

RESEARCH OUTPUTS / RÉSULTATS DE RECHERCHE

L'Union européenne et la régulation de la désinformation

Poullet, Yves; Bontridder, Noemi

Published in:

Revue du Droit des Technologies de l'information

Publication date:

2021

Document Version

le PDF de l'éditeur

[Link to publication](#)

Citation for pulished version (HARVARD):

Poullet, Y & Bontridder, N 2021, 'L'Union européenne et la régulation de la désinformation', *Revue du Droit des Technologies de l'information*, numéro 81, pp. 29-52.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

L'Union européenne et la régulation de la désinformation¹

Yves Pouillet² et Noémi Bontridder³

La désinformation, présente sur nos réseaux sociaux en particulier en ces temps de Covid, inquiète et justifie l'action des autorités publiques. Le premier souci doit être, au nom de la liberté d'expression, de définir de manière restrictive le concept et de la distinguer tant du message illicite que de l'information fautive, inexacte ou trompeuse. La désinformation doit renvoyer à l'utilisation consciente de la technologie pour amplifier le message et ce, à des fins contraires à l'intérêt général. Ainsi, il sera montré combien l'intelligence artificielle, à la fois, amplifie le phénomène mais peut également, moyennant quelques précautions, prévenir la désinformation. La lutte contre la désinformation peut suivre la voie d'une corégulation ascendante, adoptée dans un premier temps, ou d'une corégulation descendante, désormais privilégiée par l'Union européenne. Par ailleurs, l'examen des textes européens montre combien à des vues répressives du phénomène peuvent s'ajouter avec bonheur des mesures positives comme celles d'une politique d'éducation aux médias et de soutien à une presse pluraliste, critique et indépendante.



Disinformation, present on our social networks especially in these times of Covid, generates worries and justifies action from public authorities. The first concern must be, in the name of freedom of expression, to narrowly define the concept and distinguish it from both illicit message and false, inaccurate or misleading information. Disinformation must refer to the conscious use of technology to amplify the message for purposes contrary to the public interest. It will be shown that artificial intelligence systems amplify the phenomenon but can also, with some precautions, prevent disinformation. The fight against disinformation can follow the path of bottom-up co-regulation, that was adopted initially, or the one of downward co-regulation, now favored by the Europe Union. Moreover, an examination of the European texts shows that positive measures – such as media education and support for a pluralist, critical and independent press – can be adopted instead of repressive measures against the phenomenon.

1. V. de Coorebyter écrivait en conclusion d'un ouvrage récent: «Rien n'est pire que l'effacement de la frontière entre un fait et une opinion, entre une information et une allégation. Certes, il n'existe jamais de "fait" à l'état isolé, indépendant d'un cadre théorique qui

permet de le penser, ou d'un espace social qui veille à sa construction et à sa communication.... Le métier de journaliste est de se cogner aux faits, de chercher à repérer l'incertain. Il n'y a donc pas à transiger sur la différence entre le rôle des médias et celui des «ingénieurs du chaos», comme les appelle Giuliano da Empoli⁴. Les ingénieurs du chaos

¹ L'article est à jour à la date du 15 février 2021.

² Professeur émérite à l'Université de Namur, Professeur associé UCLille.

³ Chercheuse au CRIDS/NaDI.

⁴ G. DA EMPOLI, *Les ingénieurs du chaos*, Paris, JC Lattès, 2019.



sont des professionnels de la désinformation qui jouent sur deux claviers, le plus archaïque et le plus moderne : sur les colères, les peurs, la rage des milieux populaires en déshérence, et sur le recours à une informatique sophistiquée pour abreuver les électeurs de fausses informations et de promesses taillées sur mesure... »⁵.

Si incontestablement les actions de désinformation décrites par l'auteur du livre : « Les ingénieurs du chaos » mettent en cause nos démocraties, comme l'affaire *Cambridge Analytica*, la stratégie de l'ex-président Trump, le vote du Brexit, aujourd'hui la désinformation sur les mérites des vaccins, l'illustrent bien⁶, l'intervention réglementaire s'avère délicate. La difficulté de distinguer information erronée, opinion, parodie et désinformation n'est pas chose aisée, et une intervention malencontreuse et hâtive apparaîtrait facilement comme une atteinte à la liberté d'expression voire une censure. À cet égard, nous rappelons le célèbre attendu de la Cour de Strasbourg dans l'affaire *Handyside contre Royaume-Uni* : la liberté d'expression « vaut non seulement pour les "informations" ou "idées" accueillies avec faveur ou considérées comme inoffensives ou indifférentes, mais aussi pour celles qui heurtent, choquent ou inquiètent l'État ou une fraction quelconque de la population. Ainsi le veulent le pluralisme, la tolérance et l'esprit d'ouverture sans lesquels il n'est pas de "société démocratique" »⁷. Ce risque est d'autant plus grand vu

les craintes de censure exprimées par d'aucuns, sachant que les autorités publiques ne voient pas toujours d'un bon œil les critiques adressées par les médias et les journalistes, parfois perçues comme trop insistantes⁸.

2. La manipulation de ce qui est avancé comme la « vérité » est certes une pratique bien ancienne⁹, il n'en reste pas moins que les récentes avancées dans les technologies de l'information et de la communication (ou « TIC ») la rendent particulièrement aisée et efficace. À cet égard, les systèmes d'intelligence artificielle (ci-après « IA ») jouent un rôle majeur, puisque non seulement ils étendent les possibilités de création de faux contenus que l'on nomme communément les « *deepfakes* »¹⁰,

⁸ À ce propos, voy. la Déclaration conjointe du Rapporteur spécial des Nations unies sur la promotion et la protection du droit à la liberté d'opinion et d'expression, du Représentant de l'Organisation pour la sécurité et la coopération en Europe (OSCE) pour la liberté des médias, du Rapporteur spécial de l'Organisation des États américains (OEA) pour la liberté d'expression, du Rapporteur spécial sur la liberté d'expression et l'accès à l'information de la Commission africaine des droits de l'homme et des peuples (CADHP) sur la liberté d'expression et les fausses nouvelles « fake news », la désinformation et la propagande, adoptée à Vienne le 3 mars 2017, p. 1 : « Les autorités publiques dénigrent, intimident et menacent les médias, notamment en affirmant que ces derniers sont "l'opposition" ou qu'ils "profèrent des mensonges" et ont un agenda politique caché, ce qui accroît le risque de menaces et de violences contre les journalistes, sape la confiance du public dans le journalisme dans son rôle de "chien de garde public", et peut induire le public en erreur en brouillant les lignes entre la désinformation et les contenus médiatiques qui contiennent des informations pouvant faire l'objet de vérifications indépendantes ».

⁹ Pour un historique du phénomène de désinformation en Europe, voy. A. MORK (éd.), *Fake for Real – A history of forgery and falsification*, Temporary Exhibition Catalogue – Maison de l'Histoire Européenne, Luxembourg, Publications Office of the European Union, 2020.

¹⁰ Lorsqu'une technologie d'IA est utilisée pour créer du faux contenu, le produit est nommé un « *deepfake* » – mot qui combine les termes « *deep learning* » et « *fake* ». Les *deepfakes* peuvent par exemple, dans des vidéos particulièrement réalistes, faire dire à des

⁵ V. DE COOREBYTER, « L'internet : démocratie ou démagogie », in Y. POULLET (éd.), *Vie privée, liberté d'expression et démocratie dans la société numérique*, Cahier du CRIDS, n° 47, Bruxelles, Larcier, 2020, p. 252.

⁶ Sur ce point, A. RANGAPPA, « Disinformation, Democracy and the Rule of Law », 30 janvier 2019, disponible sur le site : <http://defusingdis.info/category/essays/>.

⁷ Cour eur. D.H., arrêt *Handyside c. Royaume-Uni*, 7 décembre 1976, req. n° 5493/72, § 49. Voy. également, Cour eur. D.H., arrêt *De Haes et Gijssels c. Belgique*, 24 février 1997, req. n° 19983/92, § 46 ; Cour eur. D.H., arrêt *Özgür Gündem c. Turquie*, 16 mars 2000, req. n° 23144/93, §§ 57 et 60.



mais, d'autre part, ils permettent également de diffuser de manière ciblée et personnalisée tout type de contenu, que celui-ci soit créé à l'aide de systèmes d'IA ou non, pour manipuler l'opinion des individus. En réaction à cette évolution inquiétante pour nos droits, nos libertés et notre démocratie, se développent encore d'autres systèmes d'IA, qui eux visent à lutter contre le phénomène de désinformation en ligne. Autant des systèmes de détection que des systèmes de modération de contenu sont ainsi envisagés, ou parfois même déjà utilisés par certaines plateformes. Nous exposerons ces récentes évolutions technologiques dans un premier chapitre.

3. Face à cela, les autorités publiques ne restent pas de marbre, et ce qui est sans doute le plus remarquable dans l'évolution toute récente de l'approche européenne de la régulation de la désinformation, consiste dans le passage brusque d'un système d'autorégulation ou plutôt de corégulation légère et ascendante à un système de corégulation descendante, parallèle à celui existant en matière de protection des données¹¹. On ajoutera que ce passage affecte également la régulation insidieuse et en tout cas diffuse, mais bien effective, que présentent les systèmes d'IA. Si l'IA peut être un outil utile à la lutte contre la désinformation, ses algorithmes peuvent poursuivre d'autres buts voire limiter illégitime-

ment la liberté d'expression. Ce sera l'objet de notre deuxième chapitre. L'examen du contenu de la régulation proposée réaffirme le principe de la liberté d'expression tout en mettant en place, d'une part, avec le *Digital Services Act*¹², une responsabilité accrue des intermédiaires, en particulier des plateformes considérées comme de véritables *gatekeepers*¹³ et, d'autre part, avec notamment la Résolution du Parlement européen¹⁴, une politique plus positive en faveur d'une information pluraliste et indépendante. Nous développerons ce point dans le troisième chapitre.

¹² Le lecteur/la lectrice trouvera le texte proposé par la Commission européenne: *Proposal for a Regulation of the European Parliament and of the Council on a Single Market For Digital Services (Digital Services Act) and amending Directive 2000/31/EC*, COM(2020) 825 final, 15 décembre 2020, disponible sur le site: <https://ec.europa.eu/digital-single-market/en/digital-services-act-package>.

¹³ Le terme de «*gatekeepers*» est utilisé notamment par la déclaration du Conseil des ministres de l'Union européenne: «*Major platforms which act as gatekeepers often exercise 'secondary' control over content that already complies with editorial standards and the applicable legal provisions*». Council conclusions on safeguarding a free and pluralistic media system, 27 novembre 2020, n° 37, disponible en ligne sur <https://data.consilium.europa.eu/doc/document/ST-13260-2020-INIT/en/pdf>. Ce terme exprime le fait que les plateformes d'information et les réseaux sociaux qui disposent de parts de marchés importantes gèrent l'accès à des services qui sont, dans notre société de l'information, devenus nécessaires à tout citoyen et dès lors devraient être accessibles à chacun. Le concept de service universel adopté au départ à propos des services d'accès aux réseaux, soit à l'infrastructure, apparaît comme intéressant à reprendre à propos de ces plateformes qui seraient tenues d'offrir à un prix abordable ou gratuitement un service de communication ou d'accès à l'information présentant une qualité qui serait définie par la réglementation publique. Nous reviendrons sur ce point, *infra*, n°s 17 et s.

¹⁴ Résolution du Parlement européen du 25 novembre 2020 sur le renforcement de la liberté des médias: la protection des journalistes en Europe, les discours de haine, la désinformation et le rôle des plateformes (2020/2009(INI)), disponible en ligne sur https://www.europarl.europa.eu/doceo/document/TA-9-2020-0320_FR.html.

personnes des choses qu'elles n'ont jamais dites. Ainsi, en mai 2018, un parti a publié une vidéo sur Facebook dans laquelle Trump se moque de la Belgique pour son respect de l'accord de Paris sur le climat: <https://www.facebook.com/watch/?v=10155618434657151>.

¹¹ Sur les différentes approches de la régulation et la comparaison entre la régulation de la protection des données et celle de la désinformation, voy. Y. POULLET, «Vues de Bruxelles – Modes alternatifs de régulation et libertés dans la société du numérique», in *Enjeux internationaux des activités numériques: entre logique territoriale des États et puissance d'acteurs privés*, Colloque du 11 octobre 2019, Toulouse, Larcier, 2020, pp. 91-136.



CHAPITRE I: L'AMBIGUÏTÉ DE L'IA, TECHNOLOGIE À LA FOIS D'APPUI À LA DÉSINFORMATION ET DE LUTTE CONTRE ELLE¹⁵

4. La création de faux contenus à des fins de manipulation de la réalité s'opérait assurément déjà longtemps avant l'avènement du numérique¹⁶, mais ce dernier a rendu la manipulation de textes particulièrement aisée, et les images peuvent être manipulées de manière convaincante depuis un certain temps via des logiciels tels que *Photoshop*. Plus récemment encore, des systèmes d'IA peuvent être programmés non seulement pour générer des textes de manière automatique, mais aussi pour créer des images, des enregistrements audio ou encore des vidéos de manière de plus en plus réaliste. La manipulation de textes et d'images était, comme nous le mentionnions, déjà aisée avec des technologies moins sophistiquées, mais les systèmes d'IA rendent également la création d'enregistrements audio et de vidéos particulièrement accessible¹⁷. Or, si ces «*deep-*

fakes», comme il est convenu de les appeler, sont souvent générés à des fins de divertissement, leur présence dans les médias peut engendrer des effets singulièrement néfastes pour la société: outre le fait qu'ils alimentent indéniablement le flux de fausses informations, ils sont aussi propices à décrédibiliser les informations basées sur des faits ayant eu lieu dans le monde physique, étant donné que le doute peut s'installer à l'égard de toute information présente en ligne, y compris à l'égard de celles émanant de la presse traditionnelle ou de nos gouvernements. Ainsi, les personnes désireuses de manipuler les individus peuvent atteindre leurs objectifs de deux manières: en faisant passer de faux contenus pour vrais, ou en faisant passer de vrais contenus pour faux.

5. Si l'utilisation de systèmes d'IA afin de générer de faux contenus impacte la société de manière alarmante, tel est d'autant plus le cas des systèmes d'IA qui permettent la diffusion ciblée et à grande échelle de contenus générés, automatiquement ou non, à des fins malveillantes. En l'occurrence, le «micro-ciblage» est bien présent sur le web. Le modèle économique du web, basé sur la publicité et dès lors appelé «économie de l'attention»¹⁸,

¹⁵ Pour de plus amples détails sur le sujet, nous renvoyons le lecteur/la lectrice à notre article N. BONTRIDDER et Y. POULLET, *The role of artificial intelligence in disinformation*, 2021, preprint, disponible en ligne sur ResearchGate: https://www.researchgate.net/publication/349731942_The_role_of_artificial_intelligence_in_disinformation; cfr également, le rapport «Prioritisation uncovered – the discoverability of Public interest Content Online», élaboré par E.M. MAZZOLI et D. TAMBINI pour le Conseil de l'Europe, novembre 2001, disponible en ligne sur <https://rm.coe.int/publication-content-prioritisation-report/1680a07a57>.

¹⁶ Il a ainsi par exemple été révélé (en 1440 seulement) que la «donation de Constantin», acte par lequel l'empereur romain Constantin I^{er} aurait conféré au pape Silvestre la primauté sur les Églises d'Orient et l'*imperium* (pouvoir impérial) sur l'Occident au 4^e siècle, avait en réalité été rédigée au cours de la deuxième moitié du 8^e siècle afin de justifier l'expansion des territoires pontificaux, et qu'elle fut de surcroît longtemps encore invoquée pour asseoir les pouvoirs du pape. Voy. A. MORK (éd.), *Fake for Real – A history of forgery and falsification*, op. cit., pp. 37-40.

¹⁷ En effet, quiconque peut se procurer une fausse vidéo de la personne de son choix pour autant

qu'elle dispose de 250 photos de celle-ci, et ce pour seulement 3 dollars, et un faux enregistrement audio peut être généré pour 10 dollars tous les 50 mots. A. M. WALORSKA, *Deepfakes and disinformation*, Friedrich Naumann Foundation for Freedom, Potsdam, mai 2020, disponible en ligne sur https://fnf-europe.org/wp-content/uploads/2020/06/fnf_deepfakes_broschuere_en_web.pdf, p. 15.

¹⁸ Cette économie se base sur les théories de H. SIMON (prix Nobel d'économie) exprimées pour la première fois en 1971 («Designing Organizations for an Information-Rich World», in *Computer, communications and the public interest* (M. GRENNBERGER (éd.)), Baltimore MD: The John Hopkins Press, 1971, pp. 37-72): «Dans un monde riche en informations, l'abondance d'informations entraîne la pénurie d'une autre ressource: la rareté devient ce que consomme l'information. Ce que l'information consomme est assez évident: c'est l'attention de ses receveurs. Donc une abondance d'informations crée une rareté de l'attention

en est la cause. Cette technologie, originellement développée afin de diffuser les publicités de manière plus efficace aux internautes¹⁹, est maintenant largement utilisée pour la diffusion ciblée de tout contenu en ligne. En effet, au plus les internautes passent de temps sur un site ou sur une plateforme, au plus les compagnies publicitaires rémunèrent ce diffuseur de contenu²⁰. Cela incite donc les acteurs du web à viser l'engagement des internautes, ces derniers étant à cette fin profilés par des systèmes d'IA qui leur présentent sur cette base du contenu qui optimisera leur engagement. Ainsi, la hiérarchisation de contenus sur *Google*, le fil d'actualité *Facebook*, la *Timeline* de *Twitter*, les systèmes de recommandation de *YouTube*, de *Spotify* ou encore de *Netflix*, nous enferment dans des « bulles informationnelles » ou « bulles de filtre »²¹: chaque individu en ligne voit une version différente de la réalité. Au-delà des nombreuses implications éthiques mais aussi juridiques de cet écosys-

tème, ce dernier peut être exploité directement par des acteurs malveillants pour manipuler l'opinion des individus. L'exemple connu est le ciblage des votants lors de campagnes électorales, tel que fut manifestement le cas lors des élections de 2016 aux États-Unis²² ou encore lors du référendum de 2016 relatif au *Brexit*, auquel la société *Cambridge Analytica* a contribué de manière significative. Nous notons en outre que ces systèmes amplifient également la diffusion de fausses informations de manière indépendante de toute intervention humaine ultérieure à leur programmation originelle. Pour illustration, ce qui fut nommé « le cercle vicieux de YouTube », peut être défini comme suit: 1. les contenus semant la discorde performant mieux; 2. le système d'IA les promeut afin d'optimiser le temps passé sur la plateforme et; 3. puisque ce type de contenu est plus regardé, les créateurs de contenus en créent davantage²³.

Les *bots* sociaux (*social bots*) contribuent également à la diffusion ciblée et à grande échelle de la désinformation en ligne. Ces *bots* (abréviation pour *software robots*) sont des comptes utilisateurs entièrement ou partiellement automatisés qui opèrent sur les réseaux sociaux et qui sont programmés pour communiquer et imiter le comportement humain. Ces « *artificial speakers* » peuvent être programmés pour

et le besoin de répartir efficacement cette attention parmi la surabondance des sources d'informations qui peuvent la consommer ».

¹⁹ À cet égard, « [t]he advertiser usually sets the targeting parameters (such as demographics and presumed interests), but the platform's algorithmic systems pick the specific individuals who will see the ad and determine the ad's placement within the platform ». N. MARÉCHAL ET E. ROBERTS BIDDLE, *It's Not Just the Content, It's the Business Model: Democracy's Online Speech Challenge – A Report from Ranking Digital Rights*, New America, 17 mars 2020, disponible en ligne sur <https://www.newamerica.org/oti/reports/its-not-just-content-its-business-model/>, p. 13.

²⁰ Les revenus publicitaires de *Facebook* en 2020 atteignaient ainsi 84,2 milliards de dollars. Voy. <https://www.statista.com/statistics/267031/facebook-annual-revenue-by-segment/>.

²¹ « La "bulle de filtres" ou "bulle de filtrage" (de l'anglais: "filter bubble") est un concept développé par le militant d'Internet Eli Pariser. Selon Pariser, la "bulle de filtres" désigne à la fois le filtrage de l'information qui parvient à l'internaute par différents filtres; et l'état d'"isolement intellectuel" et culturel dans lequel il se retrouve quand les informations qu'il recherche sur Internet résultent d'une personnalisation mise en place à son insu » (définition *Wikipedia*).

²² Voy. W. BENETT ET S. LIVINGSTON (éds), *The Disinformation Age: Politics, Technology, and Disruptive Communication in the United State*, Cambridge, Cambridge University Press, octobre 2020, disponible en ligne sur <https://doi.org/10.1017/9781108914628>.

²³ Ceci fut démontré par le projet *AlgoTransparency*, qui vise à démontrer l'impact des algorithmes ayant une influence prépondérante en ligne. Voy. le site consacré au projet: <https://www.algotransparency.org>. À propos des découvertes relatives à l'algorithme *YouTube*, voy. la vidéo: *The Toxic Potential of YouTube's Feedback Loop*, CADE Tech Policy Workshop, 17 novembre 2019, disponible en ligne sur <https://www.youtube.com/watch?v=Et2n0J0OeQ8&feature=youtu.be>.



publier du contenu adapté au public qu'ils sont censés intégrer pour ainsi, une fois qu'ils ont acquis un profil crédible et semblent suffisamment fiables, pouvoir disséminer de la désinformation de manière effective²⁴.

6. Face aux nombreux problèmes liés à la désinformation en ligne, qui est donc intensifiée par les technologies que nous venons de mentionner, d'autres technologies sont développées dans l'objectif de freiner la diffusion massive de contenus néfastes pour la société. Des systèmes d'IA sont ainsi conçus pour détecter les informations fausses, inexacts ou fallacieuses. Il convient d'emblée de noter que ces systèmes ne font donc aucune différence entre «désinformation» et «mésinformation», puisque l'intention de la personne à l'origine du partage doit pour ce faire être déterminée. La désinformation n'est en effet considérée comme telle que dans la mesure où une information fautive, inexacte ou fallacieuse est partagée avec l'intention de tromper²⁵. À l'égard des articles contenant des informations fausses, inexacts ou trompeuses, quatre tech-

niques sont envisagées²⁶ : 1. des systèmes d'IA peuvent être entraînés avec des données labellisées pour procéder directement à leur détection, ceci nécessitant alors d'énormes quantités de données d'apprentissage ; 2. des systèmes d'IA peuvent être programmés pour détecter l'inexactitude des faits contenus dans l'article en les comparant avec des preuves externes ou en vérifiant si l'information est appuyée par d'autres sources déterminées, ou encore en utilisant les ressources du web de manière non structurée ; 3. le style de l'article peut aussi être analysé aux fins de déterminer si celui-ci est trompeur ou non. Notons cependant que la présence (ou l'absence) d'un tel style n'est pas toujours corrélée avec la présence (ou l'absence) de contenus faux, inexacts ou trompeurs ; 4. sont également envisagées des techniques d'analyse de métadonnées plutôt que du contenu même de l'article. La plupart de ces techniques nécessitent toutefois l'analyse de textes dont la signification est souvent difficile à déterminer, ce qui engendre la détection de faux positifs ou de faux négatifs par les systèmes d'IA qui sont utilisés pour ce faire²⁷. Pour détecter les contenus générés par des systèmes d'IA (les dénommés «deepfakes»), et plus particulièrement pour détecter les fausses vidéos, le *Deepfake Detection Challenge* a été

²⁴ J. AKERS *et al.*, *Technology-Enabled Disinformation: Summary, Lessons, and Recommendations*, Technical Report UW-CSE, 2 décembre 2018, disponible en ligne sur <https://arxiv.org/abs/1812.09383>, p. 5.

²⁵ La désinformation est ainsi définie par la Commission européenne comme étant «les informations dont on peut vérifier qu'elles sont fausses ou trompeuses, qui sont créées, présentées et diffusées dans un but lucratif ou dans l'intention délibérée de tromper le public et qui sont susceptibles de causer un préjudice public. Par préjudice public on entend les menaces aux processus politiques et d'élaboration des politiques démocratiques et aux biens publics, tels que la protection de la santé des citoyens de l'Union, l'environnement ou la sécurité. La désinformation ne comprend pas les erreurs de citation, la satire, la parodie, ni les informations et commentaires partisans clairement identifiés». Commission européenne, *Lutter contre la désinformation en ligne: une approche européenne* (Communication), COM(2018) 236 final, 26 avril 2018, disponible en ligne sur <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52018DC0236>.

²⁶ J. AKERS *et al.*, *Technology-Enabled Disinformation: Summary, Lessons, and Recommendations*, *op. cit.*, p. 6.

²⁷ C. MARSDEN *et* T. MEYER, *Regulating Disinformation with artificial intelligence: Effect of disinformation initiatives on freedom of expression and media pluralism*, European Parliamentary Research Service (EPRS) – Scientific Foresight Unit (STOA), PE 624.279, Luxembourg, Publications Office of the European Union, mars 2019, disponible en ligne sur <https://op.europa.eu/en/publication-detail/-/publication/b8722bec-81be-11e9-9f05-01aa75ed71a1>, pp. 1-2. Ces auteurs notent par ailleurs que le système d'IA de Facebook ne dispose par exemple pas encore d'assez de données pour être efficace hors du portugais et de l'anglais. *Ibid.*, p. 17.

lancé fin 2019²⁸ pour encourager le développement de solutions techniques étant donné la non-effectivité de celles qui existent à ce jour. L'utilisation de systèmes d'IA s'avère plus efficace pour détecter les *bots* sociaux, puisque 99.3% des faux comptes supprimés par Facebook au cours du dernier quadrimestre de 2020, avaient été détectés par son système d'IA²⁹.

De même, et pour finir, des systèmes d'IA sont développés pour réguler directement les contenus en ligne. Marsden et Meyer mettent l'accent sur les techniques suivantes³⁰. Le filtrage ou la suppression de contenu est la technique la plus effective mais aussi la plus invasive. Alors que le filtrage s'opère avant la diffusion du contenu pour éviter le téléchargement ou la publication de celui-ci, la suppression s'opère après la publication. Le blocage de contenu consiste quant à lui à bloquer l'accès au contenu pour l'utilisateur, sans pour autant le supprimer. Ces blocages peuvent cependant, dans la plupart des cas, être contournés par l'utilisation d'un VPN (*Virtual Private Network*, ou en français, «réseau privé virtuel»). La «déprioritisation» ou «désindexation» de contenus peut aussi se faire via des systèmes d'IA pour réduire la visibilité de contenus trompeurs. On s'inquiète à cet égard de la tendance alors à affecter le pluralisme des médias, certaines sources considérées comme authentiques étant alors privilégiées, ou à réduire de

manière disproportionnée la visibilité de l'opinion des citoyens. La suppression ou la suspension de comptes sont également des mesures visant à réguler les contenus en ligne, qui sont fort invasives pour les utilisateurs. Ainsi, *Twitter* et *Facebook* ont suspendu le compte de Trump suite à sa violation alléguée des termes et conditions de ces plateformes³¹. Nous reviendrons sur ce point.

7. L'analyse du paysage technologique présent ou à venir sur le web démontre donc une approche axée sur l'innovation. Face à la constatation d'un phénomène qui est amplifié par des systèmes d'IA, la réaction semble en effet consister en l'utilisation d'autres systèmes d'IA censés atténuer les conséquences négatives de ce phénomène. Cette approche est certainement justifiée par l'intérêt moindre des plateformes à changer leurs systèmes de recommandation ou de hiérarchisation de contenus puisque, comme nous l'avons vu, leurs revenus sont générés par les publicités et donc dépendent de l'engagement des internautes. L'utilisation de systèmes d'IA pour détecter et réguler les contenus en ligne comporte pourtant de nombreux risques pour les libertés d'expression et d'information³², puisque ces systèmes ne font pas la différence entre désinformation et mésinformation. Ce qui pose problème est l'utilisation des technologies pour manipuler l'opinion des individus,

²⁸ Voy. Report of the Partnership on AI, *The Deepfake Detection Challenge: Insights and Recommendations for AI and Media Integrity*, 12 mars 2020, disponible en ligne sur <https://www.partnershiponai.org/ai-and-media-integrity-steering-committee/>.

²⁹ Facebook, *Community Standards Enforcement Report*, disponible en ligne sur <https://transparency.facebook.com/community-standards-enforcement#fake-accounts>.

³⁰ Voy. C. MARSDEN et T. MEYER, *Regulating Disinformation with artificial intelligence: Effect of disinformation initiatives on freedom of expression and media pluralism*, *op. cit.*, pp. 38-41.

³¹ Twitter, « Permanent suspension of @realDonaldTrump », 8 janvier 2021, disponible en ligne sur https://blog.twitter.com/en_us/topics/company/2020/suspension.html; Facebook, « Referring Former President Trump's Suspension From Facebook to the Oversight Board », 21 janvier 2021, disponible en ligne sur <https://about.fb.com/news/2021/01/referring-trump-suspension-to-oversight-board/>.

³² La liberté d'expression et la liberté de recevoir des informations sont énoncées à l'article 19 de la Déclaration universelle de droits de l'homme, à l'article 10 de la Convention européenne des droits de l'homme et à l'article 11 de la Charte européenne des droits fondamentaux.



l'intention de nuire étant en effet elle seule sanctionnée par le droit, et diffuser des informations fausses, inexactes ou trompeuses n'est en tant que tel pas condamnable et ne devrait pas le devenir.

CHAPITRE II: DU MODE DE RÉGULATION DE LA DÉSINFORMATION

A. De la corégulation ascendante à la corégulation descendante

8. Pour revenir à l'attitude européenne en la matière, notons que suite à une enquête lancée en 2017 – qui révélait que sur les 3.000 réponses reçues, 97% des répondants estimaient avoir déjà été confrontés aux *fake news*³³ – la Commission a lancé et nommé un «Groupe d'experts de haut niveau sur le phénomène de la désinformation» (*High Level Group on fake news and online disinformation*), chargé de produire des recommandations en la matière. Le rapport du Groupe d'experts européens de janvier 2018³⁴ a largement inspiré la position européenne exprimée dans la Communication: «Lutter contre la désinformation en ligne: une approche européenne»³⁵. Le rapport plaide en faveur d'une autorégula-

tion contrôlée. La position de la Commission ne s'écarte pas de cette option, tout en affirmant quatre principes. Le premier principe concerne la transparence, vis-à-vis des internautes, de l'origine, de la production, de la diffusion, du ciblage et de la sponsorship des informations, ce qui implique l'identifiabilité et la traçabilité des sources. Le deuxième principe insiste sur l'importance d'une offre diversifiée d'informations aux citoyens de l'Union européenne pour les encourager à prendre «des décisions éclairées guidées par l'esprit critique», ce qui implique, ajoute la Commission, l'éducation aux médias et un journalisme de qualité. Le troisième principe concerne la qualité des informations véhiculées dans les réseaux sociaux. Pour ce faire, la Commission suggère des outils tels que les indicateurs de fiabilité de l'information ou les «signaleurs de confiance». La Commission préconise également, aux fins de garantir ou d'évaluer la véracité de l'information, la mise sur pied d'organes de *fact checking* indépendants et soumis à des règles d'éthique et de transparence rigoureuses, à l'instar du Code de l'*International Fact-Checking Network*³⁶. En guise de quatrième principe, la Commission prône la discussion entre l'ensemble des parties concernées par le problème de la désinformation, ce qui implique la participation des plateformes, certes, mais également d'associations d'internautes ou de libertés civiles et d'autorités publiques comme les autorités administratives en charge de l'audiovisuel ou de la protection des données.

³³ Commission européenne, *Synopsis report of the public consultation on fake news and online disinformation*, 26 avril 2018, disponible en ligne sur <https://ec.europa.eu/digital-single-market/en/news/synopsis-report-public-consultation-fake-news-and-online-disinformation>.

³⁴ Rapport du «*High Level Group of Experts on Disinformation*», publié par la Commission européenne (Directorate-General for Communication Networks, Content and Technology): *A multi-dimensional approach to disinformation: Report of the independent High level Group on fake news and online disinformation*, Luxembourg, Publications Office of the European Union, 2018, disponible en ligne sur <https://ec.europa.eu/digital-single-market/en/news/final-report-high-level-expert-group-fake-news-and-online-disinformation>, pp. 5 et 10.

³⁵ Commission européenne, *Lutter contre la désinformation en ligne: une approche européenne*, op. cit.

³⁶ «*The International Fact-Checking Network is a unit of the Poynter Institute dedicated to bringing together fact-checkers worldwide. The IFCN was launched in September 2015 to support a booming crop of fact-checking initiatives by promoting best practices and exchanges in this field*». Sur les compétences et le fonctionnement de ce réseau, voy. le site du réseau: <https://www.poynter.org/ifcn/>.



9. La réunion dite «*Multistakeholders*»³⁷, conviée par la Commission, ne réunissait pas tout ce beau monde mais bien certaines des plus importantes plateformes en ligne (*Facebook, Google, Twitter, Mozilla, Microsoft et TikTok*)³⁸, les annonceurs ainsi que l'industrie publicitaire. Ces acteurs, sans attendre d'autres initiatives de la Commission, adoptèrent, en suivant la voie d'une autorégulation pure ou quasi pure, le «*Code of Practice on Disinformation*». La Commission européenne n'aura d'autre choix, sous peine de s'attirer la colère des signataires du Code, que de l'approuver en septembre 2018³⁹. Le Code reprend certains engagements des parties signataires ou plutôt regroupe l'ensemble des différents engagements auxquels chaque partie pourra souscrire individuellement, partiellement ou non, et y répondre de manière originale et de bonne foi⁴⁰: 1. transparence et analyse (*scrutiny*) des publicités de manière à détecter l'origine des désinformations; 2. identification de la

publicité politique et de son *sponsoring* et diffusion d'opinions (*issue based advertising*); 3. intégrité des services de manière à éviter des fausses représentations (*misrepresentations*) ou des *spams*; 4. «*empowerment*» des consommateurs, ce qui suppose de privilégier les messages authentiques, de développer des produits ou services indicateurs de la fiabilité des messages rencontrés sur la toile et d'éviter des *spams* ou des *misrepresentations*, de donner aux consommateurs l'accès à des sources alternatives et à des sites de dénonciation des messages de désinformation et d'éduquer les internautes au jugement critique; 5. aide au développement de la recherche, y compris académique, notamment pour la mise sur pied d'instruments d'IA et d'analyse automatique des contenus (*content ranking algorithms*), pour la mise sur pied d'organes «indépendants» de *fact checking* et la définition de procédures devant ces organes, et pour une politique de repérage des *chatbots* destinés à amplifier des messages de désinformation. En outre, les signataires s'engagent à publier un rapport sur les mesures prises à propos des différents objets, sur lesquels portent leurs «devoirs». Ce rapport doit en outre reprendre certains indicateurs rendant transparentes l'importance du phénomène de désinformation, les plaintes soumises et les solutions apportées.

10. Ce *Code of Practice on disinformation* représente une forme de corégulation légère que nous qualifions d'ascendante, dans la mesure où l'initiative du contenu est aux mains des acteurs privés et que sa mise en œuvre est l'objet d'un léger contrôle des autorités publiques, à travers le rapport soumis à l'examen de la Commission. L'«infodémie», terme consacré par l'Union européenne pour désigner la désinformation à propos de la pandémie Covid-19, a été l'occasion pour la

³⁷ «*The Multistakeholder Forum on Disinformation comprises two different and autonomous groups. On the one hand, major online platforms ad exchanges and their trade associations, the advertisers and agencies' associations and the European advertising self-regulatory body ("the Working Group"), and on the other hand representatives of the media, civil society, fact checkers and academia ("the Sounding Board")*».

³⁸ Alors que les quatre premières étaient présentes à l'origine de l'adoption du Code en 2018, *Microsoft* a rejoint la liste des signataires en mai 2019 et *TikTok* en juin 2020.

³⁹ Commission européenne (DG Connect), «*Code of Practice on Disinformation*», 26 septembre 2018, mis à jour le 17 juin 2019, disponible en ligne sur <https://ec.europa.eu/digital-single-market/en/news/code-practice-disinformation>.

⁴⁰ «*The purpose of this Code is to identify the actions that Signatories could put in place in order to address the challenges related to "Disinformation" In line with the Commission's Communication, the Signatories of the Code of Practice recognise the importance of efforts. ... Signatories should not be compelled by governments, nor should they adopt voluntary policies, to delete or prevent access to otherwise lawful content or messages solely on the basis that they are thought to be "false"*». *Ibid.*



DOCTRINE

Commission⁴¹ de rappeler de manière ferme les engagements nés du *Code of Practice*. Nonobstant ce rappel, la Commission a estimé que le mécanisme d'autorégulation était insuffisant, pour les motifs suivants: hétérogénéité des rapports et des mesures prises par les signataires, insuffisante couverture du *Code of Practice* par rapport au secteur à couvrir et, surtout, manque d'accès aux données nécessaires à une vérification des pratiques des signataires⁴². L'évaluation, publiée largement, est la voie ouverte par la Commission⁴³ pour modifier

radicalement son approche de la régulation de la désinformation et justifier une approche réglementaire plus proche de celle utilisée à propos de la protection des données: une «corégulation descendante»⁴⁴ marquée par la priorité de la réglementation publique qui laisse certes une marge de manœuvre étroite aux codes de conduite, labels et certifications privés. Ce sera l'objet du projet de règlement appelé *Digital Services Act* (DSA) qui, à travers une modification substantielle des articles 12 et suivants de la directive du 8 juin 2000 sur le commerce électronique en matière de responsabilité des intermédiaires, entend régler la question de la lutte contre la désinformation. Le retour à une réglementation publique de la désinformation semble donc bien acquis⁴⁵.

⁴¹ Communication conjointe au Parlement européen, au Conseil européen, au Conseil, au Comité économique et social européen et au Comité des régions: *Lutter contre la désinformation concernant la Covid-19 – Démêler le vrai du faux*, Bruxelles, 10 juin 2020, JOIN(2020) 8 final, disponible en ligne sur <https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=CELEX%3A52020JC0008>.

⁴² C'est ce que conclut la Commission après deux ans de fonctionnement du «*Code of Practice*», Commission Staff Working Document, *Assessment of the Code of Practice on Disinformation – Achievements and areas for further improvement*, 10 septembre 2020, SWD(2020) 180 final, accessible sur le site: <https://ec.europa.eu/digital-single-market/en/news/assessment-code-practice-disinformation-achievements-and-areas-further-improvement>: «*However, the SWD highlights that, in order to ensure a complete and consistent application across stakeholders and Member States, the Code should be further improved in several areas by providing commonly-shared definitions, clearer procedures, more precise and more comprehensive commitments, as well as transparent key performance indicators (KPIs) and appropriate monitoring. Participation should be broadened to include other relevant stakeholders, in particular from the advertising sector. Moreover, the lack of access to data allowing for an independent evaluation of emerging trends and threats posed by online disinformation, as well as the absence of meaningful KPIs, are fundamental shortcomings of the current Code. At present, it remains difficult to precisely assess the timeliness and impact of signatories' actions, as the Commission and public authorities are still very much reliant on the willingness of platforms to share information and data. Therefore, a more structured model for cooperation between platforms and the research community should be developed*».

⁴³ «*The information and findings set out in this assessment will support the Commission's reflections on pertinent policy initiatives, including the European Democracy Action, as well as the Digital Services Act, which will aim*

to fix overarching rules applicable to all information society services». *Ibid.*

⁴⁴ Cette corégulation descendante est celle que, dès 2003 dans le cadre de l'accord interinstitutionnel «*Mieux légiférer*» (J.O.C.E., C 321, 31 décembre 2003, pp. 1-5), les autorités européennes avaient souhaité. L'article 18 de l'accord définissait, en effet, la corégulation, comme «*le mécanisme par lequel un acte législatif communautaire confère la réalisation des objectifs définis par l'autorité législative aux parties concernées reconnues dans le domaine (notamment les opérateurs économiques, les partenaires sociaux, les organisations non gouvernementales ou les associations)*». Sur cette notion, la distinction entre autorégulation, corégulations ascendante et descendante, voy. nos réflexions in Y. Poullet, «*Technologies de l'information et de la communication et 'co-régulation': une nouvelle approche?*», *Liber amicorum Michel Coipel*, Bruxelles, Kluwer, 2004, p. 173; Y. Poullet, «*How to regulate Internet?: New Paradigms for Internet Governance*», in J. Berleur et alii (éds), *Variations sur le droit de la société de l'information*, Cahier du CRIDS, n° 20, pp. 130 et s.

⁴⁵ À cet égard, outre la loi française qui s'attaque à la question limitée de la publicité politique sur Internet en temps d'élection et celle allemande floue et critiquable à bien des égards (sur ces deux premières initiatives législatives, voy. les commentaires et les références de M. Hanot et A. Michel, «*Entre menaces pour la vie en société et risques réglementaires: les Fake News – un danger pour la démocratie*», in Y. Poullet (éd.), *Vie Privée, liberté d'expression et démocratie dans la société numérique*, Cahier du CRIDS,



11. À l'inverse, la place des codes de conduite, des standards et des labels ou certificats de conformité semble réduite. Le DSA leur réserve les articles 34 et 35 qui rappellent ceux du RGPD à propos des mêmes questions. Il s'agit de promouvoir l'écriture et la mise en place de tels dispositifs, «en tenant compte en particulier des défis spécifiques que représente la lutte contre les différents types de contenus illicites et les risques systémiques, conformément au droit de l'Union, notamment en matière de concurrence et de protection des données à caractère personnel»⁴⁶ et, au moins pour les «*very large platforms*», de répondre à l'invitation adressée par la Commission, de rédiger les codes de conduite en matière de mesures prises pour faire face aux «risques systémiques» (nous reviendrons sur cette notion, *infra*, n° 19) avec les autres plateformes larges ou non mais également toutes les parties intéressées dont les autres services d'intermédiation et les associations de la société civile.

B. L'intelligence artificielle comme mode de régulation de la désinformation

12. Les réflexions qui précèdent ont montré combien l'utilisation des technologies, et en particulier celles de l'IA, peuvent, d'une part, favoriser la désinformation mais également,

d'autre part, aider à la lutte contre celle-ci par l'analyse des contenus, de la source et des destinataires des messages. Nous avons également montré combien l'utilisation de l'IA à des fins de lutte contre la désinformation peut apparaître comme contreproductive économiquement pour des plateformes qui vivent de la publicité et du temps passé sur leurs sites. Par ailleurs, la sélection proposée par les algorithmes de l'IA peut être consciemment ou inconsciemment biaisée par la volonté de privilégier certains messages, sites ou sources, proches des opérateurs ou, à l'inverse, de contrecarrer d'autres éloignés des intérêts de ces derniers. Derrière ces algorithmes puissants, peut se cacher une véritable régulation de la liberté d'expression. Si nous pouvons accepter que la technologie rende effective la réglementation⁴⁷ et, en l'occurrence, balise, à l'instar ou plutôt conformément aux conditions posées par l'article 10 de la CEDH qui consacre la liberté d'expression, nous ne pouvons admettre qu'elle aille au-delà⁴⁸. Mettre la technologie au service d'une régulation conforme au droit et légitime est certes un autre défi. La technologie, en particulier de l'IA, apparaît comme un outil merveilleux d'effectivité de la régulation, mais cet outil ne peut être conforme et légitime que si la délégation donnée à ceux qui le mettront en place est placée sous le contrôle d'organes multidisciplinaires, suffisamment indépendants et experts. Ce contrôle n'est pas simple vu l'opacité du fonctionnement des applications d'IA et la volonté sans doute trop confiante de leur attribuer une autonomie de décision.

n° 47, pp. 178 et s.), on se référera surtout à la proposition de loi anglaise présentée au Parlement et qui attribue un rôle central à l'OFCOM: «*Online Harms White Paper: Full Government Response to the consultation Presented to Parliament by the Secretary of State for Digital, Culture, Media & Sport and the Secretary of State for the Home Department by Command of Her Majesty*», 15 décembre 2020: «Très concrètement, la 'House of Commons', qui approuve le rapport, estime qu'un Code d'éthique développé par des experts et contrôlé par un régulateur public indépendant, doit être rendu obligatoire par la loi. Ce régulateur doit pouvoir exiger toute information des entreprises soumises à son autorité, accueillir les plaintes particulières des internautes, auditer les algorithmes utilisés par les entreprises».

⁴⁶ Article 35, § 1, du DSA.

⁴⁷ Y. POULLET, «Technology and Law: from alliance to challenges», in U. GASSER (éd.), *Information Quality Regulation: Foundations, Perspectives and Applications*, Baden-Baden, 2004.

⁴⁸ À propos du risque de voir l'IA devenir un mode de régulation *ultra legem*, voy. parmi d'autres articles de l'auteure, M. HILDEBRANDT, «Law as Computation in the Era of Artificial Legal Intelligence: Speaking law to the power of statistics», in *University of Toronto Law Journal*, 2018, Volume 68, n° 1, pp. 12-35.



13. Le DSA, projet de règlement publié par la Commission européenne⁴⁹, reprend à cet égard quelques dispositions limitées à propos des systèmes d'IA utilisés par les services d'intermédiation. Il s'agit, essentiellement, d'obligations de transparence, *a priori*, dans le cadre de la publication obligatoire des termes et conditions de l'utilisation des services. L'article 12 de la réglementation proposée énonce: «Ces informations comprennent des informations sur les politiques, procédures, mesures et outils utilisés aux fins de la modération du contenu, y compris la prise de décision algorithmique et l'examen humain. Elles sont rédigées dans un langage clair et sans ambiguïté et sont mises à la disposition du public dans un format facilement accessible»⁵⁰. Il s'agit ensuite d'informations délivrées *a posteriori*, c'est-à-dire lorsque le service d'intermédiation décide de bloquer, de suspendre ou de retirer un message ou un site: l'article 15.2.c. oblige alors le service intermédiaire à fournir à l'émetteur, «le cas échéant, des informations [(claires et facilement compréhensibles et aussi précises et spécifiques que raisonnablement possible)] sur l'utilisation de moyens automatisés pour prendre la décision, y compris lorsque la décision a été prise à l'égard d'un contenu détecté ou identifié à l'aide de moyens automatisés». Notons que, lorsqu'il s'agit de plateformes en ligne, l'obligation de transparence accroît son objet quant à l'obligation de *reporting* à propos des systèmes de modération des contenus avec paramètres utilisés et finalités (article 23.1.c.), et à propos des systèmes de sélection des publicités (article 24). En matière de «*very large platforms*», le DSA prévoit en outre une obligation particulière de *risk assessment* (article 26) des systèmes de détection des contenus illégaux ou d'informations contraires

à leurs termes et conditions⁵¹, des systèmes de recommandation, de modération et de sélection des publicités (article 26.2) et un audit indépendant externe (article 28). Nous analyserons ces différentes dispositions *infra*.

CHAPITRE III: LA RÉGULATION EUROPÉENNE DE LA DÉSINFORMATION: ENTRE ATTITUDE NÉGATIVE DE CENSURE ET ATTITUDE POSITIVE DE PROMOTION D'UNE INFORMATION PLURALISTE ET CRITIQUE

14. La lutte contre la désinformation peut suivre différentes attitudes. La première est négative: elle entend intervenir dans les flux d'information, détecter et sanctionner ce qui pourrait être néfaste, mais le fait au-delà des opérations de manipulation conscientes de l'information des citoyens ou d'une partie d'entre eux. Il importe de manier avec beaucoup de précaution cette première attitude. Elle risque de contredire la liberté d'expression, surtout si elle est laissée à l'initiative d'acteurs privés dont il importe alors de souligner la responsabilité d'autant plus grande qu'ils sont les «*gatekeepers*» de nos espaces informationnels⁵². À ce titre, les États pourraient leur imposer des obligations de service universel qui porteraient par exemple sur l'adoption d'une politique de qualité de l'information, de respect de la vie privée, de protection des consommateurs, ... La seconde est positive: elle s'appuie au

⁴⁹ Digital Services Act, *op. cit.*

⁵⁰ Article 12, § 1, du DSA.

⁵¹ Article 26, § 2: «Lorsqu'elles procèdent à des évaluations des risques, les très grandes plateformes en ligne doivent tenir compte, en particulier, de la manière dont leurs systèmes de modération du contenu, leurs systèmes de recommandation et leurs systèmes de sélection et d'affichage de la publicité influent sur l'un ou l'autre des risques systémiques [...], notamment la diffusion potentiellement rapide et large de contenus illicites et d'informations incompatibles avec leurs termes et conditions».

⁵² À cet égard, *voy. supra*.

contraire sur la liberté d'expression, et en particulier des médias et sur leur déontologie, pour introduire dans l'espace public informationnel que constituent les réseaux sociaux et internet, la dimension critique, pluraliste qui contribue à la richesse du débat public et à l'intérêt général.

A. L'attitude négative

15. Deux points nous paraissent devoir être relevés d'emblée. Le premier est important. Le DSA mêle allégrement la question des contenus illégaux à celle de la désinformation. Est-ce à dire que les deux s'équivalent et méritent le même sort ? Sans doute, la frontière entre un message raciste ou de haine et des allégations ayant pour effet d'attiser le racisme et la haine n'est pas facile à tracer.

Le discours de l'ex-président Trump, suite auquel *Twitter* et *Facebook* ont décidé de bloquer son compte, peut servir d'exemple, sous réserve de ce qui sera dit plus bas à propos des conditions générales contractuelles invoquées par ces deux entreprises pour bloquer le compte de l'ex-président.

On pourrait considérer qu'il s'agit d'un message illégal, dans la mesure où il appelle à une violence ... immédiate, ajouterait la jurisprudence américaine particulièrement soucieuse de la liberté d'expression. Il n'empêche, le contenu du premier, le message illégal, même si ses contours sont parfois imprécis (par exemple, qu'est-ce qu'un message terroriste ?), est clairement sanctionné en soi par le droit ; le second, la désinformation, ne l'est que dans la mesure où l'utilisation de la technologie et du discours révèle l'intention d'éveiller ces sentiments indignes. Il nous semble qu'à défaut de ces éléments de contexte qui laissent clairement entrevoir une manipulation des faits afin d'obtenir le résultat consciemment poursuivi, il n'y a pas lieu de s'en prendre à un message isolé ou un blog personnel dont la lecture ou la

vue peuvent induire certes de tels sentiments. N'est-ce pas là le prix à payer pour garantir la liberté d'opinion, au risque d'instaurer des censures bien plus préjudiciables pour notre fonctionnement démocratique ? Bref, ce n'est pas le contenu, qui relève de l'opinion, qui rend la désinformation indésirable pour le droit, mais bien l'utilisation de méthodes qui transforme l'opinion en manipulation des opinions d'autrui. Nous reviendrons dans la suite sur cet exemple, cherchant à appliquer les dispositions du DSA.

Relevons qu'en l'occurrence, pour bloquer le compte de l'ex-président, tant *Twitter* que *Facebook* ont invoqué la violation de leurs termes et conditions générales. Cette référence à leur régulation contractuelle privée pose problème dans la mesure où ces deux réseaux sociaux occupent en Europe plus de 60% de leurs marchés respectifs et constituent donc des passages quasi obligés. Cette liberté de définir ses propres conditions générales contractuelles pourrait être contestée ou du moins réglementée dans la mesure où la position sur le marché d'une plateforme laisse aux personnes peu de choix quant au service. On retrouve ici l'argument du service universel. Une société comme *Facebook* peut-elle mettre dans ses conditions générales qu'elle n'acceptera aucun message relatif à la critique d'elle-même ? Pourrait-elle interdire tout message dénonçant la facture environnementale du numérique ou toute critique des mesures de vaccination anti-Covid ? Pourrait-elle considérer que tout message pornographique doit être censuré ? Il nous semble que leur qualité d'opérateur d'un service universel, de « *gatekeeper* » des espaces informationnels et de communication, pourrait justifier leur obligation réglementaire d'être neutres par rapport à tout message non illégal ou qui ne constitue pas au sens strict du terme une désinformation. On note en ce sens que les très grandes plateformes doivent

veiller à la nécessité d'éviter des restrictions inutiles à l'utilisation de leur service, en tenant dûment compte des effets négatifs potentiels sur les droits fondamentaux des bénéficiaires du service⁵³. Toute censure privée dans le cas de plateformes qui, dans les faits, offrent ce que nous avons qualifié de service universel, devrait être interdite.

16. Le second point relevé est l'approche proposée. La directive du 8 juin 2000 évoquait la seule responsabilité de trois types de prestataires intermédiaires: l'opérateur de simple transport (*mere conduit*), le prestataire de service de stockage (*caching*) et, enfin, celui d'hébergement (*hosting*). La distinction de ces seules trois catégories est depuis longtemps critiquée. La complexité des métiers de l'internet justifie des catégories nouvelles et des obligations nouvelles pour ces catégories professionnelles d'intermédiaires qui, grâce à l'interactivité et aux performances de l'outil, assument des rôles qui ne peuvent plus être qualifiés de simple hébergement. En particulier, on souligne les multiples fonctions assumées par les plateformes qui, du fait de leurs prestations et de leur modèle économique, ne peuvent plus être qualifiées de simples hébergeurs au vu de leur connaissance des sites hébergés et de leurs prestations dictées par le contenu de ces sites voire des messages échangés avec ces derniers. La nécessité de faire apparaître de nouvelles catégories et de mieux jauger les responsabilités des membres de chacune de ces catégories en fonction de la réalité de leur métier semblait donc évidente. À cette première réflexion, le projet européen de règlement en ajoute une seconde: au sein du métier spécifique que constitue celui des plateformes, ne faut-il pas introduire une dissymétrie dans l'attribution des responsabilités en fonction de la part de marché occupée par

certaines plateformes⁵⁴? Une telle place leur confère à la fois une puissance économique sur d'autres acteurs du net et constitue une porte d'entrée pour une grande majorité des internautes. Ce faisant, ne peut-on pas considérer que ces plateformes très grandes (*very large platforms*), selon l'expression du DSA, constituent un véritable service universel⁵⁵, qu'il soit d'accès à l'information ou à la communication avec autrui? Comme le note le considérant n° 56 du projet: «Les très grandes plateformes en ligne sont utilisées d'une manière qui influence fortement la sécurité en ligne, la formation de l'opinion publique et du discours, ainsi que le commerce en ligne. La façon dont elles conçoivent leurs services est généralement optimisée au profit de leurs modèles d'affaires souvent axés sur la publicité et peut causer des préoccupations sociétales. En l'absence d'une réglementation et d'une application efficaces, elles peuvent établir les règles du jeu, sans identifier et atténuer efficacement les risques et les préjudices sociétaux et économiques qu'elles peuvent causer».

Ainsi, le projet de règlement distingue parmi les services intermédiaires⁵⁶: les services inter-

⁵³ Considérant n° 59 du DSA.

⁵⁴ Soit, selon l'article 25, § 1, du DSA: «La présente section s'applique aux plateformes en ligne fournissant leurs services à un nombre mensuel moyen de bénéficiaires actifs du service au sein de l'Union égal ou supérieur à 45 millions».

⁵⁵ Sur l'approche des plateformes comme opérateurs d'un service universel qui doit pouvoir être régulé par l'autorité publique, voy. Y. Poullet, «Internet of the future: Achieving Transparency, Pluralism and Democracy», European Parliament, DG Internal policies of the Union, Briefing Note, IP/A/ITRE/WS-IC-2008-139, janvier 2009.

⁵⁶ Article 2, f), du DSA: «"service intermédiaire" signifie l'un des services suivants: – un service de "simple transport" qui consiste à transmettre dans un réseau de communication des informations fournies par un destinataire du service, ou à fournir un accès à un réseau de communication; – un service de "mise en cache" qui consiste à transmettre dans un réseau de communication des informations fournies par un destinataire du service, impliquant le stockage

médiaires offrant une infrastructure réseau : fournisseurs d'accès Internet, y compris les registres de noms de domaine ; les services de pur hébergement tels que le *cloud* et les services de *webhosting* ; les plateformes en ligne réunissant vendeurs et consommateurs tels que les places de marché en ligne, les magasins d'applications, les plateformes d'économie collaborative et les plateformes de médias sociaux ; et, enfin, les très grandes plateformes en ligne en raison des risques particuliers dans la diffusion de contenus illégaux et la création de préjudices collectifs comme les atteintes à la démocratie ou la discrimination de groupes sociaux. Des règles spécifiques sont prévues dès lors pour les plateformes atteignant plus de 10% des 450 millions de consommateurs en Europe. Par ailleurs, l'application du projet entend soumettre certaines obligations à tous les intermédiaires qui offrent un service dans l'Union ou un État membre de l'Union, soit parce qu'ils y ont des utilisateurs, soit parce que c'est la cible de leurs activités⁵⁷.

17. Le but de créer un espace informationnel de confiance s'obtient par diverses mesures valables cumulativement suivant la catégorie d'intermédiaires dont l'entreprise en question relève. Ainsi, tous les services intermédiaires se doivent de respecter les règles déduites des droits fondamentaux et de coopérer avec les autorités (rapports d'activités et mise en œuvre des décisions des autorités judiciaires ou administratives).

Les plateformes en ligne⁵⁸ sont tenues par des obligations supplémentaires : détecter et s'opposer à la présence en ligne de produits, de services ou de contenus, y compris par l'action de « signaleurs de confiance » (« *trusted flaggers* »)⁵⁹, accessibles aux internautes et avec lesquels les plateformes sont priées de coopérer ; on évoque, premièrement, la nécessité pour les plateformes et également, sur la base du règlement de 2019 relatif aux intermédiaires en ligne⁶⁰, les places de marché en ligne (*online market places*)⁶¹, de

⁵⁸ Article 2, h), du DSA : « "plateforme en ligne" vise tout fournisseur de service d'hébergement qui, à la demande d'un bénéficiaire du service, stocke et diffuse au public des informations, à moins que cette activité ne soit une caractéristique mineure et purement accessoire d'un autre service qui, pour des raisons objectives et techniques, ne peut être utilisé sans cet autre service, et pour autant que l'intégration de cette caractéristique à l'autre service ne soit pas un moyen de contourner l'applicabilité du présent règlement ».

⁵⁹ Cfr l'article 19, qui définit le statut et les conditions de la reconnaissance de ces « signaleurs de confiance ».

⁶⁰ Règlement (UE) 2019/1150 du Parlement européen et du Conseil du 20 juin 2019 promouvant l'équité et la transparence pour les entreprises utilisatrices de services d'intermédiation en ligne, disponible en ligne sur <https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=CELEX:32019R1150>. Ce règlement vise également les plateformes en ligne.

⁶¹ Les places de marché en ligne collectent les consentements des internautes donnés via les cookies placés par les sites membres de la place et mettent à disposition de telles informations aux sociétés qui souhaitent envoyer des publicités ciblées, et ce en fonction des paramètres définis par ces dernières. En d'autres termes, l'ouverture d'une page sur un site web qui utilise la publicité comportementale en ligne, entraîne la transmission ou la possibilité de transmission à de nombreuses entreprises connues mais souvent inconnues par l'internaute. La « publicité comportementale en ligne » (aussi dénommée « publicité programmatique ») repose dans nos pays sur deux systèmes principaux, les deux fonctionnant sur une technologie connue sous le nom de « *real time bidding* » (« offres en temps réel ») : ainsi, l'« OpenRTB » est utilisé virtuellement par toutes les grandes entreprises du secteur de la publicité et des médias et l'« Authorized Buyers », le système RTB exclusif de Google, anciennement appelé « DoubleClick Ad Exchange » (ou « AdX ») et

automatique, intermédiaire et temporaire de ces informations, dans le seul but de rendre plus efficace la transmission ultérieure des informations à d'autres destinataires, à leur demande ; – un service d'« hébergement » qui consiste à stocker des informations fournies par un destinataire du service et à sa demande ».

⁵⁷ On retrouve là les critères d'extension « *ratione loci* » repris dans le RGPD.



tracer les utilisateurs professionnels présents sur leurs sites afin d'identifier les vendeurs de biens ou services illégaux; deuxièmement, la possibilité pour les utilisateurs de contester les décisions de modération de contenus, prises par les plateformes; troisièmement, les mesures de transparence imposées aux plateformes en ligne, en particulier les algorithmes qui guident les recommandations⁶². On ajoute à toutes ces obligations mais, cette fois seulement pour les très grandes plateformes en

ligne, l'obligation d'effectuer des évaluations des risques systémiques provoqués par ou liés au fonctionnement et à l'utilisation de leurs services (article 26) et de prendre des mesures raisonnables et efficaces visant à atténuer ces risques (article 27). Elles doivent également se soumettre à des audits externes et indépendants (article 28). Dans les cas où de très grandes plateformes en ligne utiliseraient des systèmes de recommandation (article 29) ou afficheraient de la publicité en ligne sur leur interface en ligne (article 30), elles font également l'objet d'obligations de transparence spécifiques. Elles accorderont l'accès à des données aux chercheurs pour comprendre comment évolue le risque en ligne; elles nomment un ou plusieurs agents de conformité pour assurer le respect des obligations énoncées dans le règlement (article 32) et se voient imposer des obligations spécifiques et supplémentaires en matière de rapports sur la transparence (article 33). Enfin, une structure pan-européenne est créée pour aider les États membres à jouer leur rôle de contrôle et de mise en œuvre du règlement.

récemment renommé «Authorized Buyers» (acheteurs autorisés). Ces deux systèmes permettent de produire des publicités personnalisées et ciblées sur les sites internet. Sur ces places de marché et les questions de vie privée posées par leur existence, voy. l'excellent rapport de l'ICO (l'APD du Royaume Uni), *Update report into adtech and real time bidding*, 20 juin 2019, accessible à l'adresse: <https://ico.org.uk/media/about-the-ico/documents/2615156/adtech-real-time-bidding-report-201906-dl191220.pdf>.

⁶² Cfr à ce sujet les Lignes directrices adoptées par la Commission, le 4 décembre 2020 en ce qui concerne les mécanismes de *ranking*: «*Guidelines on ranking transparency pursuant to Regulation (EU) 2019/1150*», accessibles à l'adresse: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020XC1208%2801%29>. On rappelle le contenu de l'article 5 du règlement de 2019 cité *supra*, note 60: «1. Les fournisseurs de services d'intermédiation en ligne indiquent dans leurs conditions générales les principaux paramètres déterminant le classement, et les raisons justifiant l'importance relative de ces principaux paramètres par rapport aux autres paramètres. 2. Les fournisseurs de moteurs de recherche en ligne indiquent les principaux paramètres qui, individuellement ou collectivement, sont les plus importants pour déterminer le classement ainsi que l'importance relative de ces principaux paramètres, en fournissant une description facilement et publiquement accessible, énoncée dans une formulation claire et compréhensible, sur les moteurs de recherche en ligne de ces fournisseurs. Ils tiennent cette description à jour. 3. Lorsque les principaux paramètres incluent la possibilité d'influer sur le classement contre toute rémunération directe ou indirecte versée par les entreprises utilisatrices ou les utilisateurs de sites internet d'entreprise au fournisseur concerné, ce fournisseur présente également une description de ces possibilités et des effets de cette rémunération sur le classement, conformément aux exigences énoncées aux paragraphes 1 et 2. [...]».

18. Les articles 6 et 7 constituent des principes fondamentaux valables pour tous les services intermédiaires. L'article 6 refuse de dégager les intermédiaires de leurs responsabilités sur la seule base du fait qu'ils aient implémenté d'initiative des systèmes en particulier d'IA pour veiller à détecter, repérer et supprimer des contenus illicites, ou à en rendre l'accès impossible, ou qu'ils prennent les mesures nécessaires pour se conformer aux exigences du droit de l'Union. L'article 7 n'oblige en aucune manière les «fournisseurs de services intermédiaires à surveiller les informations qu'ils transmettent ou stockent, ou de rechercher activement des faits ou des circonstances révélant des activités illicites». Ceci dit, l'utilisation de tout système de détection, y compris utilisant des technologies d'IA,



doit faire l'objet d'une information claire dans les conditions générales contractuelles⁶³. Si ce système conduit à un retrait d'informations ou à un blocage d'accès au compte, l'intermédiaire se doit, selon l'article 15, d'informer la personne qui fait l'objet de la mesure, de fournir «un exposé clair et spécifique des motifs de cette décision» et, en particulier, si tel est le cas, de l'utilisation de systèmes automatiques de détection. On ajoute que toute plateforme, peu importe sa taille, doit permettre l'accès facile à la personne incriminée à l'organe de règlement des litiges extrajudiciaires laissé au choix de ce dernier, ceci sans exclure *in fine* le recours judiciaire. Les obligations renforcées des très grandes plateformes sont justifiées par le fait que leur importance sur le marché, jointe aux technologies qu'elles utilisent, créent des risques qualifiés de «systémiques». Le considérant n° 57 du DSA décrit ces fameux risques. Le premier risque concerne l'ampleur de la diffusion de contenus illicites permis par les plateformes en ligne occupant une place importante sur le marché. On retrouve ici le cas cité (*supra*, n° 15) du message de l'ex-président Trump relatif à la «montée au Capitole»⁶⁴. Si, en effet, le

message de ce dernier était illégal, ce qui reste à démontrer, il est clair que sa diffusion par les plateformes a pu recevoir une portée sans commune mesure avec celle qu'aurait la parole d'un humble citoyen⁶⁵. Le deuxième risque concerne «l'incidence du service sur l'exercice des droits fondamentaux, tels que protégés par la Charte des droits fondamentaux, y compris la liberté d'expression et d'information, le droit à la vie privée, le droit à la non-discrimination et les droits de l'enfant. De tels risques peuvent découler, par exemple, de la conception des systèmes algorithmiques utilisés par la très grande plateforme en ligne ou de l'usage abusif de ses services par la soumission de notifications abusives ou d'autres méthodes visant à empêcher la liberté d'expression ou à entraver la concurrence». Le troisième risque vise l'utilisation des mécanismes mis en place par la plateforme, comme le système de recommandations, afin de manipuler autrui lors d'élections, diffuser des messages intentionnellement erronés qui mettent en danger la santé publique, la démocratie, etc.⁶⁶. Dans

⁶³ Cfr l'article 12 relatif aux conditions générales contractuelles: «Ces renseignements ont trait, notamment, aux politiques, procédures, mesures et outils utilisés à des fins de modération des contenus, y compris la prise de décision fondée sur des algorithmes et le réexamen par un être humain. Ils sont énoncés clairement et sans ambiguïté et sont publiquement disponibles dans un format facilement accessible». L'article 23, lorsque l'intermédiaire est une plateforme en ligne, oblige ce dernier à indiquer dans le rapport annuel, qu'il doit publier «tout recours à des moyens automatisés à des fins de modération de contenus, y compris une spécification des objectifs précis, des indicateurs de la précision des moyens automatisés pour atteindre ces objectifs et des éventuelles mesures de sauvegarde appliquées».

⁶⁴ Rappelons que nous émettons (*supra*, n° 15) des doutes quant à la validité de conditions générales contractuelles qui, dans le cas de très grandes plateformes, constitueraient une véritable censure privée.

⁶⁵ Cfr à ce propos, cet extrait du considérant n° 57: «Par exemple, et sans préjudice de la responsabilité personnelle du bénéficiaire du service de très grandes plateformes en ligne du fait de l'éventuelle illégalité de son activité au regard du droit applicable, cette diffusion ou ces activités peuvent constituer un risque systémique important lorsque l'accès à ce contenu peut être amplifié par l'intermédiaire de comptes ayant une audience particulièrement étendue».

⁶⁶ Cfr toujours le considérant n° 57: «Une troisième catégorie de risques concerne la manipulation intentionnelle et, souvent, coordonnée du service de la plateforme, avec un impact prévisible sur la santé, le discours civique, les processus électoraux, la sécurité publique et la protection des mineurs, compte tenu de la nécessité de sauvegarder l'ordre public, de protéger la vie privée et de lutter contre les pratiques commerciales frauduleuses et trompeuses. De tels risques peuvent survenir, par exemple, par la création de faux comptes, l'utilisation de bots et d'autres comportements automatisés ou partiellement automatisés, ce qui peut conduire à la diffusion rapide et généralisée d'informations illégales ou incompatibles avec les conditions générales d'une plateforme en ligne».



le cas du message « trumpiste », sans doute, cela a-t-il été également le cas ? On peut considérer, encore faudrait-il analyser les termes du message, qu'il remettait en cause l'ordre public et il est fort à parier que le relais du message de l'ex-président a dû bénéficier de la recommandation générée automatiquement par un système d'IA et que ce dernier savait que la diffusion de son message en profiterait. Si donc, aux yeux du droit européen, les plateformes pouvaient bloquer le message présidentiel, en tant que contenu illégal ou source de désinformation, il leur incombait non seulement d'en informer le titulaire du compte mais, en outre, de le renseigner, dans un langage clair et accessible, sur les moyens de contester la décision et, le cas échéant, d'agir devant les tribunaux.

19. Ce sont ces risques systémiques qui expliquent la *responsabilité soci(ét)ale*⁶⁷ de ces plateformes de grande ampleur et justifient les obligations supplémentaires mises à leur charge. Les articles 26 et 27 du DSA requièrent de telles plateformes des mesures d'évaluation de ces risques⁶⁸ et, le cas échéant leur atténua-

tion⁶⁹. Ces mesures, au rang desquelles figurent bien évidemment les systèmes d'IA utilisés, doivent faire l'objet d'un audit externe par un organisme indépendant. Cet audit doit, selon le vœu des auteurs du texte, s'opérer avec la participation de représentants des bénéficiaires du service, de représentants de groupes potentiellement touchés par leurs services, d'experts indépendants et d'organisations de la société civile. Ses recommandations devront être prises en compte⁷⁰ et, selon

-
- (a) la diffusion de contenus illicites par l'intermédiaire de leurs services ;
 - (b) tout effet négatif pour l'exercice des droits fondamentaux relatifs au respect de la vie privée et familiale, à la liberté d'expression et d'information, à l'interdiction de la discrimination et aux droits de l'enfant, tels que consacrés aux articles 7, 11, 21 et 24 de la Charte, respectivement ;
 - (c) la manipulation intentionnelle de leur service, y compris via l'utilisation non authentique ou l'exploitation automatisée de leur service, avec un effet négatif avéré ou prévisible sur la protection de la santé publique, des mineurs, du discours civique, ou des effets avérés ou prévisibles en lien avec les processus électoraux et la sécurité publique.

2. Lorsqu'elles procèdent à des évaluations des risques, les très grandes plateformes en ligne tiennent notamment compte de la manière dont leurs systèmes de modération des contenus, systèmes de recommandation et systèmes de sélection et d'affichage de la publicité influencent tout risque systémique visé au paragraphe 1, y compris la diffusion potentiellement rapide et à grande échelle de contenus illicites et d'informations incompatibles avec leurs conditions générales».

⁶⁹ Article 27: « Les très grandes plateformes en ligne mettent en place des mesures d'atténuation raisonnables, proportionnées et efficaces, adaptées aux risques systémiques spécifiques identifiés en application de l'article 26. [...] ».

⁷⁰ Article 28, § 4: « Les très grandes plateformes en ligne qui reçoivent un rapport d'audit qui n'est pas positif tiennent dûment compte de toute recommandation opérationnelle qui leur est adressée en vue de prendre les mesures nécessaires à leur mise en œuvre. Dans le mois suivant la réception de ces recommandations, elles adoptent un rapport de mise en œuvre des recommandations d'audit énonçant ces mesures. Si elles ne mettent pas en œuvre les recommandations opérationnelles, elles en fournissent les motifs dans

⁶⁷ La responsabilité sociétale des entreprises (RSE) est définie dans le cadre de référence fédéral comme « un processus permanent d'amélioration dans le cadre duquel les entreprises intègrent de manière volontaire, systématique et cohérente des considérations d'ordre social, environnemental et économique dans la gestion globale de l'entreprise ; à cet égard, la concertation avec les parties prenantes de l'entreprise fait partie intégrante du processus », définition reprise par le ministère des Affaires économiques sur le site: <https://economie.fgov.be/fr/themes/entreprises/developper-et-gerer-une/>.

⁶⁸ Article 26: « 1. Les très grandes plateformes en ligne recensent, analysent et évaluent, à compter de la date d'application visée au second alinéa de l'article 25, paragraphe 4, puis au moins une fois par an, tout risque systémique important trouvant son origine dans le fonctionnement et l'utilisation faite de leurs services au sein de l'Union. Cette évaluation des risques est spécifique à leurs services et comprend les risques systémiques suivants :



l'article 33, publiées semestriellement par les plateformes elles-mêmes avec leur propre évaluation des risques et les mesures d'atténuation prises. En particulier, des précautions doivent être prises en matière d'utilisation des systèmes automatiques de modération de contenu⁷¹. Bien que la liberté contractuelle des prestataires de services intermédiaires mérite d'être rappelée⁷², il convient de limiter cette liberté en établissant certaines règles d'ordre public sur le contenu, et sur l'application des conditions de ces fournisseurs, dans l'intérêt de la transparence, de la protection des bénéficiaires du service et de l'évitement des résultats injustes ou arbitraires⁷³. Les mêmes précautions doivent être prises en ce qui concerne la sécurité du système contre l'utilisation des plateformes pour un envoi massif de messages. En ce qui concerne les systèmes de recommandation, la même précaution est prise : « La manière dont les informations sont hiérarchisées et présentées sur l'interface en ligne d'une très grande plateforme afin de faciliter et d'optimiser l'accès aux informations pour les

bénéficiaires du service revêt une importance capitale pour les activités de la plateforme. Cela consiste, par exemple, à suggérer, classer et hiérarchiser les informations de manière algorithmique, en les distinguant par le texte ou par d'autres représentations visuelles, ou en organisant de toute autre manière les informations fournies par les bénéficiaires. Ces systèmes de recommandation peuvent avoir une incidence significative sur la capacité des bénéficiaires à récupérer les informations en ligne et à interagir avec elles. Ils jouent également un rôle important dans l'amplification de certains messages, la diffusion virale de l'information et la stimulation du comportement en ligne. Par conséquent, les très grandes plateformes en ligne devraient garantir que les bénéficiaires sont informés de manière appropriée et peuvent influencer les informations qui leur sont présentées. Elles devraient présenter clairement les principaux paramètres de ces systèmes de recommandation d'une manière facilement compréhensible afin que les bénéficiaires comprennent comment l'information est hiérarchisée à leur intention. Elles devraient également veiller à ce que les bénéficiaires puissent avoir d'autres options concernant les principaux paramètres, notamment des options qui ne relèvent pas du profilage du bénéficiaire »⁷⁴.

Dernier point, plus contestable au regard du nécessaire respect de la liberté d'expression : le texte laisse la possibilité d'intervention pour la Commission européenne en cas de crise au sens large « comme des tremblements de terre, des ouragans, des pandémies et d'autres menaces transfrontalières graves à la santé publique [sans doute la lutte contre la pandémie actuelle justifie-t-elle cet ajout], à la guerre et aux actes de terrorisme »⁷⁵. Ainsi,

le rapport de mise en œuvre des recommandations d'audit et exposent les mesures alternatives qu'elles auraient éventuellement prises pour résoudre tout cas de manquement recensé».

⁷¹ Article 2, p) : « "modération des contenus" vise les activités entreprises par les fournisseurs de services intermédiaires destinées à détecter et à repérer les contenus illicites ou les informations incompatibles avec leurs conditions générales, fournis par les bénéficiaires du service, et à lutter contre ces contenus ou informations, y compris les mesures prises qui ont une incidence sur la disponibilité, la visibilité et l'accessibilité de ces contenus illicites ou informations, telles que leur rétrogradation, leur retrait ou le fait de les rendre inaccessibles, ou sur la capacité du bénéficiaire à fournir ces informations, telles que la suppression ou la suspension du compte d'un utilisateur ».

⁷² Rappelons nos doutes quant à la possibilité dans le cas de très grandes plateformes d'imposer par là une censure privée. Voy. à cet égard, le considérant n° 59 du DSA et notre réflexion sur la qualification de service universel de communication et d'information rappelés *supra*, n° 15 *in fine*.

⁷³ Considérant n° 38 du DSA.

⁷⁴ Considérant n° 62.

⁷⁵ Voy. considérant n° 71 : « [temps de crise] où, par exemple, les plateformes en ligne peuvent être utilisées à mauvais escient pour la diffusion rapide de



la Commission peut élaborer, en concertation ou non avec les autorités des États membres et les organisations de la société civile, un protocole de crise⁷⁶ qui comprendra l'obligation d'afficher les principales informations relatives à la situation de crise et le renforcement du contrôle du respect des obligations mises à charge des intermédiaires. La Commission veillera à limiter dans le temps ces mesures, à respecter les libertés et à assurer la transparence des mesures prises. La Commission justifie cette intervention précisément par les risques systémiques liés à l'existence des plateformes. On note que ces interventions de l'autorité publique suivront l'initiative de la Commission qui peut lancer l'élaboration de protocoles de crise pour coordonner une réponse rapide, collective et transfrontalière dans l'environnement en ligne.

B. L'attitude positive

20. La Résolution du Parlement européen a certainement été marquée par la multiplication des désinformations qui ont circulé à propos de la pandémie Covid-19 et qui ont exigé une réaction rapide des autorités européennes⁷⁷.

contenus illégaux ou de désinformation ou lorsque le besoin se fait sentir d'une diffusion rapide d'informations fiables. Compte tenu du rôle important des très grandes plateformes en ligne dans la diffusion de l'information dans nos sociétés et au-delà des frontières, ces plateformes devraient être encouragées dans l'élaboration et l'application de protocoles de crise spécifiques. Ces protocoles de crise ne devraient être activés que pour une période limitée et les mesures adoptées devraient également se limiter à ce qui est strictement nécessaire pour remédier aux circonstances extraordinaires».

⁷⁶ Article 37.

⁷⁷ Ainsi, dès le 10 juin 2020, la Communication conjointe de la Commission et du Haut représentant de l'Union pour les affaires étrangères et la politique de sécurité au Parlement européen, au Conseil européen, au Conseil, au Comité économique et social européen et au Comité des Régions, intitulée «Lutter contre la désinformation concernant la Covid-19 – Démêler le vrai du faux», JOIN(2020) 8 final, disponible en

Un rapport spécial du 24 avril 2020 de l'EEAS⁷⁸ liste d'ailleurs les faits de désinformation significatifs à ce propos, qu'il s'agisse de montages vidéos, d'articles de presse, d'attaques y compris physiques contre des lieux publics ou des journalistes, de messages adressés dans les réseaux sociaux, ce que le rapport qualifie d'«infodémie». Depuis⁷⁹, en la matière, se sont ajoutés bien d'autres faits explicables par des raisons stratégiques ou économiques. Cette «infodémie» brouille considérablement les messages officiels des instances nationales, européennes et internationales et contrecarre les initiatives qu'elles souhaitent prendre. Il est clair qu'au-delà du débat Covid-19, bien d'autres débats, tels que ceux relatifs à la 5G ou aux élections américaines, rendent la question de la désinformation plus que prégnante pour l'avenir de nos démocraties. Cette Résolution est importante car elle développe des pistes plus positives qui, au lieu de limiter la liberté d'expression, trouvent dans cette dernière un moyen de lutter contre la désinformation. Sans doute, n'est-elle pas la seule dans le concert des textes européens, mais elle en constitue l'expression la plus articulée. La réaffirmation de la liberté d'expression comme pilier de la

ligne sur <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020JC0008>.

⁷⁸ Special Report update of the European External Action Service (EEAS), 14 avril 2020, «*Short Assessment of Narratives and Disinformation around the COVID-19/ Coronavirus Pandemic*», accessible en ligne sur <https://euvsdisinfo.eu/eeas-special-report-update-2-22-april/>.

⁷⁹ Résolution, point B: «considérant que la crise a mis en lumière le rôle essentiel des journalistes, qui fournissent aux citoyens des informations fiables et vérifiées; qu'il convient ainsi de redoubler d'efforts pour leur assurer un environnement de travail sûr, dans des conditions adéquates; que, dans le contexte de la lutte contre la corruption et la mauvaise administration dans l'Union européenne, le journalisme d'investigation devrait bénéficier d'une attention particulière».



démocratie européenne est centrale dans le texte parlementaire⁸⁰.

21. La première idée défendue est la réaffirmation du rôle de quatrième pouvoir et donc le rôle essentiel au service de la démocratie, que doivent jouer, en particulier⁸¹, les organes de presse⁸² en tant qu'organismes

d'édition, responsables d'un contenu, pluralistes, indépendants et rassemblant des journalistes habités par une forte déontologie. Le Parlement souligne les dangers d'une presse de plus en plus contrôlée par certains pouvoirs privés et, tantôt réduite au silence par la menace externe voire par les gouvernements, tantôt devenue un outil de propagande pour l'action gouvernementale. Elle décrit la paupérisation d'une presse privée des revenus publicitaires et soumise aux impératifs d'une opinion qui souhaite sensationnalisme et réaction immédiate⁸³. Par ailleurs, elle note qu'un journalisme de haute qualité déontologique et donc critique par rapport à ses sources « est primordial pour la démocratie »⁸⁴. Le Parlement plaide pour un refinancement du secteur des médias⁸⁵ et conclut à ce propos : « les mesures de lutte contre la désinformation devraient viser à favoriser la pluralité des opinions, grâce à la promotion du journalisme de haute qualité, qui fournit des informations fiables, fondées sur les faits et vérifiées »⁸⁶. Le renforcement de la protection des médias dignes de confiance par l'apposition de labels de qualité délivrés par la profession elle-même semble être une piste qui, en accroissant la visibilité pour le simple

⁸⁰ Résolution, point 6 : « [...] le futur mécanisme de l'Union sur la démocratie, l'État de droit et les droits fondamentaux doit consacrer la liberté des médias, y compris la liberté artistique, comme un pilier essentiel d'un régime démocratique; invite la Commission, dans ce contexte, à recueillir des informations et des statistiques sur la liberté et le pluralisme des médias dans tous les États membres ».

⁸¹ La jurisprudence de la Cour eur. D.H. tend à conférer, dans une approche fonctionnelle, l'octroi de l'ensemble des droits et libertés liés à l'article 10 à toute personne, peu importe la profession de journaliste. Sur cette jurisprudence, voy. Th. LEONARD et Y. Poullet, « L'intérêt général comme arbitre du débat entre vie privée et liberté d'expression », in Y. Poullet (éd.), *Vie privée, liberté d'expression et démocratie dans la société numérique*, Cahier du CRIDS, n° 47, Bruxelles, Larcier, 2020, pp. 79 et s., en particulier les analyses des arrêts récents dans les affaires *Satamedia* et *Buyvids*.

⁸² Il s'agit bien ici des organismes de presse et non des messages que chacun de nous postent sur les réseaux sociaux, parfois qualifiés de presse électronique. Certes, eux aussi dans la mesure où ils sont la manifestation de notre liberté d'expression, contribuent au débat public mais ne peuvent être qualifiés de presse faute d'un environnement qui garantit la qualité du propos et dès lors ne peuvent bénéficier de la protection particulière que la Constitution et les lois accordent à une presse par ailleurs responsable. Sur la tendance à accorder aux expressions d'intérêt général des citoyens sur Internet, les mêmes protections qu'à la presse, voy. J. ENGLEBERT, « La liberté d'expression à l'heure de l'internet », in Y. Poullet (éd.), *Vie privée, liberté d'expression et démocratie dans la société numérique*, Cahier du CRIDS, n° 47, Bruxelles, Larcier, 2020, pp. 123 et s.). On connaît l'arrêt *Castells* (Cour eur. D.H., 23 avril 1992, série A, n° 236, § 43) : « Il ne faut pas oublier le rôle éminent de la presse dans un État de droit. [...] La liberté de la presse fournit aux citoyens l'un des meilleurs moyens de connaître et juger les idées et attitudes de leurs dirigeants. Elle donne en particulier aux hommes politiques l'occasion de refléter et commenter les soucis de l'opinion publique. Elle permet à chacun de participer au libre jeu du débat politique qui se trouve au cœur même de la notion de société démocratique ». Voy. aussi, les

arrêts plus anciens *Sunday Times et Garden & Observer*, de 1991, considérant la presse comme chien de garde d'une démocratie.

⁸³ Résolution, point 7 : « souligne que le rôle des médias du service public est irremplaçable et insiste sur le fait qu'il est essentiel de garantir et de maintenir leur indépendance vis-à-vis de toute ingérence politique; souligne, en outre, également la nécessité de garantir l'indépendance financière des opérateurs privés et les conditions de viabilité de leurs activités afin d'éviter la captation de médias; réitère, dans ce contexte, l'appel du Parlement à mettre en place un ambitieux plan d'action de l'Union en faveur des médias; condamne les tentatives des gouvernements de certains États membres de réduire au silence les médias critiques et indépendants, ainsi que de restreindre la liberté et le pluralisme des médias ».

⁸⁴ Résolution, point 7.

⁸⁵ Résolution, points 3 et 21.

⁸⁶ Résolution, point 34.



citoyen des canaux d'information fiables dans la manière dont l'allégation est contrôlée avant de devenir, le cas échéant, information⁸⁷ et d'être alors seulement publiée⁸⁸, permettraient en contrepoint de démasquer les désinformations. Dans ce contexte, il est intéressant de noter la volonté des autorités européennes de mettre elles-mêmes sur pied des plateformes d'informations, afin d'«assurer une communication renforcée et proactive dans toutes les langues officielles en cas de situations d'urgence publique majeure, telles que la pandémie, afin de veiller à ce que les citoyens européens aient accès à des informations exactes, faciles à exploiter et vérifiées»⁸⁹.

Le pluralisme de la presse est fondamental. Le Parlement en appelle à un plan d'action européen pour le développement d'un paysage des médias, «dynamique et pluraliste»⁹⁰. Pour assurer cette liberté des journalistes d'agir sans crainte de représailles voire de violences, le Parlement plaide pour l'adoption de règles assurant la protection des journalistes et des «lanceurs d'alerte», protection y compris physique mais également contre le harcèlement et les représailles de quelque nature qu'elles soient⁹¹.

22. Au-delà des réflexions qui appuient les développements repris ci-dessus (*supra*, n°s 17 et s.) à propos de la responsabilité des plateformes, en particulier des plateformes «très larges» et des mesures qu'elles doivent prendre pour contrer la désinformation, le Parlement souligne, sans doute de manière trop discrète, le rôle essentiel que ces plateformes jouent dans la discussion et les débats publics essentiels à notre démocratie⁹². Surtout, il plaide pour une collaboration avec les laboratoires de recherche et les communautés de *fact checkers* pour améliorer le fonctionnement de ces plate-

par la prévention ou en insistant pour que de réelles enquêtes soient menées lorsque des violations sont commises; souligne que les femmes journalistes sont particulièrement exposées au harcèlement et aux intimidations et qu'elles doivent donc faire l'objet de protections supplémentaires; fait part de sa vive préoccupation au sujet de l'augmentation des attaques contre les femmes journalistes et professionnelles des médias; invite une nouvelle fois les États membres à adopter une approche qui tient compte des questions d'égalité des sexes lorsqu'ils envisagent des mesures visant à assurer la sécurité des journalistes». *Cfr* sur la question de la protection y compris physique des journalistes et les intimidations opérées auprès de ces derniers ou des ONG, le rapport de l'*Academic Network on European* (P. BARD *et alii*), *Citizenship Rights, «Ad hoc Request SLAPP in the EU context»*, 29 mai 2020, accessible sur le site: https://ec.europa.eu/info/sites/info/files/ad-hoc-literature-review-analysis-key-elements-slapp_en.pdf. SLAPP est l'acronyme utilisé par la Commission européenne pour: «*Strategic Lawsuits Against Public Participation*», soit «des poursuites judiciaires sans fondement ou exagérées et d'autres formes légales d'intimidation lancées par des organes de l'État, des sociétés commerciales et des individus au pouvoir contre des parties plus faibles – journalistes, organisations de la société civile, défenseurs des droits de l'homme et autres – qui expriment des critiques ou transmettent des messages gênants pour les puissants, sur une question publique».

⁸⁷ Pour reprendre la distinction proposée par V. de Coorebyter, citée en note 5.

⁸⁸ Des initiatives existent en ce sens. Ainsi, on distinguera des labels comme celui proposé par «*Journalism Trust Initiative*» des reporters dits sans frontières qui se concentre sur les méthodes de production de l'information ou celui, finlandais, qui consiste à adhérer à une déontologie, dont le respect est contrôlé par le conseil de déontologie journalistique. Sur ce point, voy. M. HANOT, «Un conseil de déontologie pour restaurer la confiance entre médias d'information et public», *Rapport annuel du Conseil de déontologie journalistique*, 2016-2017, pp. 7 et 8.

⁸⁹ Résolution, point 50.

⁹⁰ Résolution, point 20.

⁹¹ Résolution, point 11: «[...] demande la mise en œuvre des orientations de l'Union relatives à la liberté d'expression en ligne et hors ligne, qui précisent que l'Union prendra toutes les mesures nécessaires pour assurer la protection des journalistes, que ce soit

⁹² Point 35: «[...] rappelle que les plateformes en ligne font partie de la sphère publique en ligne, où s'organise le débat public; demande à la Commission de veiller à ce que les plateformes respectent les droits fondamentaux et la liberté d'expression».

formes dans leur détection et leur analyse des dommages créés par la désinformation⁹³.

23. L'action, en faveur de ce que les textes européens appellent la «*media and information literacy*»⁹⁴, est également remarquable. Il s'agit d'introduire dès les premières étapes des programmes d'éducation des cours, mais également tout au long de leurs formations, voire de la vie, des enseignements qui

permettent d'aiguiser l'analyse critique par les internautes des messages présents sur le net. Le Parlement insiste sur la nécessité d'élargir cette formation aux valeurs démocratiques et aux libertés individuelles et sur le rôle que doivent jouer les organisations médiatiques locales et communautaires. Les plateformes sont invitées à participer à cet effort.

CONCLUSIONS

24. Lors de ses récentes conclusions relatives à la sauvegarde de médias libres et pluralistes, le Conseil des ministres de l'Union européenne soulignait avec raison: «*with regard to the importance of freedom of speech, states and administrative regulatory authorities as well as private platform providers should abstain from defining quality content or the reliability of content itself*»⁹⁵. Ce primat de la liberté de la presse est à rappeler au moment où les graves atteintes à la démocratie orchestrées lors de votes, et plus récemment la lutte contre la Covid-19, ont mis en évidence les dangers de la désinformation. Au nom de la vérité et de l'urgence de l'intérêt public, hier celui de la démocratie, aujourd'hui, celui de la santé publique, on conçoit la volonté de certains de combattre l'hérésie et de mêler les notions de «*misinformation*» et de «*disinformation*». Il ne peut être question de les confondre. La liberté d'expression – si elle peut connaître des limites: l'appel à la violence, les messages racistes, la pédopornographie – s'entend, comme la Cour européenne des droits de l'homme le rappelle

⁹³ Point 40: «se félicite du lancement du projet "Observatoire européen des médias numériques (EDMO)", qui améliorera les connaissances scientifiques disponibles en matière de désinformation en ligne, favorisera la mise en place d'un marché des services de vérification de l'authenticité des informations dans l'Union, soutiendra la création d'une communauté transfrontière et multidisciplinaire composée de vérificateurs de faits et de chercheurs universitaires qui collaboreront avec les parties intéressées pour déceler, analyser et exposer les menaces potentielles en matière de désinformation, y compris dans le domaine de la Covid-19, par exemple».

⁹⁴ Cfr points 46 et s.: «invite la Commission et les États membres à redoubler d'efforts pour renforcer les politiques en matière d'éducation qui promeuvent l'éducation aux médias et à l'information, donnent aux citoyens les moyens de réfléchir de manière critique et les aident à détecter la désinformation; souligne, à cet égard, que le maintien de l'indépendance éditoriale au sein des médias nationaux et locaux et la mise en place de projets d'éducation aux médias sont essentiels pour le renforcement de la résilience, de la sensibilisation et de l'éducation, afin de lutter efficacement contre la propagande, la désinformation et les manipulations; estime qu'un programme et des efforts continus d'éducation aux médias dans toutes les catégories d'âge sont de la plus haute importance pour renforcer la résilience de la société face à ces menaces pour l'espace numérique; invite, à cet égard, à travailler en étroite coopération avec les États membres et les organisations de la société civile afin d'élaborer des programmes d'enseignement sur l'éducation à l'information, aux médias et aux données; observe que l'éducation aux médias est une compétence de plus en plus essentielle et critique pour les citoyens; fait remarquer qu'afin de toucher un public plus large et autant de tranches d'âge que possible, il est important d'intensifier les initiatives d'éducation aux médias par l'intermédiaire des plateformes de médias sociaux, y compris les stratégies efficaces d'éducation aux médias destinées aux personnes âgées et aux groupes les plus vulnérables [...]».

⁹⁵ Council of the European Union, «*Council conclusions on safeguarding a free and pluralistic media system*», 27 novembre 2020, n° 37, accessible sur le site: <https://data.consilium.europa.eu/doc/document/ST-13260-2020-INIT/en/pdf>.



DOCTRINE

depuis l'arrêt *Handyside*⁹⁶, de l'acceptation d'opinions, fussent-elles provocatrices, dérangeantes voire fausses.

Ce que, par contre, on peut concevoir, c'est que l'utilisation consciente des outils technologiques permette la manipulation voulue des opinions, non pour ouvrir le débat public, mais dans la poursuite d'une finalité mettant en danger l'intérêt public. Il revient d'abord à ceux qui constituent les portes d'entrée de nos espaces publics d'information et de communication de mettre en place, d'une part, des systèmes qui ne favorisent pas ce type de comportement et, d'autre part, de détecter les manœuvres qu'en aval prennent un certain nombre d'acteurs en s'appuyant sur des outils technologiques de déformation ou de propagation de la fausse information. Sur ces deux points, il était utile que les autorités publiques interviennent. L'autorégulation ou la corégulation descendante ne peuvent suffire là où les exigences de la lutte contre la désinformation vont à l'encontre du modèle d'affaires

sur lequel repose le fonctionnement des plateformes. Sans doute, faudra-t-il inventer de nouvelles manières pour ces plateformes de rentabiliser leurs investissements? C'est ainsi que les autorités de l'Union européenne justifient les nombreuses obligations d'audit indépendant des systèmes, en particulier d'IA, utilisés par les plateformes, et les précautions prises vis-à-vis des décisions que les systèmes de modération pourraient prendre et ce, afin d'assurer le droit de chacun de contester les «vérités» sorties de l'ordinateur.

Cette solution nous apparaît insuffisante. Le rôle de l'État n'est-il pas de lutter contre la désinformation, non en s'arrogeant les pouvoirs de censure mais, au contraire, en promouvant des médias pluralistes capables d'une analyse critique de l'information et ce par des acteurs respectueux d'une déontologie? Le rôle de l'État n'est-il pas de promouvoir auprès de chacun une éducation au regard critique? La liberté d'expression et la vitalité de nos démocraties sont à ce prix.

⁹⁶ Ainsi que l'a relevé la Cour dans l'arrêt *Handyside*, la liberté d'expression constitue l'un des fondements essentiels d'une société démocratique. Sous réserve du paragraphe 2 de l'article 10 (art. 10-2), elle vaut «non seulement pour les informations ou idées accueillies avec faveur ou considérées comme inoffensives ou indifférentes, mais aussi pour celles qui heurtent, choquent ou inquiètent l'État ou une fraction quelconque de la population» (arrêt *Handyside* c. Royaume-Uni, 7 décembre 1976, p. 23, par. 43).

