Premièrement, celles-ci sont tenues de mettre en conformité leur politique de modération de contenu avec les normes juridiques européennes. Deuxièmement, elles doivent mettre en place des mesures organisationnelles et de gestion des risques. Troisièmement, le législateur européen a instauré une gouvernance européenne constituée d'autorités publiques (nationales et européenne) et d'acteurs privés reconnus par les autorités publiques, exigeant ainsi des Etats de mettre en place les autorités pertinentes. Du point de vue du discours de haine, ce nouveau dispositif permet de mettre en place une gouvernance permettant de surveiller, de lutter et de réprimer la propagation de contenus relevant de ce type de discours.

Pour terminer, il faut dire qu'il est bien entendu trop tôt pour évaluer la pertinence du modèle européen 2.0 en matière d'encadrement des activités en ligne et de son impact sur la propagation des discours de haine. Toutefois, les Etats non-membres de l'Union européenne ne devraient pas utiliser un tel argument pour justifier une attitude attentiste. D'une part, les atteintes causées par les réseaux sociaux à la démocratie et aux individus ont été largement documentées depuis maintenant plusieurs années et montrent que ce problème connaît une évolution très inquiétante, au point de mettre en danger le système démocratique sur la base duquel les Etats fonctionnent. Une telle situation exige donc d'engager une action étatique. D'autre part, l'expérience européenne offre un cadre réglementaire dont plusieurs éléments ont déjà été éprouvés. En effet, plusieurs mesures du DSA puisent directement leurs origines dans les législations nationales et d'autres proviennent des études sur la régulation dont l'efficacité a été testée dans d'autres domaines (notamment les analyses de gestion des risques). Autrement dit, les mesures du DSA ne sont finalement pas des nouveautés en soi et elles permettront d'outiller progressivement les autorités judiciaires et administratives afin de les aider à cerner et analyser le phénomène des contenus en ligne.

Bibliographie

Ouvrages et articles scientifiques

Abbott, K. W., Levi-Faur, D., & Snidal, D. (2017). *Introducing Regulatory Intermediaries*. The ANNALS of the American Academy of Political and Social Science, 670(1), 6-13. <https://doi.org/10.1177/0002716217695519>.

Alkiviadou, N. (2019). Hate speech on social media networks: towards a regulatory framework? *Information & Communications Technology Law*, 28(1). <https://doi.org/10.1080/13600834.2018.1494417>

Andersen, J. (2018). Archiving, ordering, and searching: search engines, algorithms, databases, and deep mediatization. *Media, Culture & Society*, 40(8), <https://doi.org/10.1177/0163443718754652>

Banks, J. (2010). Regulating hate speech online. *International Review of Law, Computers & Technology*, 24(3), pp. 233-239. <https://doi.org/10.1080/13600869.2010.522323>

Benkler, Y. (2006) *The Wealth of Networks*. Yale University Press, New Haven.

Biettri, E. (2023). A Genealogy of Digital Platforms Regulation. *Georgetown Law Technology Review*, 7(1), pp. 1- 68. Disponible sous : [A Genealogy of Digital Platform Regulation by Elettra Bietti :: SSRN](#)

Botero Arcila, B., Griffin, R. (2023). *Social media platforms and challenges for democracy, rule of law and fundamental rights*. Etude pour le parlement européen. Disponible sous : [https://www.europarl.europa.eu/RegData/etudes/STUD/2023/743400/IPOL_STU\(2023\)743400_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2023/743400/IPOL_STU(2023)743400_EN.pdf)

Bukovska, B., Callamard, A., Parrmer, S. (2010). *Towards an interpretation of article 20 of the ICCPR : Thresholds for the prohibition of incitement to hatred*. www.ohchr.org/Documents/Issues/Expression/ICCPR/Vienna/CRP7Callamard.pdf.

Cohen, J. (2019). *Between Truth and Power – The Legal Constructions of Informational Capitalism*. Oxford University Press

Colombier, N., M'chirgui, Z., Pénard, T. (2010). An Empirical Analysis of Internet Providers' Interconnection Strategies. *Revue d'économie industrielle*, 131(3). <https://doi.org/10.4000/rei.4178>

Crain, M. (2016). The limits of transparency: Data brokers and commodification. *New Media & Society*, 20(1). <https://doi.org/10.1177/1461444816657096>

Gelber, K. (2019). Differentiating hate speech: a systemic discrimination approach. *Critical Review of International Social and Political Philosophy*, 24(4). <https://doi.org/10.1080/13698230.2019.1576006>

Gillespie, T. (2018). *Custodians of the Internet*. New Haven/London: Yale University Press.

Goldman, E. (2020). An Overview of the United States' Section 230 Internet Immunity. In Frosio, G. (ed.) *The Oxford Handbook of Online Intermediary Liability*. Oxford University Press.

Gorwa, R., Binns, R., Katzenbach, C. (2020). Algorithmic content moderation: technical and political challenges in the automation of platform governance. *Big Data & Society*, 7(1). <https://doi.org/10.1177/2053951719897945>.

Klonick, K. (2018). The New Governors. *Harvard Law Review*, 131(6), 1598-1670.

- Kuczerawy, A. (2015). Intermediary liability & freedom of expression: Recent developments in the EU notice & Action initiative. *Computer Law & Security Review*, 31(1), 46-56. <https://doi.org/10.1016/j.clsr.2014.11.004>
- Maarouf, A., Pröllochs, N & Feurriegel, S. (2022). The Virality of Hate Speech on Social Media. <https://doi.org/10.48550/arXiv.2210.13770>
- Majone, G. (2006). *Agenda Setting*. Moran, M., Rein, M., Goodin, R.E. (eds). *The Oxford Handbook of Public Policy*. Oxford University Press.
- Neudert, L.-M. (2023). Regulatory capacity capture: The United Kingdom's online safety regime. *Internet Policy Review*, 12(4). <https://doi.org/10.14763/2023.4.1730>
- Stockmann, D. (2023). Tech companies and the public interest: the role of the state in governing social media platforms. *Information, Communication & Society*, 26(1), 1-15. <https://doi.org/10.1080/1369118X.2022.2032796>
- Subeid, S., Wheatley, S., Mukherjee, A., Ntube Ngane, S. (2011). The Role of the Special Rapporteurs of the United Nations Human Rights Council in the Development and Promotion of International Human Rights Norms. *The International Journal of Human Rights*, 15(2), 155-161. <https://doi.org/10.1080/13642987.2011.537463>
- Thiel, M., Maslanik, J. (2010). *Transnational actors*. Oxford Research Encyclopedias: International studies. <https://doi.org/10.1093/acrefore/9780190846626.013.105>
- Weber, R.H. (2021). *Internet Governance at the Point of No Return*. EIZ Publishing.
- Yannopoulos, G. (2017). *The Immunity of Internet Intermediaries Reconsidered?*. In Taddeo, M., Floridi, L. (eds), *The Responsibilities of Online Service Providers*. Springer, pp. 43-59.
- Zuboff, S. (2020). *L'âge du capitalisme de surveillance*. Zulma Essays.

Ressources

Nations Unies

- Nations Unies, Comité des droits de l'homme. (2011). Observation générale n°34, article 19 du Pacte international relatif aux droits civils et politiques, CCPR/C/GC/34. Disponible sous : <https://documents-dds-ny.un.org/doc/UNDOC/GEN/G11/453/32/PDF/G1145332.pdf?OpenElement>
- Nations Unies, Comité des droits de l'homme. (1983). Observation générale n°11, article 20 du Pacte international relatif aux droits civils et politiques. Disponible sous : https://tbinternet.ohchr.org/_layouts/15/treatybodyexternal/TBSearch.aspx?Lang=fr&TreatyID=8&DocTypeID=11
- Nations Unies, Commission des droits humains. (1998). Résolution 1998/42, Document E/CN.4/1998/42. Disponible sous : <http://hrlibrary.umn.edu/UN/1998/Res042.html>
- Nations Unies, Comité pour l'élimination de la discrimination raciale. (2023). Rapport CERD/C/CHE/10-12. Disponible sous : https://tbinternet.ohchr.org/_layouts/15/treatybodyexternal/Download.aspx?symbolno=CERD%2fC%2fCHE%2f10-12&Lang=en
- Nations Unies, Comité pour l'élimination de la discrimination raciale. (2013). Recommandation générale n°35 sur la Convention internationale sur l'élimination de toutes les formes de discrimination

raciale. CERD/C/GC/35. Disponible sous :
<https://undocs.org/Home/Mobile?FinalSymbol=CERD%2FC%2FGC%2F35&Language=E&DeviceType=Desktop&LangRequested=False>

Nations Unies, Conseil de sécurité. (2005). Résolution S/RES/1624. Disponible sous :
<https://documents-dds-ny.un.org/doc/UNDOC/GEN/N05/510/53/PDF/N0551053.pdf?OpenElement>.

Nations Unies, Conseil des droits de l'homme. (2011). Résolution A/HRC/RES/16/18. Disponible sous :
<https://undocs.org/Home/Mobile?FinalSymbol=A%2FHRC%2FRES%2F16%2F18&Language=E&DeviceType=Desktop&LangRequested=False>

Nations Unies, Conseil des droits de l'homme. (2008). Résolution 7/36. Disponible sous :
https://ap.ohchr.org/documents/F/HRC/resolutions/A_HRC_RES_7_36.pdf

Nations Unies, Haut-Commissariat aux droits de l'homme. (2012). Plan d'action de Rabat. Disponible sous [La liberté d'expression contre l'incitation à la haine : le HCDH et le Plan d'action de Rabat | OHCHR](#).

Nations Unies, Rapporteur spécial sur la promotion et la protection du droit à la liberté d'opinion et d'expression. (2019). Rapport A/74/486 de Monsieur David Kaye. Disponible sous :
<https://undocs.org/Home/Mobile?FinalSymbol=A%2F74%2F486&Language=E&DeviceType=Desktop&LangRequested=False>.

Nations Unies, Rapporteur spécial sur la promotion et la protection du droit à la liberté d'opinion et d'expression. (2018). Rapport A/HRC/38/35 de Monsieur David Kaye. Disponible sous :
<https://undocs.org/Home/Mobile?FinalSymbol=A%2FHRC%2F38%2F35&Language=E&DeviceType=Desktop&LangRequested=False>.

Nations Unies, Rapporteur spécial sur la promotion et la protection du droit à la liberté d'opinion et d'expression. (2017). Rapport A/HRC/35/22 de Monsieur David Kaye. Disponible sous :
<https://www.undocs.org/Home/Mobile?FinalSymbol=A%2FHRC%2F35%2F22&Language=E&DeviceType=Desktop&LangRequested=False>.

Nations Unies, Rapporteur spécial sur la promotion et la protection du droit à la liberté d'opinion et d'expression. (2016). Rapport A/HRC/32/38 de Monsieur David Kaye. Disponible sous :
<https://undocs.org/Home/Mobile?FinalSymbol=A%2FHRC%2F32%2F38&Language=E&DeviceType=Desktop&LangRequested=False>.

Nations Unies, Rapporteur spécial sur la promotion et la protection du droit à la liberté d'opinion et d'expression. (2013). Rapport A/HRC/23/40 de Monsieur Franck La Rue. Disponible sous :
<https://undocs.org/Home/Mobile?FinalSymbol=A%2FHRC%2F23%2F40&Language=E&DeviceType=Desktop&LangRequested=False>.

Nations Unies, Rapporteur spécial sur la promotion et la protection du droit à la liberté d'opinion et d'expression. (2012). Rapport A/67/357 de Monsieur Franck La Rue. Disponible sous
<https://undocs.org/Home/Mobile?FinalSymbol=A%2F67%2F357&Language=E&DeviceType=Desktop&LangRequested=False>

Nations Unies, Rapporteur spécial sur la promotion et la protection du droit à la liberté d'opinion et d'expression. (2011). Rapport A/HRC/17/27 de Monsieur Franck La Rue. Disponible sous
<https://undocs.org/Home/Mobile?FinalSymbol=A%2F67%2F357&Language=E&DeviceType=Desktop&LangRequested=False>

Nations Unies, Rapporteur spécial sur la promotion et la protection du droit à la liberté d'opinion et d'expression. (2011). Rapport A/66/290 de Monsieur Franck La Rue. Disponible sous <https://documents-dds-ny.un.org/doc/UNDOC/GEN/N11/449/79/PDF/N1144979.pdf?OpenElement>.

Nations Unies, Rapporteur spécial sur la promotion et la protection du droit à la liberté d'opinion et d'expression. (2007). Rapport A/HRC/4/27 de Monsieur Ambeyi Ligabo. Disponible sous : <https://documents-dds-ny.un.org/doc/UNDOC/GEN/G07/101/82/PDF/G0710182.pdf?OpenElement>.

Nations Unies, Rapporteur spécial sur la promotion et la protection du droit à la liberté d'opinion et d'expression. (2005). Rapport E/CN.4/2005/64 de Monsieur Ambeyi Ligabo. Disponible sous https://ap.ohchr.org/documents/alldocs.aspx?doc_id=11160

Nations Unies, Rapporteur spécial sur la promotion et la protection du droit à la liberté d'opinion et d'expression. (2001). Rapport E/CN.4/2001/64 de Monsieur Abid Hussain. Disponible sous : <https://documents-dds-ny.un.org/doc/UNDOC/GEN/G01/111/24/PDF/G0111124.pdf?OpenElement>.

Nations Unies, Rapporteur spécial sur la promotion et la protection du droit à la liberté d'opinion et d'expression. (2000). Rapport E/CN.4/2000/63 de Monsieur Abid Hussain. Disponible sous <https://documents-dds-ny.un.org/doc/UNDOC/GEN/G00/102/60/PDF/G0010260.pdf?OpenElement>

Nations Unies, Rapporteur spécial sur la promotion et la protection du droit à la liberté d'opinion et d'expression. (1999). Rapport E/CN.4/1999/64 de Monsieur Abid Hussain. Disponible sous <https://documents-dds-ny.un.org/doc/UNDOC/GEN/G99/107/67/PDF/G9910767.pdf?OpenElement>

Nations Unies, Rapporteur spécial sur la promotion et la protection du droit à la liberté d'opinion et d'expression. (1998). Rapport E/CN.4/1998/40 de Monsieur Abid Hussain. Disponible sous <https://documents-dds-ny.un.org/doc/UNDOC/GEN/G98/103/13/PDF/G9810313.pdf?OpenElement>

Nations Unies, Rapporteur spécial sur la promotion et la protection des droits de l'homme et des libertés fondamentales dans la lutte antiterroriste. (2010). Rapport A/HRC/16/51 de Monsieur Martin Scheinin. Disponible sous <https://documents-dds-ny.un.org/doc/UNDOC/GEN/G10/178/99/PDF/G1017899.pdf?OpenElement>.

Nations Unies. (2023). *Discours de haine*. Disponible sous : <https://www.un.org/fr/hate-speech/understanding-hate-speech/what-is-hate-speech>

Nations Unies. (2002). Déclaration et programme d'action de Durban. Disponible sous : https://www.ohchr.org/sites/default/files/Documents/Publications/DurbanDecProgAction_fr.pdf

Conseil de l'Europe

Comité des ministres du Conseil de l'Europe. (1997). Recommandation n°R(97)20 sur le « discours de haine ». Disponible sous : <https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=090000168050116d>

Commission européenne contre le racisme et l'intolérance. (2015). Recommandation de politique générale n°15. Disponible sous : <https://rm.coe.int/recommandation-de-politique-generale-n-15-de-l-ecri-sur-la-lutte-contr/16808b5b03>

Service de l'exécution des arrêts de la Cour Européenne des droits de l'homme. (2022). Crimes de haine et discours de haine. Disponible sous : <https://rm.coe.int/thematic-factsheet-hate-crime-fra-docx/1680a96866>

Allemagne

Deutscher Bundestag. (2022). Antwort der Bundesregierung, Digital Services Act, Drucksache 20/2308. 17. Juni 2022. Disponible sous : [Drucksache 20/2308 \(bundestag.de\)](https://www.bundestag.de/Drucksache/20/2308)

Deutscher Bundestag. (2017). Gesetzesentwurf der Fraktionen der CDU/CSU und SPD. Drucksache 18/12356, p.1. Disponible sous : <https://dserver.bundestag.de/btd/18/123/1812356.pdf>.

Deutscher Bundestag. (2017). Stenografischer Bericht, 18. Wahlperiode – 235. Sitzung. Berlin, Freitag, den 19. Mai 2017. Disponible sous : <https://dserver.bundestag.de/btp/18/18235.pdf>.

Deutscher Bundestag. (2017). Stenografischer Bericht. 18. Wahlperiode – 235. Sitzung. Berlin, Freitag, den 30. Juni 2017. Disponible sous : <https://dserver.bundestag.de/btp/18/18244.pdf>.

Deutscher Bundestag. (2017). Beschlussempfehlung und Bericht des Ausschusses für Recht und Verbraucherschutz. Drucksache 18/13013. Disponible sous : <https://dserver.bundestag.de/btd/18/130/1813013.pdf>.

Deutscher Bundesrat. (2017). Beschluss des Bundesrates – Gesetz zur Verbesserung der Rechtsdurchsetzung in sozialen Netzwerken. Drucksache 536/17. Disponible sous : <https://dserver.bundestag.de/brd/2017/0536-17.pdf>.

France

Assemblée nationale. (2020). Texte n°419, 13 mai 2020. Disponible sous : https://www.assemblee-nationale.fr/dyn/15/textes/15t0419_texte-adoptee-seance#

Assemblée nationale. (2020). Proposition de loi n°388, 22 janvier 2020. Disponible sous : https://www.assemblee-nationale.fr/dyn/15/textes/15t0388_texte-adoptee-seance#

Assemblée nationale. (2020). Rapport n°2583, 15 janvier 2020. Disponible sous : https://www.assemblee-nationale.fr/dyn/15/rapports/cion_lois/15b2583_rapport-fond#

Assemblée nationale. (2019). Proposition de loi n°1785 visant à lutter contre la haine sur internet. Disponible sous : https://www.assemblee-nationale.fr/dyn/15/textes/15b1785_proposition-loi.

Assemblée nationale. (2019). Rapport n°2062, 19 juin 2019. Disponible sous : https://www.assemblee-nationale.fr/dyn/15/rapports/cion_lois/15b2062_rapport-fond#

Assemblée nationale. (2019). Analyse du scrutin n°2039, 9 juillet 2019. Disponible sous : [https://www2.assemblee-nationale.fr/scrutins/detail/\(legislature\)/15/\(num\)/2039](https://www2.assemblee-nationale.fr/scrutins/detail/(legislature)/15/(num)/2039)

Commission mixte paritaire. (2020). Rapport n°239, 8 janvier 2020. Disponible sous : <https://www.senat.fr/rap/19-239/19-239.html>

Conseil constitutionnel. (2020). Décision n°2020-801, 18 juin 2020. Disponible sous : <https://www.conseil-constitutionnel.fr/decision/2020/2020801DC.htm>

Délégation interministérielle à la lutte contre le racisme, l'antisémitisme et la haine anti-LGBT. (2020). Lutter contre la haine en ligne. Disponible sous : [Lutter contre la haine en ligne | DILCRAH](https://www.lutte-contre-la-haine-en-ligne.fr/)

Délégation interministérielle à la lutte contre le racisme, l'antisémitisme et la haine anti-LGBT. (2018). Plan national de lutte contre le racisme et l'antisémitisme 2018-2020. Disponible sous : <https://www.gouvernement.fr/plan-national-de-lutte-contre-le-racisme-et-l-antisemitisme-2018-2020>

Ministère de la justice. (2021). Lutte contre la haine en ligne. Disponible sous : [Lutte contre la haine en ligne | Ministère de la justice](https://www.lutte-contre-la-haine-en-ligne.fr/)

Ministère de la justice. (2020). Circulaire relative à la lutte contre la haine en ligne, 24 novembre 2020. Disponible sous : [JUSD2032620C.pdf \(justice.gouv.fr\)](#).

Sénat. (2020). Texte n°64, 26 février 2020. Disponible sous : <https://www.senat.fr/leg/tas19-064.html>

Sénat. (2020). Rapport n°299, 5 février 2020. Disponible sous : <https://www.senat.fr/rap/I19-299/I19-299.html>

Sénat. (2019). Commission des lois, Rapport n°197 (2019-2020), 11 décembre 2019. Disponible sous : <https://www.senat.fr/rap/I19-197/I19-197.html>

Sénat. (2019). Texte n°36 rectifié, 17 décembre 2019. Disponible sous : <https://www.senat.fr/leg/tas19-036.html>

Autriche

Bundesministerium Justiz. (2020). Regierungsvorlage, Erläuterung zum Gesetzestext. Disponible sous : [fname_848587.pdf \(parlament.gv.at\)](#)

Bundesministerium Justiz. (2020). Ministerialentwurf betreffend Bundesgesetz, mit dem ein Bundesgesetz über Massnahmen zum Schutz der Nutzer auf Kommunikationsplattformen erlassen wird, 9 septembre 2020. Disponible sous : <https://www.parlament.gv.at/gegenstand/XXVII/ME/49>

Bundesrat. (2020). Beschluss im Bundesrat 154/BNR. Disponible sous : <https://www.parlament.gv.at/gegenstand/XXVII/BNR/154>

Nationalrat. (2020). Bericht des Verfassungsausschusses, 25 novembre 2020. Disponible sous : https://www.parlament.gv.at/dokument/XXVII/I/509/fname_851399.pdf

Nationalrat. (2020). Beschluss des Nationalrates vom 10. Dezember 2020. Disponible sous : <https://www.parlament.gv.at/gegenstand/XXVII/BNR/154>

Union Européenne

Commission européenne. (2020). *Impact assessment accompanying the proposal for a regulation of the European parliament and of the council*. Disponible sous : <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52020SC0348>

Commission Européenne. (2020). Proposition de loi, 15 décembre 2020. Disponible sous : <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020PC0825>

Commission Européenne. (2016). Code of conduct on countering illegal hate speech online. Disponible sous : https://commission.europa.eu/strategy-and-policy/policies/justice-and-fundamental-rights/combating-discrimination/racism-and-xenophobia/eu-code-conduct-countering-illegal-hate-speech-online_en?prefLang=fr

Commission Européenne. (2022). Code of Practice on Disinformation. Disponible sous : <https://digital-strategy.ec.europa.eu/fr/policies/code-practice-disinformation>

Conseil Européen. (2022). Communiqué de presse du 4 octobre 2022. Disponible sous : <https://www.consilium.europa.eu/fr/press/press-releases/2022/10/04/dsa-council-gives-final-approval-to-the-protection-of-users-rights-online/>

Cour Européenne de Justice. (2023). Google Ireland Limited, Meta platforms Ireland Limited, Tik Tok Technology Limited contre Kommunikationsbehörde Austria (KommAustria), Arrêt du 9 novembre 2023, Affaire C-376-22. Disponible sous : <https://curia.europa.eu/juris/documents.jsf?num=C-376/22>

Parlement européen. (2022). Verbatim, débats du 19 janvier 2022. Disponible sous : https://www.europarl.europa.eu/doceo/document/CRE-9-2022-01-19_EN.pdf.

Parlement européen. (2022). Textes adoptés, 5 juillet 2022. Disponible sous : https://www.europarl.europa.eu/doceo/document/TA-9-2022-0269_FR.html

Royaume-Uni

Joint Committee on the Draft Online Safety Bill. (2021). Report of Session 2021-22, 14 December 2021. Disponible sous : <https://committees.parliament.uk/publications/8206/documents/84092/default/>

House of Commons. (2021). Online Safety Act, 1st reading. Disponible sous : <https://bills.parliament.uk/bills/3137/stages/16299>

House of Commons. (2022). Online Safety Act, 2nd reading. Disponible sous : <https://bills.parliament.uk/bills/3137/stages/16300>

House of Lords. (2023) Online Safety Act, 1st reading. Disponible sous : <https://bills.parliament.uk/bills/3137/stages/17303>

House of Lords. (2023) Online Safety Act, 2nd reading. Disponible sous : <https://bills.parliament.uk/bills/3137/stages/17304>

House of Lords. (2023). 3rd reading. Disponible sous : <https://bills.parliament.uk/bills/3137/stages/17870>

UK Government. (2022). Government Response to the Report of the Joint Committee on the Draft Online Safety Bill, March 2022. Disponible sous : https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1061446/E02721600_Gov_Resp_to_Online_Safety_Bill_Accessible_v1.0.pdf

UK Government. (2020). Department for Digital, Culture, Media & Sport and Home Office, Online Harms White Paper: Full government response to the consultation. Disponible sous : <https://www.gov.uk/government/consultations/online-harms-white-paper/outcome/online-harms-white-paper-full-government-response>

UK Government. (2019). Government Online Harms White Paper, April 2019. Disponible sous : https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/973939/Online_Harms_White_Paper_V2.pdf

UK Government. (2018). Government response to the Internet Safety Strategy Green Paper, May 2018. Disponible sous : https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/708873/Government_Response_to_the_Internet_Safety_Strategy_Green_Paper_-_Final.pdf

Autres

Euronews. (2024). Platform rules to apply this month as member states lag on oversight personnel. Edition du 7 février 2024. Disponible sous : [Platform rules to apply this month as member states lag on oversight personnel | Euronews](#)

Le Temps. (2024). Le nombre d'actes antisémites a explosé en 2023. Edition du 15 février 2024. Disponible sous : [Le nombre d'actes antisémites a explosé en 2023 - Le Temps](#)

Légipresse. (2020). Mise en œuvre du « Parquet numérique » prévu par la loi. Edition du 27 novembre 2020. Disponible sous : [Mise en œuvre du « Parquet numérique » prévu par la loi Avia - Communications électroniques \(legipresse.com\)](#)

Manila Principles on Intermediary Liability. Disponible sous : <https://manilaprinciples.org/principles.html>

Nations Unies, Haut-Commissariat aux droits de l'homme. (2011). Conclusions et recommandations issues des quatre ateliers d'experts organisés par le HCDH en 2011 et adoptés par les experts à Rabat, Maroc, le 5 octobre 2012. Disponible sous : https://www.ohchr.org/sites/default/files/Rabat_draft_outcome_FR.pdf

Ortiz-Opsina, E. (2019). The rise of social media. Disponible sous : <https://ourworldindata.org/rise-of-social-media>

The Santa Clara Principles on Transparency and Accountability in Content Moderation. Disponible sous : <https://santaclaraprinciples.org/>

US Congress, Congressional Research Service. (2024). Section 230: An Overview. Disponible sous : <https://crsreports.congress.gov/product/pdf/R/R46751>

Textes de lois

Allemagne. Gesetz zur Verbesserung der Rechtsdurchsetzung in sozialen Netzwerken. Loi du 1^{er} septembre 2017. Disponible sous <https://www.gesetze-im-internet.de/netzdg/BJNR335210017.html>

Allemagne. Gesetz zur Änderung des Netzwerkdurchsetzungsgesetzes. Loi du 3 juin 2021. Disponible sous : [NetzDGÄndG Gesetz zur Änderung des Netzwerkdurchsetzungsgesetzes \(buzer.de\)](#)

Allemagne. Strafgesetzbuch. Disponible sous : <https://www.gesetze-im-internet.de/stgb/>

Autriche. Kommunikationsplattformengesetz. Loi du 23 décembre 2020. Disponible sous : <https://www.ris.bka.gv.at/GeltendeFassung.wxe?Abfrage=Bundesnormen&Gesetzesnummer=20011415&FassungVom=2023-07-09>

Autriche. Strafgesetzbuch. Disponible sous : <https://www.jusline.at/gesetz/stgb>

Autriche. Bundesgesetz über die Einrichtung einer Kommunikationsbehörde Austria. Loi du 30 mars 2001. Disponible sous : [2001_32_I.pdf \(bka.gv.at\)](#)

Autriche. Koordinator für digitale Dienste Gesetz. Loi du 30 décembre 2023. Disponible sous : [BGBLA_2023_I_182.pdfsig \(bka.gv.at\)](#).

Etats-Unis. Communications Act of 1934. Disponible sous : <https://transition.fcc.gov/Reports/1934new.pdf>

France. Code pénal. Disponible sous : https://www.legifrance.gouv.fr/codes/texte_lc/LEGITEXT000006070719/2020-12-16

France. Décret n°2021-1853 du 27 décembre 2021 relatif aux compétences et à l'organisation de l'Autorité de régulation de la communication audiovisuelle et numérique en matière de protection du droit d'auteur et des droits voisins. Disponible sous : [Décret n° 2021-1853 du 27 décembre 2021 relatif aux compétences et à l'organisation de l'Autorité de régulation de la communication](#)

audiovisuelle et numérique en matière de protection du droit d'auteur et des droits voisins - Légifrance (legifrance.gouv.fr)

France. Loi n°2020-766 du 24 juin 2020 visant à lutter contre les contenus haineux sur Internet. Disponible sous : <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000042031970>

France. Loi n°8925 du 17 janvier 1989 modifiant la loi n°86-1067 du 30 septembre 1986 relative à la liberté de communication. Disponible sous : [Loi n° 89-25 du 17 janvier 1989 modifiant la loi n° 86-1067 du 30 septembre 1986 relative à la liberté de communication - Légifrance \(legifrance.gouv.fr\)](#)

France. Loi n°86-1067 du 30 septembre 1986 relative à la liberté de communication. Disponible sous : [Loi n° 86-1067 du 30 septembre 1986 relative à la liberté de communication \(Loi Léotard\) - Légifrance \(legifrance.gouv.fr\)](#)

France. Loi du 29 juillet 1881 sur la liberté de la presse. Disponible sous : <https://www.legifrance.gouv.fr/loda/id/LEGISCTA000006089701>

Nations Unies. Convention internationale sur l'élimination de toutes les formes de discrimination raciale du 4 janvier 1969. Disponible sous : <https://www.ohchr.org/fr/instruments-mechanisms/instruments/international-convention-elimination-all-forms-racial>.

Nations Unies. Pacte international relatif aux droits civils et politiques du 23 mars 1976. Disponible sous : <https://www.ohchr.org/fr/instruments-mechanisms/instruments/international-covenant-civil-and-political-rights>

Nations Unies. Protocole facultatif à la convention relative aux droits de l'enfant concernant la vente d'enfants, la prostitution des enfants et la pornographie mettant en scène des enfants. Disponible sous : <https://www.ohchr.org/fr/instruments-mechanisms/instruments/optional-protocol-convention-rights-child-sale-children-child>.

Nations Unies. Convention pour la prévention et la répression du crime de génocide. Disponible sous : <https://www.ohchr.org/fr/instruments-mechanisms/instruments/convention-prevention-and-punishment-crime-genocide>.

Royaume-Uni. Online Safety Act. Loi du 26 octobre 2023. Disponible sous : [Online Safety Act 2023 \(legislation.gov.uk\)](#)

Union Européenne. Directive sur les services électroniques du 8 juin 2000. Disponible sous : [Directive - 2000/31 - EN - e-commerce directive - EUR-Lex \(europa.eu\)](#)

Union Européenne. Règlement sur les services numériques du 19 octobre 2022. Disponible sous : https://eur-lex.europa.eu/legal-content/EN/TXT/?toc=OJ%3AL%3A2022%3A277%3ATOC&uri=uriserv%3AOJ.L_2022.277.01.0001.01.ENG