Etude

Les désordres de l'information : quels enjeux pour la liberté d'expression, le vivre ensemble et la démocratie ?

Jean-François Furnémont, Président du CLARA

1. Introduction

Le 24 février 2017, un site de fake news enregistré aux Etats-Unis a cloné le site du quotidien Le *Soir* pour faire de fausses allégations concernant le candidat à l'élection présidentielle Emmanuel Macron. Cette histoire a été partagée par Marion Maréchal-Le Pen, nièce de la candidate présidentielle du Front national Marine Le Pen, affirmant que la campagne électorale d'Emmanuel Macron était soutenue financièrement par l'Arabie Saoudite. Le faux site internet semblait très plausible et était même équipé d'hyperliens menant au site internet original du journal[1].

Les nouvelles fausses et déformées ne sont pas un phénomène nouveau, tout comme les théories du complot, les rumeurs et la propagande. Les politiciens ont toujours utilisé tous les moyens disponibles pour gagner des votes pendant les campagnes électorales. Et les médias ont toujours partagé des nouvelles exagérées avec des titres choquants, parce que c'est ce qui attire l'attention et vend de la publicité. Ces phénomènes se sont développés bien avant Internet et les médias sociaux, en même temps que les nouvelles ont commencé à circuler largement - avec l'invention de l'imprimerie en 1439 - beaucoup plus longtemps que les nouvelles vérifiées, « objectives », qui ont émergé en force il y a un peu plus d'un siècle. Dès le début, les fausses nouvelles ont eu tendance à être sensationnalistes et extrêmes, conçues pour enflammer les passions et les préjugés[2]. La motivation était la même que pour les créateurs de fausses nouvelles aujourd'hui : influence, pouvoir et gain financier.

L'environnement moderne de l'information a rendu possibles des méthodes entièrement nouvelles de production, de diffusion et d'amplification de la désinformation. Jamais auparavant il n'a été aussi facile de produire de la désinformation, tout en ayant la possibilité de rester complètement anonyme et donc libre de toute responsabilité. La vitesse à laquelle elle se répand, grâce aux nouvelles technologies, et l'ampleur de son influence, grâce à sa présence durable en ligne et à ses possibilités de partage, la placent dans une catégorie totalement différente de ses prédécesseurs.

En outre, les médias eux-mêmes, à la recherche de publicité, sont devenus de plus en plus dépendants des réseaux sociaux pour diffuser leur contenu et atteindre le public. Les plateformes en ligne, telles que les réseaux sociaux, les moteurs de recherche, les agrégateurs d'informations, les applications de messagerie ou les plateformes de partage de vidéos, sont devenues essentielles à l'échange d'informations et de nouvelles. Ainsi, les plateformes assument le rôle des médias, même si elles ne produisent généralement pas leur propre contenu, mais agissent simplement comme intermédiaires pour le contenu produit par leurs utilisateurs - y compris les médias.

La technologie a non seulement permis de nouveaux moyens de diffusion de contenu, mais elle a également donné la parole à ceux qui ne l'avaient pas auparavant, ou dont la portée était limitée, et leur a permis d'amplifier leurs messages à une échelle sans précédent. Le développement d'internet et, plus tard, des réseaux sociaux, a démocratisé l'espace public et a permis aux citoyens de participer à un débat public d'une manière qui n'était pas possible auparavant. Mais il y a aussi un revers à la médaille. Les politiciens peuvent désormais s'adresser directement à leurs électeurs via les réseaux sociaux et les services de messagerie, ce qui ouvre la voie à la

[1] https://www.bbc.com/news/world-europe-39265777
[2] https://www.politico.com/magazine/story/2016/12/fake-news-history-long-violent-214535/

manipulation. Les théoriciens du complot, les entreprises, les machines de propagande ou tout autre créateur de discours, grâce aux algorithmes utilisés par les réseaux sociaux, peuvent désormais cibler des groupes de population spécifiques définis sur la base de caractéristiques telles que le sexe, le lieu, les opinions religieuses ou politiques, l'éducation, le statut économique, etc.

D'une part, le volume et la variété des nouvelles et des sources ont augmenté, ce qui contribue potentiellement à la démocratisation et au pluralisme du débat public. D'autre part, la quantité d'informations et de sources en ligne rend difficile pour les citoyens de faire la distinction entre les informations crédibles et trompeuses. Notre écosystème de l'information est devenu dangereusement pollué et nous divise au lieu de nous connecter[3]. Nous vivons à une époque de désordre de l'information, ce qui a une autre implication dangereuse: il sape gravement la confiance du public dans les médias professionnels et d'autres institutions. Il déforme la capacité des gens à donner un sens au monde qui les entoure et menace les processus démocratiques dans le monde entier[4].

Le problème devient encore plus complexe lorsque nous ajoutons à l'équation la fracture numérique dans laquelle les générations plus âgées sont laissées pour compte non seulement en termes d'accès et d'utilisation des technologies, mais surtout en termes de compréhension et d'approche critique de l'environnement numérique, ce qui les rend particulièrement vulnérables à la désinformation et à ses dangers. Il est d'une importance cruciale de soutenir les initiatives d'éducation aux médias et à l'information afin de donner aux personnes âgées les moyens d'être des citoyens numériques actifs. Cette analyse de Clara démontrera en détail les mesures qui peuvent être prises pour résoudre ce problème urgent afin d'aider les générations âgées sur deux fronts: combler le fossé technologique entre les générations et permettre aux générations âgées de profiter des aspects positifs des développements technologiques; fournir un niveau élevé de sensibilisation aux dangers et aux défis en ligne, en particulier dans le domaine de la désinformation et de la propagande.

[3] Wardle, C. (2020), Comprendre le trouble de l'information.
https://firstdraftnews.org/long-form-article/understanding-information-disorder/
[4] EuroDIG (2018), Trouble de l'information: causes, risques et remèdes.
https://eurodigwiki.org/wiki/Information_disorder:_causes,_risks_and_remedies._%E2%80%93_PL_02_2018

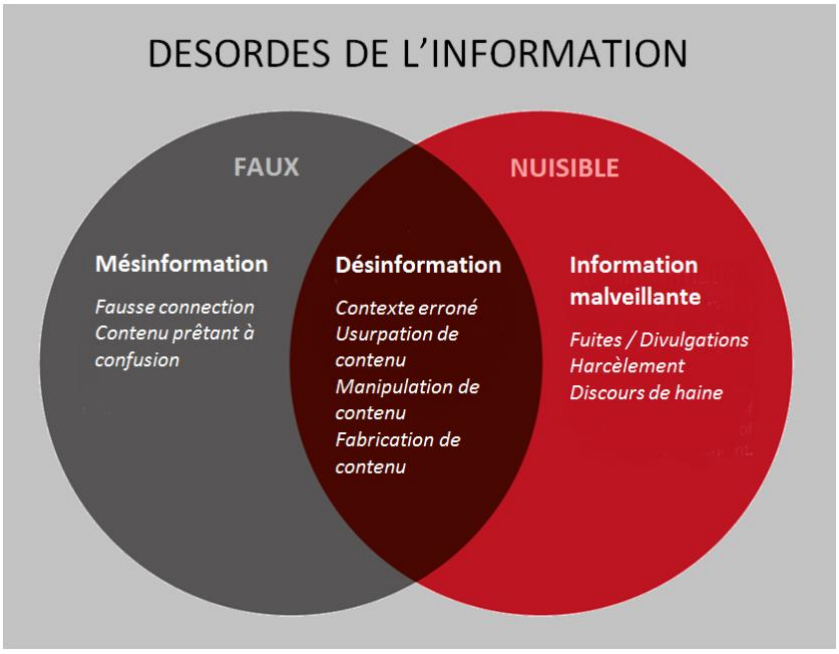
2. Désordre de l'information – quels sont les défis pour la liberté d'expression, le vivre ensemble et la démocratie ?

2.1. Notions de désinformation

2.1.1. Définitions

Le terme « désordre de l'information » se rapporte à tous les défis de l'environnement informationnel moderne divisé en trois catégories : la désinformation, la désinformation et la désinformation. C'est le cadre conceptuel introduit par le rapport sur le *« Désordre de l'information : vers un cadre interdisciplinaire pour la recherche et l'élaboration des politiques »*[5] commandé par le Conseil de l'Europe en 2017 dans le but d'examiner de manière exhaustive le désordre de l'information, ses défis connexes, et de définir les moyens de lutter contre la « pollution de l'information ».
Les auteurs identifient trois types différents de troubles de l'information en fonction de la fausseté de l'information, de l'intention de l'acteur et du type de préjudice:

- la « mésinformation » est une information fausse mais sa diffusion n'est pas destinée à nuire ;
- la « désinformation » est une information fausse délibérément diffusée pour nuire ;
- « l'information malveillante » est une information authentique diffusée dans le but de nuire, souvent en rendant publique des informations destinées à rester privées.



Ce cadre conceptuel a été largement adopté et développé par de nombreux chercheurs, institutions et organisations.

[5] Wardle, C. et Derakhshan, H. (2017). Désordre de l'information : Vers un cadre interdisciplinaire pour la recherche et l'élaboration de politiques. Strasbourg: Conseil de l'Europe. https://rm.coe.int/information-disorder-report-version-august-2018/16808c9c77

Comme on peut le constater, trois éléments clés se dégagent des définitions relatives à la désinformation : la désinformation concerne (a) les fausses informations, (b) diffusées avec une intention spécifique (malveillante ou de mauvaise foi) et (c) cause certains préjudices.

Les tentatives subséquentes de définition de la désinformation se sont éloignées du critère de l'intention de causer un préjudice, étant donné que l'intention est difficile à prouver[6]. Au lieu de cela, l'intention de tromper et la possibilité de causer un préjudice public sont soulignées, comme dans la définition fournie par le Plan d'action pour la démocratie européenne (2020) qui définit la désinformation comme *« un contenu faux ou trompeur qui est diffusé dans l'intention de tromper ou d'obtenir un gain économique ou politique et qui peut causer un préjudice public »*. Cette définition révèle en outre les éléments qui servent de motifs à leur large diffusion, à savoir le gain économique (par exemple, par le biais d'appâts à clics) et les motifs politiques, y compris l'ingérence (étrangère) dans les processus démocratiques/électoraux[7].
Le groupe d'experts de haut niveau de l'Union européenne sur les fausses informations et la désinformation en ligne reconnaît que, bien qu'elle ne soit pas nécessairement illégale, la désinformation peut néanmoins être préjudiciable aux citoyens et à la société dans son ensemble. Le risque de préjudice comprend les menaces aux processus démocratiques, y compris l'intégrité des élections, et aux valeurs démocratiques qui façonnent les politiques publiques dans divers secteurs, tels que la santé, la science, les finances et plus encore.

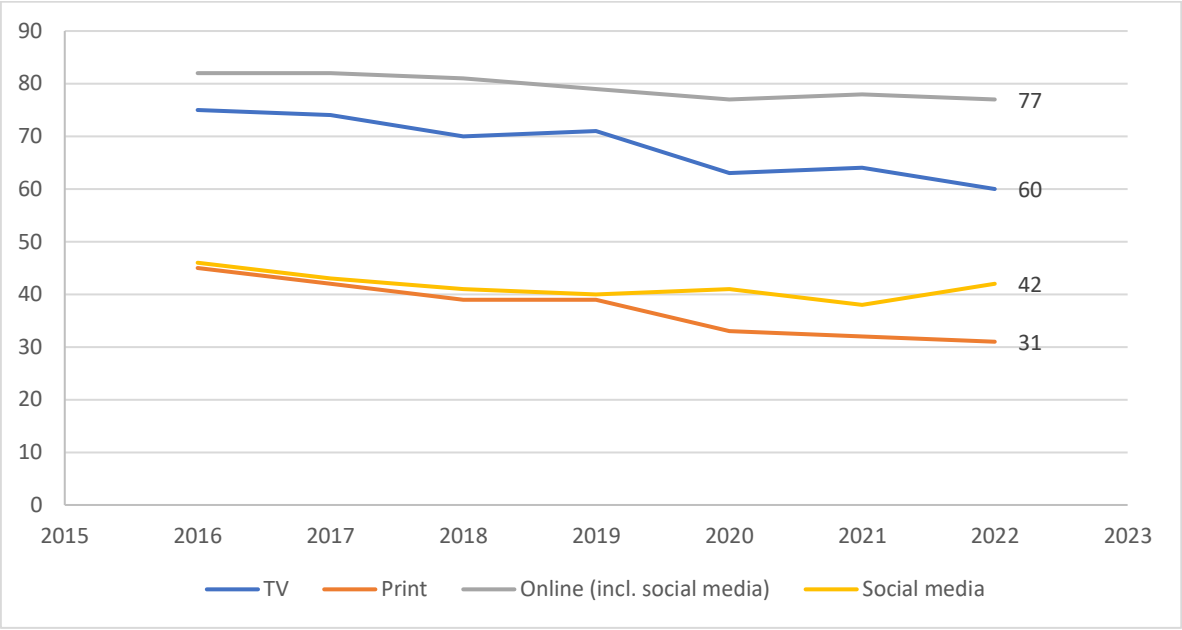
La désinformation, même si elle manque d'intention malveillante, peut également avoir des conséquences néfastes. Elle se propage lorsque la personne ne se rend pas compte qu'elle partage des informations fausses ou trompeuses, que ce soit par naïveté, ignorance, examen insuffisant, rapidité... La pandémie de Covid-19 a montré à quel point le partage d'informations fausses, non vérifiées et non sourcées peut être problématique. Il est donc nécessaire de prêter attention à ce phénomène ainsi qu'aux mesures visant à le prévenir.

2.1.2. Diffusion

Même si les médias audiovisuels traditionnels continuent de servir de principale source d'information, en particulier pour les générations plus âgées, les recherches suggèrent que les citoyens qui obtiennent des informations en ligne le font de plus en plus par le biais d'intermédiaires[8], et moins directement par l'intermédiaire des médias en ligne. Les données pour la Belgique publiées dans le Reuters Institute Digital News Report 2022 montrent que le pourcentage de personnes utilisant les médias sociaux comme source d'information est en hausse[9] :

Sources des nouvelles (%)

[6] McGonagle, T. and Katie Pentney, K. „From risk to reward? The DSA's risk-based approach to disinformation", in Cappello M. (ed.), *Unravelling the Digital Services Act package*, IRIS Special, European Audiovisual Observatory, Strasbourg, 2021, p. 46.
[7] Ibid., p. 47.
[8] https://europa.eu/eurobarometer/surveys/detail/2832
[9] Newman, Nic et al., Reuters Institute Digital News Report 2022. https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2022-06/Digital_News-Report_2022.pdf



Source: Reuters Institute 2022 Digital News Report

Le rôle des médias dans cette chaîne de valeur est très important. Avec l'essor des médias en ligne et l'évolution des tendances de consommation des médias, la publicité s'est également déplacée en ligne. La publicité numérique bénéficie particulièrement de la publicité dite ciblée qui se concentre sur les caractéristiques, les intérêts et les préférences spécifiques d'un consommateur. Cela peut prendre la forme d'un ciblage contextuel (diffusion d'annonces basées sur le contenu d'un site internet que l'utilisateur visite) ou d'un ciblage comportemental (annonces basées sur les préférences et l'historique de recherche des utilisateurs).

Compte tenu de la prévalence de la publicité numérique, les médias sont obligés de se déplacer en ligne et de développer des modèles commerciaux alternatifs pour assurer leur viabilité et rester pertinents. Malheureusement, cela se fait souvent au détriment du professionnalisme. Outre la baisse globale des revenus publicitaires qui rend souvent les organes de presse dépendants de sources de pouvoir politique ou économique, les impératifs de rapidité, d'économie de l'attention[10] et de journalisme piège à clics[11] ont un impact important sur la façon dont les décisions éditoriales sont prises. Cela constitue un terrain fertile pour la diffusion d'informations incorrectes, non vérifiées et trompeuses (souvent involontaires).

Les médias ont toujours un rôle clé à jouer dans la production d'informations et d'autres contenus d'intérêt général, mais les plateformes en ligne, telles que les réseaux sociaux, les moteurs de recherche ou les plateformes de partage de vidéos, ont assumé le rôle d'intermédiaires entre les médias et leur public. En outre, ils distribuent également du contenu produit et téléchargé par leurs utilisateurs. Ce faisant, leur rôle n'est pas simplement automatique et passif, c'est-à-dire qu'ils ne sont pas seulement des plateformes techniques de stockage de contenu. Grâce à leur

[10] Terme utilisé pour décrire le fait que les entreprises de technologie comptent sur la captation de l'attention des utilisateurs pour gagner de l'argent, en le vendant à des annonceurs. https://www.humanetech.com/youth/the-attention-economy

[11] « Clickbait » est un titre sensationnaliste qui vous encourage à cliquer sur un lien vers un article, une image ou une vidéo. Une fois que vous cliquez, le site internet hébergeant le lien génère des revenus auprès des annonceurs, mais le contenu réel est généralement d'une qualité et d'une précision douteuses. Les sites internet utilisent des pièges à clics pour attirer autant de clics que possible, augmentant ainsi leurs revenus publicitaires. https://edu.gcfglobal.org/en/digital-media-literacy/what-is-clickbait/1/

conception, à leur tri algorithmique et à leur classement des contenus et des systèmes de recommandation, ces intermédiaires numériques façonnent l'environnement informationnel dans lequel les citoyens interagissent avec les nouvelles provenant des médias et d'autres sources.

2.1.2.1. Systèmes de recommandation

Les plateformes en ligne dotées de fonctions de recherche et/ou de recommandation ont une influence considérable sur la disponibilité, l'accessibilité, la visibilité, la repérabilité et l'importance de contenus particuliers. Cette influence est obtenue, en partie, grâce à l'utilisation de la personnalisation algorithmique (ou systèmes de recommandation)[12].

Les systèmes de recommandation sont utilisés pour aider les utilisateurs à naviguer dans la grande quantité d'informations disponibles en ligne. Ils suggèrent du contenu qui est ou pourrait être pertinent pour les utilisateurs, comme du texte à lire, des vidéos ou des films à regarder, des produits à acheter. En d'autres termes, ils créent algorithmiquement des recommandations personnalisées de médias et d'autres contenus, en fonction des caractéristiques et des préférences des utilisateurs individuels.

En plus de recommander du contenu, les plateformes modèrent (filtrent et suppriment) et classent (donnent une plus grande visibilité à certains contenus ou sources). Les algorithmes dictent ce que nous trouvons et lisons en premier sur internet, car ils fonctionnent en évaluant sélectivement ce que l'utilisateur aimerait voir, en fonction de ses données personnelles telles que l'emplacement, les clics et les likes précédents, l'historique de recherche... Le but ultime de la recommandation et du classement du contenu est de maximiser l'engagement des utilisateurs, ce qui, à son tour, augmente la visibilité du contenu. Cela inclut le partage, le like, les commentaires, les clics et, finalement, le temps passé à visionner certains contenus, qui sont également mesurés par les plateformes. Tout leur modèle économique repose sur cette formule : plus l'engagement des utilisateurs est élevé, plus les revenus publicitaires sont élevés.

Un tel mode de diffusion du contenu, motivé principalement par les politiques internes et l'intérêt commercial des entreprises privées et non par l'intérêt public, déterminé par des algorithmes et non par des décisions éditoriales, a un impact énorme sur la visibilité de la désinformation et sur la manière dont elle est diffusée davantage.

Une autre préoccupation pour le pluralisme et la démocratie est le manque notoire de transparence concernant les politiques de modération de contenu des plateformes en ligne et les critères selon lesquels leurs algorithmes sélectionnent, classent et recommandent différents types de contenu à leurs utilisateurs. Ce problème a été souligné par beaucoup, malgré les engagements pris par les grandes plateformes en ligne de *« prendre les mesures nécessaires pour permettre un accès conforme aux données pour les activités de vérification des faits et de recherche »* et de *« coopérer en fournissant des données pertinentes sur le fonctionnement de leurs services, y compris des données pour une enquête indépendante par des chercheurs universitaires et des informations générales sur les algorithmes »*[13].

[12] McGonagle, T. and Pentney, K., 2021, p. 45.
[13] Code de bonnes pratiques contre la désinformation (2018).
https://digital-strategy.ec.europa.eu/fr/library/2018-code-practice-disinformation

Par exemple, le groupe des régulateurs européens des services de médias audiovisuels (ERGA) a surveillé la mise en œuvre du code de bonnes pratiques de l'Union européenne contre la désinformation en 2019. Dans le cadre de ce travail, l'Ofcom, le régulateur britannique, a surveillé la mise en œuvre des engagements pris par Google, Facebook et Twitter lors de la campagne électorale précédant les élections nationales du 12 décembre 2019, l'une des principales conclusions étant le manque de transparence des critères de ciblage : « *les informations sur les personnes qui ont été exposées à une publicité, leur degré d'engagement, et l'information sur les critères de ciblage utilisés est relativement limitée tant pour les chercheurs que pour les utilisateurs. Dans son échantillonnage de publicités sur Facebook et YouTube, l'Ofcom a rencontré des cas de publicités semblant cibler des groupes électoraux spécifiques, ce qui suggère la possibilité que des critères de ciblage plus sophistiqués soient disponibles pour les annonceurs que ceux capturés dans les bibliothèques publicitaires des plates-formes* »[14].

2.1.2.2. Bulles de filtre et chambres d'écho

Deux caractéristiques particulières de l'environnement des médias sociaux peuvent exacerber davantage le désordre de l'information, à savoir les bulles de filtres et les chambres d'écho. Ces phénomènes sont étroitement liés à la tendance humaine naturelle à préférer l'information qui confirme nos propres attitudes et visions du monde. En raison du filtrage algorithmique sur les médias sociaux, comme expliqué ci-dessus, les utilisateurs sont le plus souvent exposés précisément aux attitudes avec lesquelles ils sont déjà d'accord à l'avance. Cela produit un effet de chambre d'écho, une chambre d'écho étant définie comme *« un environnement dans lequel quelqu'un ne rencontre que des opinions et des croyances similaires aux siennes et n'a pas à envisager d'alternatives »*[15].

La sélection personnalisée du contenu prive donc les utilisateurs de points de vue et d'opinions différents, et donc une opportunité de dialogue démocratique. En plus de nous empêcher d'avoir accès à des points de vue et opinions alternatifs, les chambres d'écho favorisent la propagation de la désinformation. Comme l'a souligné la Commission européenne dans sa communication intitulée *« Lutter contre la désinformation en ligne: une approche européenne », « en facilitant le partage de contenus personnalisés entre utilisateurs partageant les mêmes idées, les algorithmes accentuent indirectement la polarisation et renforcent les effets de la désinformation »*[16].

Par exemple, les « Facebook Files » publiés par le Washington Post au début de 2021 ont révélé les résultats de recherches internes menées par Facebook pour tenter de comprendre la propagation des idées qui contribuent à la réticence à la vaccination. La recherche a montré qu'un petit nombre d'utilisateurs semble jouer un rôle important : seulement 10 des 638 segments de population contenaient 50% de tout le contenu sur l'hésitation vaccinale sur la plate-forme[17].

Un autre aspect mérite d'être mentionné à cet égard : dans la surabondance de sources et de contenus, il est devenu de plus en plus difficile de juger de leur crédibilité - cela signifie que les gens dépendent de plus en plus de leurs amis et des membres

[14] https://erga-online.eu/wp-content/uploads/2020/05/ERGA-2019-report-published-2020-LQ.pdf
[15] https://www.oxfordlearnersdictionaries.com/us/definition/english/echo-chamber
[16] Communication de la Commission au Parlement européen, au Conseil, au Comité économique et social européen et au Comité des Régions, Lutter contre la désinformation en ligne: une approche européenne, 2018.
https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:52018DC0236&from=EN
[17] https://www.washingtonpost.com/technology/2021/03/14/facebook-vaccine-hesistancy-qanon/

de leur famille pour les guider dans l'écosystème de l'information. En outre, les informations transmises en temps réel entre pairs de confiance sont beaucoup moins susceptibles d'être contestées[18].

2.1.2.3. Techniques de diffusion manipulatrices: faux comptes, bots et trolls

Afin d'amplifier leur portée et leur impact sur l'opinion publique, les campagnes de désinformation utilisent une gamme de techniques de diffusion, telles que les faux comptes sur les réseaux sociaux, les bots et le trollage organisé, qui représente une utilisation abusive et une manipulation des services de la plateforme en ligne.

Les faux comptes sont *des « profils simulés qui n'ont pas d'utilisateur authentique derrière eux »*[19]. Ceux-ci peuvent être sous la forme de « bots », soit des comptes contrôlés par un logiciel et servant le même objectif. Ils peuvent aimer, partager et commenter sur les plateformes de médias sociaux, ou effectuer des interactions simples, comme le retweeting automatique. Une autre tactique de manipulation consiste à embaucher les soi-disant trolls – des humains qui publient derrière un faux nom d'utilisateur, dont la tâche est de cibler spécifiquement ceux qui expriment des opinions différentes, par exemple en publiant des attaques personnelles pour les saper et finalement les faire taire, afin de manipuler les opinions et le débat public, gonfler la popularité, etc. Les trolls font généralement partie de campagnes orchestrées, parfois à grande échelle, appelées « usines à trolls ». Les usines à trolls sont également utilisées par les gouvernements du monde entier dans des campagnes de désinformation parrainées par l'État[20], telles que la campagne de la Russie pour façonner l'opinion internationale autour de son invasion de l'Ukraine, utilisée déjà en 2014. Les rapports montrent que le Kremlin a recruté et entraîné des trolls en ligne afin de diffuser son message dans la section des commentaires des principaux sites internet américains: *« Les documents montrent des instructions fournies aux commentateurs qui détaillent la charge de travail attendue d'eux. Au cours d'une journée de travail moyenne, les Russes doivent publier des articles de presse 50 fois. Chaque blogueur doit maintenir six comptes Facebook publiant au moins trois messages par jour et discutant de l'actualité en groupe au moins deux fois par jour. À la fin du premier mois, ils devraient avoir gagné 500 abonnés et obtenir au moins cinq messages sur chaque article par jour. Sur Twitter, les blogueurs devraient gérer 10 comptes avec jusqu'à 2 000 abonnés et tweeter 50 fois par jour »*[21].

Un autre exemple est la Chine, qui a payé des gens pour publier des millions de messages fabriqués sur les médias sociaux chaque année, dans le cadre d'un effort visant à *« distraire régulièrement le public et à changer de sujet »* de toute question politique qui menace d'inciter à des manifestations[22].

De nombreuses techniques de manipulation ont également été reconnues par les plateformes en ligne comme une utilisation abusive de leurs services, et celles-ci prennent en effet des mesures afin de les prévenir et de les détecter, ainsi que de supprimer les faux comptes[23]. Cependant, ce n'est pas toujours une tâche facile. La distinction entre les comptes inauthentiques et authentiques devient de plus en plus floue. Les comptes peuvent être piratés, achetés ou loués, et certains utilisateurs

[18] Wardle C. et Derakhshan H. (2017).
[19] Communication de la Commission.
[20] https://www.theguardian.com/media/2016/nov/06/troll-armies-social-media-trump-russian
[21] https://www.buzzfeednews.com/article/maxseddon/documents-show-how-russias-troll-army-hit-america#.jnn07ep33
[22] Wardle, C. et Derakhshan, H. (2017).
[23] Par exemple, Twitter affirme qu'il supprime un million de comptes automatisés chaque jour: https://abcnews.go.com/Technology/wireStory/twitter-removes-million-spam-accounts-day-86382214

« font don » de leurs informations d'identification à des organisations qui publient en leur nom[24]. Il existe même des comptes hybrides « cyborgs » contrôlés à la fois par des algorithmes et des humains, où un humain prend périodiquement le contrôle d'un compte de bot, ce qui rend le compte plus difficile à identifier en tant que bot.

Les mesures visant à prévenir les abus par ces techniques de manipulation ont également été soulignées dans tous les instruments européens visant à lutter contre la désinformation, comme cela sera expliqué ci-dessous, mais aussi au niveau de certains États membres. Par exemple, en France, l'ancien régulateur des médias audiovisuels CSA[25] a adopté une recommandation[26] au titre de la loi de 2018 relative à la lutte contre la manipulation de l'information dans laquelle les plateformes en ligne sont encouragées à mettre en place des *« procédures appropriées permettant la détection de comptes diffusant de fausses informations à grande échelle », « des procédures proportionnées destinées à entraver les actions de ces comptes »,* et un *« espace d'information facilement accessible fournissant aux utilisateurs des informations claires et détaillées sur les pratiques susceptibles d'entraîner une action de l'exploitant (création d'un nombre anormal de comptes, partage de contenus à des tarifs anormaux, utilisation d'informations fausses, volées ou trompeuses, etc.) ».*

Ainsi, le type de comportement ciblé est un nombre anormal de comptes, des taux anormaux de partage et l'utilisation de renseignements faux ou trompeurs.[27]

2.1.3. L'appel et le pouvoir de la désinformation

Comme expliqué ci-dessus, le partage de la désinformation est motivé par trois facteurs: gagner de l'argent; d'avoir une influence politique, étrangère ou nationale; ou même de causer des problèmes pour le plaisir[28]. La désinformation, d'autre part, décrit un faux contenu partagé par une personne qui ne se rend pas compte qu'il est faux ou trompeur. Souvent, une désinformation est reprise par quelqu'un qui la partage avec ses réseaux, involontairement et de bonne foi, croyant qu'il aide. Il y a aussi des facteurs socio-psychologiques qui doivent être pris en compte : *« En ligne, les gens exécutent leur identité. Ils veulent se sentir liés à leur « tribu », qu'il s'agisse de membres du même parti politique, de parents qui ne vaccinent pas leurs enfants, de militants préoccupés par le changement climatique ou de ceux qui appartiennent à une certaine religion, race ou groupe ethnique »*[29].

La Communication de la Commission européenne a également souligné le rôle des utilisateurs dans la diffusion de la désinformation, qui a tendance à voyager plus rapidement sur les médias sociaux en raison de la propension des utilisateurs à partager du contenu sans aucune vérification préalable. Les fausses nouvelles ont tendance à se propager de manière plus virale, car elles tirent parti de la conception

[24] https://theconversation.com/how-many-bots-are-on-twitter-the-question-is-difficult-to-answer-and-misses-the-point-183425
[25] Transformé en ARCOM à compter du 1er janvier 2022.
[26] https://www.csa.fr/Reguler/Espace-juridique/Les-textes-adoptes-par-l-Arcom/Les-deliberations-et-recommandations-de-l-Arcom/Recommandations-et-deliberations-du-CSA-relatives-a-d-autres-sujets/Recommandation-n-2019-03-du-15-mai-2019-du-Conseil-superieur-de-l-audiovisuel-aux-operateurs-de-plateforme-en-ligne-dans-le-cadre-du-devoir-de-cooperation-en-matiere-de-lutte-contre-la-diffusion-de-fausses-informations
[27] https://erga-online.eu/wp-content/uploads/2021/03/ERGA-SG2-Report-2020-Notions-of-disinformation-and-related-concepts-final.pdf
[28] Wardle C. (2020).
[29] Ibid..

de la plate-forme en ligne pour générer des revenus et sont centrées sur des questions qui font appel aux émotions des lecteurs.

Tout d'abord, comme déjà mentionné, le modèle économique même des plateformes en ligne est axé sur les algorithmes et la publicité: il est basé sur le clic, ce qui facilite le placement de publicités sur des sites internet qui publient des contenus sensationnalistes faisant appel à la curiosité et aux émotions des utilisateurs, y compris la désinformation. En outre, comme l'a fait observer la Commission dans sa Communication, la désinformation est un outil d'influence puissant et peu coûteux – et souvent économiquement rentable: « *À ce jour, la plupart des cas connus concernaient des articles écrits, parfois complétés par des images authentiques ou du contenu audiovisuel sorti de leur contexte. Mais une technologie nouvelle, abordable et facile à utiliser est maintenant disponible pour créer de fausses images et du contenu audiovisuel (ce qu'on appelle des « deep fakes »), offrant des moyens plus puissants de manipuler l'opinion publique ».*

Une autre raison pour laquelle le contenu tel que les théories du complot et les faux récits a tant de succès est qu'il joue sur les émotions des gens, telles que les peurs profondément enracinées et d'autres vulnérabilités. Dans sa Déclaration de 2019 sur les capacités de manipulation des processus algorithmiques, le Comité des Ministres du Conseil de l'Europe note qu'une manipulation inacceptable *« peut prendre la forme d'une influence subliminale, exploiter des vulnérabilités existantes ou des biais cognitifs, et/ou empiéter sur l'indépendance et l'authenticité de la prise de décision individuelle ».* La Déclaration souligne également les dangers des outils d'apprentissage automatique contemporains pour les sociétés démocratiques et leur possibilité de manipuler et de contrôler les comportements sociaux et politiques en raison de *leur « capacité croissante non seulement à prédire les choix, mais aussi à influencer les émotions et les pensées et à modifier un plan d'action anticipé, parfois de manière subliminale »* [30].

La littérature met également l'accent sur le pouvoir de *la « familiarité »* du contenu, et l'une des techniques les plus efficaces pour que les gens approuvent certains récits est sa répétition. C'est précisément pourquoi les campagnes de désinformation emploient des robots qui « aiment » ou « partagent » automatiquement des histoires, créant ainsi un faux sentiment de popularité et de familiarité du contenu. Le rapport du Conseil de l'Europe susmentionné reconnaît également ce fait et souligne quatre caractéristiques des stratégies efficaces de désinformation qui rendent un message plus attrayant et donc plus susceptible d'être approuvé et largement partagé :[31]

- provoquer une réaction émotionnelle;
- répétition;
- un aspect visuel fort;
- un récit puissant.

[30] Conseil de l'Europe (2019), Déclaration du Comité des Ministres sur les capacités de manipulation des processus algorithmiques. https://ccdcoe.org/uploads/2019/09/CoE-190213-Declaration-on-manipulative-capabilities-of-algorithmic-processes.pdf

[31] Wardle, C. et Derakhshan, H. (2017).

2.2. Désordre de l'information : défis pour la liberté d'expression et la démocratie

Dans sa Communication, la Commission européenne note que « *l'exposition des citoyens à la désinformation à grande échelle, y compris aux informations trompeuses ou carrément fausses, constitue un défi majeur pour l'Europe. La désinformation érode la confiance dans les institutions et dans les médias numériques et traditionnels, et nuit à nos démocraties en entravant la capacité des citoyens à prendre des décisions éclairées. La désinformation soutient aussi souvent des idées et des activités radicales et extrémistes... La propagation de la désinformation affecte également les processus d'élaboration des politiques en faussant l'opinion publique. Les acteurs nationaux et étrangers peuvent utiliser la désinformation pour manipuler les politiques, les débats sociétaux et les comportements dans des domaines tels que le changement climatique, la migration, la sécurité publique, la santé et la finance. La désinformation peut également diminuer la confiance dans la science et les preuves empiriques* ».

La désinformation représente une grave préoccupation pour la démocratie et le bien-être public. Alors que certaines notions couvertes par le concept de désinformation peuvent constituer des informations illégales et illicites – et sont donc sanctionnées par la loi – les informations préjudiciables ne sont pas réglementées par la loi. L'information préjudiciable fait souvent référence à des choses qui, bien que dans les limites de la loi, sont considérées comme socialement indésirables. Cela peut inclure des théories du complot fabriquées, des opinions politiques extrêmes, des théories médicales erronées telles que dans le mouvement anti-vaccination, ou des reportages trompeurs. Bien que potentiellement nuisibles, ces notions font partie d'un débat social sain et bénéficient donc de la protection du droit à la liberté d'expression[32].

Si elle n'est pas nécessairement illégale ou illégale, pourquoi la désinformation est-elle dangereuse? Comme cela a été démontré à de nombreuses reprises dans le monde entier, elle peut infliger des dommages considérables aux citoyens et à la société dans son ensemble. Elle menace les processus politiques démocratiques, met en danger les processus électoraux, influence négativement les valeurs démocratiques qui façonnent les politiques publiques dans différents secteurs tels que la santé, la science, l'éducation, les finances, etc. Elle peut également nuire à la sécurité de l'État et à la stabilité d'une société, y compris la sécurité des citoyens. Le risque posé par la désinformation et les dommages qu'elle peut causer sont devenus particulièrement visibles pendant la pandémie de Covid-19, lorsqu'une expression « infodémie » a été introduite par l'Organisation mondiale de la santé (OMS) pour décrire *« trop d'informations, y compris des informations fausses ou trompeuses, dans des environnements numériques et physiques pendant une période d'explosition d'une maladie »*, ce qui crée de la confusion

[32] Joris van Hoboken et al. (2019). Le cadre juridique relatif à la diffusion de la désinformation par les services Internet et à la réglementation de la publicité politique. Institut du droit de l'information (IViR), Université d'Amsterdam.

et des comportements à risque qui peuvent nuire à la santé, ce qui entraîne une méfiance à l'égard des autorités sanitaires et sape la réponse de santé publique. Un accent particulier devrait être mis précisément sur les conséquences négatives que la désinformation peut avoir sur la confiance du public et l'engagement des citoyens. Un autre exemple de ceci sont les récits autour du changement climatique. La recherche examinant l'impact de l'exposition aux théories du complot liées au climat a montré que l'exposition à de telles théories créait un sentiment d'impuissance, entraînant un désengagement de la politique et une probabilité réduite pour les gens de faire de petits changements qui réduiraient leur empreinte carbone[33].

La désinformation est une préoccupation croissante pour les citoyens européens. Selon l'Eurobaromètre de mars 2018 sur les fausses nouvelles et la désinformation en ligne, 85% des citoyens de l'Union européenne perçoivent les fausses nouvelles comme un problème dans leur pays, 83% les perçoivent comme un problème pour la démocratie en général et 73% sont préoccupés par la désinformation en ligne pendant les périodes préélectorales. Les inquiétudes du public ont notamment été amplifiées par la crise sanitaire du Covid-19[34].

Plusieurs sujets de préoccupation peuvent être mis en évidence en ce qui concerne la désinformation et son influence néfaste, dont les plus importants comprennent la propagande et l'influence étrangère, les discours de haine et les crises sanitaires.

2.2.1. Propagande et influence étrangère

L'utilisation de la propagande comme arme de guerre n'est – encore une fois – pas nouvelle : l'histoire regorge de tels exemples. Mais ce phénomène n'a jamais été aussi agressif, en partie parce qu'il est fortement influencé par les campagnes de désinformation qui se propagent à grande vitesse via internet et les téléphones mobiles, en utilisant de nouvelles techniques sophistiquées telles que l'intelligence artificielle qui rend difficile la distinction entre les faux et la réalité. L'essor des médias sociaux a considérablement changé la nature des campagnes électorales, ouvrant de nouvelles voies de communication entre les politiques et l'électorat et contournant les médiateurs traditionnels, tels que les journalistes[35]. En effet, les nouvelles technologies sont exploitées pour la production de désinformation ou de contenu source de division, par de multiples acteurs, y compris des États, des partis politiques, des politiciens et d'autres individus ou entreprises puissants, pour de multiples motifs - politiques, idéologiques ou

[33] Wardle, C. et Derakhshan, H. (2017).
[34] https://europa.eu/eurobarometer/surveys/detail/2183
[35] Höller, Maximilien (2021). La composante humaine dans les médias sociaux et les fausses nouvelles: la performance des leaders d'opinion britanniques sur Twitter pendant la campagne du Brexit. https://www.tandfonline.com/doi/full/10.1080/13825577.2021.1918842

commerciaux[36]. Cela peut venir des plus hauts représentants du pouvoir exécutif : aux États-Unis, le Washington Post a calculé que durant son mandat, 30 573 fausses allégations ou mensonges ont été diffusés par le président Donald Trump (principalement via son compte Twitter). Au Brésil, un juge de la Cour suprême a ordonné en août 2021 l'ouverture d'une enquête contre le président brésilien Jair Bolsonaro, pour diffusion de fausses informations. La décision a été prise suite à une demande du Tribunal électoral supérieur d'ouvrir une enquête contre le chef de l'Etat pour ses attaques contre le système de vote électronique et la légitimité des élections de 2022.

Une autre caractéristique inséparable du paysage de la désinformation et de la propagande est le concept d'influence étrangère. Comme l'a noté la Commission européenne, cela se produit lorsque des acteurs étrangers et des pays tiers sont engagés dans des opérations d'influence ciblées visant à saper le débat démocratique et à exacerber la polarisation sociale. La Communication de la Commission, en particulier, établit un lien entre les campagnes de désinformation en ligne de masse et les acteurs « étrangers » et les « pays tiers »: « *Les campagnes de désinformation en ligne de masse sont largement utilisées par une série d'acteurs nationaux et étrangers pour semer la méfiance et créer des tensions sociétales, avec de graves conséquences potentielles pour notre sécurité. En outre, les campagnes de désinformation menées par des pays tiers peuvent faire partie de menaces hybrides pour la sécurité intérieure, y compris les processus électoraux, en particulier en combinaison avec des cyberattaques. Par exemple, la doctrine militaire russe reconnaît explicitement la guerre de l'information comme l'un de ses domaines* ». Il a été reconnu que l'UE est souvent la cible de campagnes de désinformation visant à saper ses institutions, ses politiques, ses actions et ses valeurs. En 2015, la task-force East Stratcom a été créée au sein du Service européen pour l'action extérieure afin de lutter contre les campagnes de désinformation provenant de Russie.

En effet, beaucoup a été écrit et dit sur l'impact de la propagande russe sur l'environnement de l'information et les processus démocratiques en Europe, aux États-Unis et au-delà, en particulier en s'attachant à influencer le cours des processus électoraux afin de favoriser un candidat, voire de nuire à un autre. Les développements entourant l'élection présidentielle américaine de 2016 et le referendum sur le Brexit sont des exemples frappants de cette ingérence. En France, les médias publics Russia Today France et Radio Sputnik ont été accusés d'avoir diffusé de fausses informations pendant la campagne présidentielle de 2017, avec l'intention claire de nuire au candidat Emmanuel Macron.

Le résultat de l'élection présidentielle américaine de 2016 a ouvert de nombreuses questions quant au pouvoir de la désinformation et de l'influence étrangère. L'ingérence

[36] « Désinformation et liberté d'opinion et d'expression », rapport de la Rapporteuse spéciale sur la promotion et la protection du droit à la liberté d'opinion et d'expression, Irene Khan. Conseil des droits de l'homme, Assemblée générale des Nations Unies, avril 2021, p. 4.

russe, comme le révèlent des sources officielles, impliquait de nombreuses tactiques, y compris le piratage des données d'inscription des électeurs, la campagne d'Hillary Clinton et d'autres opposants de Trump, mais une partie importante de cette campagne consistait à diffuser de la propagande sur les médias sociaux. Les plans de la Russie pour influencer les élections américaines ont commencé en avril 2014 avec le développement d'une « ferme à trolls » appelée Internet Research Agency, pour mener une campagne sur les médias sociaux qui favorisait Donald Trump et dénigrait Hillary Clinton. L'Internet Research Agency a également cherché à *« provoquer et amplifier la discorde politique et sociale aux États-Unis »*[37]*.*

Des articles fabriqués et de la désinformation ont été diffusés à partir de médias contrôlés par le gouvernement russe, RT et Sputnik, pour être popularisés sur des comptes pro-russes sur Twitter et d'autres médias sociaux : « *Dans une opération qui a coûté des millions de dollars, les Russes ont étudié les groupes politiques américains, se sont déplacés pour recueillir des renseignements dans plusieurs États et ont développé un réseau de faux comptes qu'ils ont utilisés pour infecter l'électorat américain. Tout au long de 2016, ils ont publié du contenu controversé sur des sujets tels que Black Lives Matter, l'immigration et le contrôle des armes à feu; ils ont acheté des publicités politiques critiquant Clinton ; Et ils ont injecté des hashtags comme #Hillary4Prison et #TrumpTrain vers leurs masses d'adeptes* ».[38]

Le pouvoir de la désinformation qui se répand en période électorale est bien illustré par l'exemple suivant, qui a révélé que les histoires fabriquées peuvent être partagées plus largement que les histoires des médias grand public, comme l'a révélé une analyse de BuzzFeed News: *« Au cours des trois derniers mois de la campagne présidentielle américaine, les fausses nouvelles électorales les plus performantes sur Facebook ont généré plus d'engagement que les principales histoires des principaux organes de presse tels que le New York Times, Washington Post, Huffington Post, NBC News, et d'autres... Au cours de ces mois critiques de la campagne, 20 fausses histoires électorales très performantes provenant de sites de canulars et de blogs hyperpartisans ont généré 8 711 000 partages, réactions et commentaires sur Facebook. Au cours de la même période, les 20 reportages électoraux les plus performants de 19 grands sites d'information ont généré un total de 7 367 000 partages, réactions et commentaires sur Facebook ».* Sur les 20 fausses histoires électorales les plus performantes identifiées dans l'analyse, toutes sauf trois étaient ouvertement pro-Donald Trump ou anti-Hillary Clinton[39].

Alors que les motifs des campagnes de désinformation et de propagande sont principalement de nature politique, il y a des exemples où elles ont été motivées par un simple gain financier. L'un des exemples les plus cités est celui d'adolescents d'une petite ville de Macédoine du Nord qui ont créé des dizaines de sites internet pro-Trump, ciblant

[37] https://en.wikipedia.org/wiki/Russian_interference_in_the_2016_United_States_elections
[38] https://time.com/5565991/russia-influence-2016-election/
[39] https://www.buzzfeednews.com/article/craigsilverman/viral-fake-election-news-outperformed-real-news-on-facebook

les lecteurs américains et partageant des histoires sensationnalistes et fabriquées sur les élections américaines. Leur seule motivation était de gagner de l'argent grâce à la publicité en créant un engagement utilisateur à grande échelle sur Facebook.

2.2.2. Discours de haine et violence

La diffusion de désinformation via internet est souvent liée à la diffusion de discours de haine (discutée plus en détail dans une étude du CLARA de 2021). Même si, comme nous l'avons déjà souligné, la désinformation ne concerne pas les contenus illicites en soi, ses conséquences dans la pratique peuvent être extrêmement dangereuses et nuisibles, avoir le même effet ou amplifier l'effet du discours de haine.

La désinformation met souvent délibérément en évidence les différences et les divisions, qu'elles soient entre les partisans de différents partis politiques, nationalités, races, ethnies, groupes religieux, classes socio-économiques... Ces types de messages permettent aux idées discriminatoires et incendiaires d'entrer dans le discours public et d'être traitées comme des faits. Une fois ancrées, ces idées peuvent à leur tour être utilisées pour créer des boucs émissaires, normaliser les préjugés, durcir les mentalités du « nous contre eux » et, dans des cas extrêmes, même catalyser et justifier la violence[40].

Dans certaines parties du monde, la désinformation dirigée contre les personnes en raison de leur identité religieuse, ethnique ou raciale a effectivement conduit à la violence. Un exemple notoire est la campagne de désinformation menée par l'État au Myanmar, qui a publié en ligne des photographies trafiquées et mal étiquetées relatives aux événements de 2017, au cours desquelles les forces de sécurité du Myanmar ont mené une campagne brutale de nettoyage ethnique contre les musulmans rohingyas. Dans son récent rapport, Amnesty International affirme que les algorithmes de Facebook ont amplifié et promu de manière proactive le contenu qui incitait à la violence, à la haine et à la discrimination contre les Rohingyas : *« Dans les mois et les années qui ont précédé et pendant les atrocités de 2017, Facebook au Myanmar est devenu une chambre d'écho de contenu anti-Rohingya virulent. Des acteurs liés à l'armée birmane et à des groupes nationalistes bouddhistes radicaux ont systématiquement inondé la plate-forme Facebook d'incitations ciblant les Rohingyas, semant la désinformation concernant une prise de contrôle imminente du pays par les musulmans et cherchant à dépeindre les Rohingyas comme des envahisseurs sous-humains. La diffusion massive de messages prônant la haine incitant à la violence et à la discrimination à l'égard des Rohingyas, ainsi que d'autres contenus anti-Rohingyas déshumanisants et discriminatoires, a jeté de*

[40] Wardle, C. et Derakhshan, H. (2017)

l'huile sur le feu d'une discrimination de longue date et a considérablement augmenté le risque d'une flambée de violence de masse »[41].

2.2.3. Crises sanitaires

La pandémie de Covid-19 s'est accompagnée d'une vague massive d'informations fausses et trompeuses, de fausses déclarations aux conséquences dangereuses et de théories du complot. Celles-ci allaient d'allégations relativement inoffensives telles que l'ail prévient l'infection; les affirmations qui diminuent la confiance dans les autorités, par exemple que les installations 5G propagent le virus; à des cas très dangereux tels que la consommation d'eau de Javel peut guérir les infections du coronavirus – par exemple, le Centre antipoison belge a enregistré une augmentation de 15% du nombre d'incidents liés à l'eau de Javel[42]. Il n'est pas toujours facile pour le citoyen ordinaire de faire la distinction entre les informations vérifiées et les informations inexactes ou trompeuses, et cela est particulièrement vrai en temps de crise et de situations nouvelles telles que la dernière pandémie, alimentées par le manque de connaissances scientifiques sur la maladie, l'absence de traitements efficaces et les déclarations souvent contradictoires des autorités sanitaires elles-mêmes.

Pendant les crises sanitaires, il y a une surproduction d'informations provenant de sources multiples, dont la qualité, ainsi que la vitesse à laquelle elles sont diffusées, créent des impacts sociaux et sanitaires – un phénomène qualifié d'infodémie par l'OMS, comme déjà mentionné ci-dessus. Une revue systématique sur l'infodémie et la désinformation sur la santé menée par cette institution en 2022 a montré que les effets de l'infodémie, de la désinformation, et des fausses nouvelles constatés par 10 études comprennent :

- réduire la volonté des patients de se faire vacciner,
- faire obstacle aux mesures visant à contenir les foyers de maladies,
- provoquer l'interruption physique de l'accès aux soins de santé,
- amplifier et promouvoir la discorde pour renforcer la crise politique,
- augmentation de la peur sociale, de la panique, du stress et des troubles mentaux,
- améliorer la mauvaise affectation des ressources,
- affaiblir et ralentir les interventions de contre-mesures,
- exacerber la création de contenu de mauvaise qualité[43].

Alors que les médias traditionnels ont sans aucun doute contribué à la diffusion de fausses informations sur la pandémie dans leurs efforts pour rester pertinents, tout en

[41] Amnesty International (2022). Myanmar : L'atrocité sociale : Meta et le droit à un recours pour les Rohingyas.
https://www.amnesty.org/en/documents/ASA16/5933/2022/en/
[42] https://ec.europa.eu/info/live-work-travel-eu/coronavirus-response/fighting-disinformation/tackling-coronavirus-disinformation_en
[43] https://www.who.int/europe/news/item/01-09-2022-infodemics-and-misinformation-negatively-affect-people-s-health-behaviours--new-who-review-finds

souffrant des effets économiques dévastateurs de la crise, les études ci-dessus ont montré que les médias sociaux ont joué un rôle important dans la propagation de rumeurs et la spéculation sur le contenu lié à la santé pendant les pandémies. Quatre études ont examiné la proportion de désinformation sur la santé sur les médias sociaux et ont constaté qu'elle atteignait jusqu'à 51% dans les publications associées aux vaccins, jusqu'à 28% dans les publications associées à COVID-19 et jusqu'à 60% dans les publications liées aux pandémies. Parmi les vidéos YouTube sur les maladies infectieuses émergentes, 20 à 30% contenaient des informations inexactes ou trompeuses.

D'autre part, il a été constaté que les médias sociaux et les médias traditionnels, lorsqu'ils sont utilisés adéquatement, pourraient être utiles pendant la communication de crise et pendant les pandémies de maladies infectieuses émergentes. Les médias sociaux peuvent améliorer l'acquisition de connaissances, la sensibilisation, la conformité et les comportements positifs à l'égard de l'adhésion aux protocoles et aux comportements cliniques en matière d'infection, et leur rôle était particulièrement important à une époque où les citoyens sont invités à rester à la maison. Plus que jamais, l'importance de soutenir un journalisme de qualité basé sur des informations vérifiées et des mesures pour lutter contre la propagation et les effets néfastes de la désinformation afin de protéger l'espace médiatique hors ligne et en ligne est devenue évidente.

Les efforts visant à lutter contre la désinformation, tels que les campagnes de vérification des faits, les initiatives volontaires entreprises par les plateformes en ligne, les mesures réglementaires et législatives, les efforts d'éducation aux médias et à l'information, etc., seront développés ci-dessous.

2.3. Mesures de lutte contre la désinformation

Le problème du désordre de l'information a été largement reconnu à tous les niveaux : par les États, par la société civile et même par les plateformes elles-mêmes. Cependant, l'absence d'un consensus plus large sur la manière de traiter cette question et sur les responsabilités a donné lieu à des approches diverses et parfois divergentes, tant au niveau européen que national, allant de l'autorégulation par les plateformes en ligne aux approches de corégulation, en passant par les mesures d'identification et de suppression imposées par les États[44].

2.3.1. Vérification des faits

La vérification des faits est un processus qui cherche à vérifier des informations factuelles, afin de promouvoir la véracité et l'exactitude des contenus. Cette pratique s'est considérablement développé ces dernières années en tant que moyen de lutter contre la

[44] McGonagle, T. et Katie Pentney K. (2021), p. 43.

désinformation, alimentée par les préoccupations concernant les informations fausses et trompeuses et les théories du complot diffusées sur les médias sociaux, mais aussi dans les médias grand public. De nombreux médias professionnels et agences de presse ont mis en place des équipes de vérification des faits, comme par exemple le Washington Post ou l'Agence France-Presse (AFP), dont l'équipe AFP Fact Check dispose d'un réseau mondial de journalistes surveillant le contenu en ligne dans différentes langues, en tenant compte des cultures et des politiques locales, et en travaillant à trouver et à démystifier les contenus faux, nuisibles et manipulateurs.

Les organisations indépendantes de vérification des faits sont également devenues importantes dans de nombreux pays. EU Disinfo Lab, par exemple, est une organisation indépendante à but non lucratif basée à Bruxelles qui se concentre sur la lutte contre les campagnes de désinformation sophistiquées ciblant l'Union européenne, ses États membres, ses institutions fondamentales et ses valeurs fondamentales[45]. En plus de surveiller et de démystifier les activités de désinformation sur les principales plateformes, ces organisations s'engagent dans des activités de recherche et de plaidoyer, identifient les tendances et les menaces et alertent les militants et les chercheurs ou formulent des recommandations politiques.

Le rôle des organisations de vérification des faits dans la lutte contre la désinformation a été reconnu au niveau de l'Union européenne, qui a intensifié ses efforts pour soutenir les vérificateurs de faits et les chercheurs européens sur la désinformation. En 2020, il a contribué à la création de l'Observatoire européen des médias numériques (EDMO) [46], qui sert de plaque tournante européenne permettant aux vérificateurs de faits, aux universitaires et aux chercheurs de collaborer entre eux et d'établir des liens actifs avec les organisations médiatiques et les experts en éducation aux médias, et de fournir un soutien aux décideurs politiques. Il a mis en place plusieurs pôles régionaux dans différents États membres, dont la Belgique, chacun de ces pôles servant de communauté multidisciplinaire en rassemblant des chercheurs universitaires, des vérificateurs de faits, des professionnels des médias et d'autres parties prenantes concernées pour créer un réseau afin de :

- détecter et analyser les campagnes de désinformation, ainsi que produire du contenu pour aider les médias grand public et locaux et les autorités publiques à dénoncer les campagnes de désinformation préjudiciables ;
- organiser des activités d'éducation aux médias au niveau national ou multinational ;
- fournir un soutien aux autorités nationales pour le suivi des politiques des plateformes en ligne et de l'écosystème des médias numériques[47].

[45] https://www.disinfo.eu/
[46] https://edmo.eu/
[47] https://digital-strategy.ec.europa.eu/en/news/eight-proposals-selected-become-first-hubs-european-digital-media-observatory

L'EDMO joue également un rôle important dans la mise en œuvre du Code de bonnes pratiques renforcé sur la désinformation (qui sera développé plus en détail ci-dessous). En outre, le rôle particulier des vérificateurs de faits en général est reconnu dans le nouveau cadre législatif de l'Union européenne pour les services numériques.

Au cours des dernières années, nous avons assisté à la propagation de sites internet autodéclarés de vérification des faits ou de comptes de médias sociaux. Comme l'avertit EDMO[48], le but de ces faux fact-checkers est de semer la confusion et le doute dans toutes les informations qui apparaissent dans l'espace public, de sorte qu'il serait de plus en plus difficile pour les citoyens d'évaluer ce qui est vrai et ce qui est faux. Cette « militarisation de la vérification des faits » est particulièrement présente depuis le début de l'agression russe : EDMO cite l'exemple d'un site russe qui publie quelques corrections précises, mais démystifie également des faux inexistants prétendument publiés par la partie ukrainienne, en créant par exemple des images fabriquées de chars brûlés.

Il a été démontré que la publication d'informations vérifiées a généralement un effet positif en termes de correction d'informations inexactes. Cependant, cet effet est plus faible dans les contextes polarisés (comme pendant les campagnes électorales) et parmi certains publics (partisans ayant des convictions profondes) [49]. La recherche en sciences sociales montre que, parce que les vérificateurs de faits proviennent souvent de groupes idéologiquement et socialement opposés, il peut y avoir un grave manque de confiance de la part du groupe cible réel. Cela peut rendre leur utilisation inefficace ou même contre-productive[50].

Une étude examinant dans quelle mesure la vérification des faits pourrait corriger certaines des affirmations inexactes et trompeuses propagées par Marine Le Pen lors de l'élection présidentielle française de 2017, a également montré que la vérification des faits peut améliorer la précision des connaissances factuelles du public, mais qu'elle a beaucoup moins d'impact sur leurs croyances. Même armés des faits, les auditoires peuvent continuer à croire des arguments fondés sur des informations incorrectes ou à agir selon ces croyances antérieures, ou, comme l'étude l'a conclu, *« le succès dans la correction des connaissances factuelles ne se traduit pas par un impact sur les intentions de vote »*[51].

Le dernier cas en date inclut le Brésil, dont les élections présidentielles de 2018 servent d'exemple d'une campagne de désinformation agressive qui a inondé le public avec les grandes quantités d'histoires fabriquées qui se répandent sur Facebook et WhatsApp.

---

[48] https://edmo.eu/2022/03/17/russian-propaganda-disguising-as-fact-checking-a-statement-from-the-edmo-taskforce/
[49] https://akademie.dw.com/en/is-fact-checking-effective-a-critical-review-of-what-works-and-what-doesn't/a-55248257
[50] Joris van Hoboken et al. (2019).
[51] https://akademie.dw.com/en/is-fact-checking-effective-a-critical-review-of-what-works-and-what-doesn't/a-55248257

Alors que le pays se préparait aux élections présidentielles de 2022, les stratégies de désinformation sont également devenues plus sophistiquées. Les rapports concluent que de nombreux efforts sur la professionnalisation des médias, des initiatives concentrées d'institutions démocratiques et d'organes de presse indépendants visant à lutter contre les contenus trompeurs menaçant l'intégrité du processus électoral national, y compris l'engagement pris par les principales plateformes numériques d'imposer des contrôles plus stricts sur la diffusion de la désinformation, *« ne font qu'effleurer la surface »* du problème[52].

2.3.2. Autorégulation

L'absence de cadre législatif et de surveillance réglementaire des plateformes en ligne signifiait que ces entreprises devaient s'autoréguler, mais aussi décider quelles sources sont crédibles, et donc plus visibles, c'est-à-dire mieux classées dans l'organisation et la présentation du contenu aux utilisateurs individuels.

Lors de la création d'un compte avec l'un des médias sociaux aujourd'hui, les utilisateurs doivent accepter les conditions d'utilisation, un contrat qui spécifie les conditions dans lesquelles l'utilisateur et la plate-forme interagissent. Il est accompagné d'une page de « Lignes directrices de la communauté » qui explique aux utilisateurs quel type de comportement et de contenu est autorisé et ce qui est approprié et ce qui ne l'est pas, ainsi que d'énumérer les interdictions dont les violations peuvent entraîner la suppression de contenu ou de page, les comptes étant complètement désactivés et, dans certains cas, signalés aux autorités. Ces règles entrent généralement dans les catégories suivantes de contenu interdit ou restreint : contenu sexuel (nudité, sexe, pornographie), contenu graphique (violence et obscénité), harcèlement (abus, trollage, menaces directes), discours de haine, automutilation, activité illégale et désinformation. Les plateformes en ligne s'appuient sur les conditions d'utilisation et les lignes directrices de la communauté pour réglementer le comportement des utilisateurs et fonder leurs pratiques de modération en ligne de contenus illégaux. Ce qui rend ces documents problématiques du point de vue de la liberté d'expression, c'est qu'ils sont souvent plus stricts dans l'identification des contenus illicites en ligne à supprimer que les lois nationales ou les tribunaux des pays dans lesquels ils fournissent leurs services[53].

Les principaux outils utilisés par les plateformes en ligne pour identifier les contenus illicites en ligne sont le signalement par les utilisateurs, les mots-clés/filtres et les outils d'intelligence artificielle (IA). Cependant, les outils automatisés ont leurs limites en termes de précision, ce qui conduit à de nombreux choix erronés en termes de retrait de contenus parfaitement légitimes, ou de non-reconnaissance de contenus illicites, et doivent donc

[52] https://reutersinstitute.politics.ox.ac.uk/news/despite-efforts-fight-falsehoods-brazils-tight-election-threatened-dangerous-lies?utm_campaign=Future%20of%20Journalism&utm_medium=email&utm_source=Revue%20newsletter

[53] https://www.europarl.europa.eu/RegData/etudes/STUD/2020/652718/IPOL_STU(2020)652718_EN.pdf

souvent s'accompagner d'une modération humaine. Cela ouvre encore de la place aux erreurs, compte tenu des grandes quantités de contenu mise en ligne chaque seconde et d'un nombre limité de modérateurs humains qui disposent d'un temps très limité pour examiner chaque contenu, sans parler des défis posés par le contexte, les différentes langues et les normes culturelles.

Il existe de nombreux exemples d'initiatives d'autoréguation prises par les plateformes en ligne au fil des ans. L'un des exemples les plus connus est le programme de vérification des faits tiers de Meta[54], dans lequel cette société travaille avec un réseau mondial (plus de 90 organisations travaillant dans plus de 60 langues) de partenaires de vérification des faits qui examinent et évaluent indépendamment la désinformation potentielle sur Facebook, Instagram et WhatsApp. En outre, cette société a fondé le Conseil de surveillance[55] en 2020, un organisme externe composé d'experts indépendants auxquels les gens peuvent faire appel s'ils ne sont pas d'accord avec les décisions relatives au contenu de Meta sur Facebook ou Instagram. Le panel examine les cas sur la base des politiques et des valeurs de contenu de Facebook, tout en tenant compte des normes relatives aux droits de l'homme qui protègent la liberté d'expression. Les décisions prises par le conseil de surveillance sont contraignantes pour Facebook, rendues publiques et clairement justifiées.

Pendant la pandémie de Covid-19, plusieurs plateformes en ligne ont intensifié leurs efforts pour lutter contre la désinformation et mis en œuvre des mesures liées à la suppression de contenus (par exemple, suppression de contenus décourageant les personnes de suivre un traitement ou affirmant que des substances dangereuses sont saines), à la promotion d'informations crédibles, faisant autorité et pertinentes (par exemple, affichage dans les principales mises à jour du fil d'actualité de l'OMS et des autorités sanitaires nationales, réduction des possibilités de publicité en ligne comme le blocage de la publicité capitalisant sur la pandémie). En ce qui concerne la manipulation des élections, la plupart des plateformes ont commencé à marquer les messages politiques en tant que tels, ont ouvert des canaux permettant aux utilisateurs de signaler des contenus trompeurs et ont lancé des campagnes d'éducation aux médias. WhatsApp, par exemple, a limité les interactions des utilisateurs pour arrêter la messagerie de masse sans discernement[56].

Même si beaucoup a été fait ces dernières années pour améliorer les systèmes internes de prévention de la manipulation et de la diffusion de contenus faux et malveillants, beaucoup s'accordent à dire que ces entreprises privées, dont les modèles économiques sont basés sur le profit, ont « privatisé » la réglementation des contenus en ligne préjudiciables, ce qui est problématique du point de vue de la protection de l'intérêt public (efforts insuffisants pour prévenir les dommages, manque de transparence sur les

[54] https://www.facebook.com/formedia/mjp/programs/third-party-fact-checking
[55] https://www.oversightboard.com/
[56] https://reutersinstitute.politics.ox.ac.uk/news/despite-efforts-fight-falsehoods-brazils-tight-election-threatened-dangerous-lies

pratiques de modération des contenus), ainsi que la protection de la liberté d'expression (censure privée, suppression excessive de contenus). Ce sont précisément les lacunes que les mesures législatives introduites tant au niveau national qu'au niveau de l'Union européenne ont cherché à surmonter.

2.3.3. Approches nationales

Certains pays ont des lois spécifiques qui interdisent la désinformation ou les fausses informations. Le meilleur exemple est la France, qui a introduit une législation visant spécifiquement les fausses informations pendant les élections. En vertu de la loi de 2018 relative à la lutte contre la manipulation de l'information, pendant les trois mois précédant une élection, un juge peut ordonner d'urgence (dans les 48 heures), toutes mesures proportionnées et nécessaires pour faire cesser la diffusion de *« toute allégation ou accusation d'un fait inexact ou trompeur susceptible d'altérer la sincérité du vote à venir, qui sont délibérément diffusés de manière automatisée et massive par l'intermédiaire d'un service de communication publique en ligne »*. Le premier jugement du tribunal en vertu de cette disposition a précisé que le contenu doit être du contenu sponsorisé, c'est-à-dire le paiement de tiers pour élargir artificiellement la diffusion de l'information, et du contenu promu à l'aide d'outils automatisés tels que les robots[57].

Certains pays ont imposé, ou prévoient d'imposer, de nouvelles obligations législatives de diligence aux services en ligne afin de mettre en œuvre des mesures de lutte contre la désinformation. La loi française précitée impose aux grandes plateformes en ligne l'obligation de lutter contre la diffusion de fausses informations susceptibles de troubler l'ordre public ou d'altérer la sincérité des élections, en mettant en place certains mécanismes permettant aux utilisateurs de signaler de fausses informations, et comprend également des mesures que les plateformes peuvent mettre en place, telles que la promotion de contenus provenant de certaines entreprises de médias. De même, la proposition du gouvernement britannique sur le Livre blanc sur les services en ligne implique une nouvelle obligation légale de diligence pour les services en ligne de prendre des *« mesures raisonnables »* pour assurer la sécurité des utilisateurs et lutter contre les activités illégales et nuisibles. En ce qui concerne la désinformation, les services en ligne seraient tenus de prendre des *« mesures proportionnées et proactives »* pour minimiser la propagation de la désinformation trompeuse et préjudiciable et *« accroître l'accessibilité de contenus d'information fiables et variés »*. Fondamentalement, l'obligation légale de diligence serait appliquée par un organisme de réglementation indépendant, et l'organisme de réglementation définirait la manière dont les services en ligne rempliraient leurs obligations légales par le biais de nouveaux codes de conduite. Notamment, l'organisme de réglementation aurait des pouvoirs de décisions

[57] Joris van Hoboken et al. (2019), p. 104.

considérables, y compris de prononcer des amendes, et peut-être le pouvoir de bloquer des plateformes, comme option d'application de dernière instance[58].

La législation nationale applicable à la désinformation est dans certains cas de nature pénale, comme celles que certains États membres de l'Union européenne ont introduites dans le contexte du Covid-19. En Hongrie, par exemple, le Code pénal a été modifié pour inclure des déclarations fausses ou téméraires dans l'intention d'entraver ou d'empêcher l'efficacité des mesures de protection. La Commission européenne a toutefois averti que les lois applicables aux concepts de désinformation qui sont « trop larges » soulèvent des préoccupations particulières pour la liberté d'expression. Ces préoccupations sont également reprises par les experts, car la désinformation est un *« concept extraordinairement insaisissable à définir en droit »* et est *« susceptible de donner aux autorités exécutives un pouvoir discrétionnaire excessif pour déterminer ce qui est de la désinformation, ce qui est une erreur, ce qui est la vérité »*[59].

2.3.4. Approches transnationales

2.3.4.1. Conseil de l'Europe

Le Conseil de l'Europe attache une grande importance aux activités de recherche et de normalisation en matière de lutte contre le désordre de l'information en aidant à définir les différentes notions qui y sont attachées, en renforçant le journalisme de qualité, en favorisant un environnement médiatique indépendant et pluraliste et en renforçant l'éducation numérique et médiatique des utilisateurs d'internet.

L'un des documents les plus récents adoptés à cet égard est la Recommandation CM/Rec(2022)4 du Comité des Ministres aux Etats membres sur la promotion d'un environnement favorable au journalisme de qualité à l'ère numérique[60]. Reconnaissant les menaces croissantes que fait peser sur les démocraties la propagation de campagnes de désinformation et de propagande en ligne, ainsi que « *le passage à un environnement de plus en plus numérique, mobile et de médias sociaux a profondément modifié la dynamique de production, de diffusion et de consommation des informations et autres contenus médiatiques, et notant que, de ce fait, le journalisme de qualité rivalise pour attirer l'attention du public avec d'autres types de contenus qui ne sont pas soumis à la même législation, cadres réglementaires ou éthiques »,* il énonce un certain nombre de lignes directrices sur la promotion d'un journalisme de qualité à l'ère numérique. Il s'agit notamment d'investir dans la production d'un journalisme de qualité et d'assurer la viabilité financière du journalisme, y compris des mesures visant à assurer un partage

[58] Ibid.
[59] https://erga-online.eu/wp-content/uploads/2020/05/ERGA-2019-report-published-2020-LQ.pdf
[60] Recommandation CM/Rec(2022)4 du Comité des Ministres aux États membres sur la promotion d'un environnement favorable à un journalisme de qualité à l'ère du numérique. https://search.coe.int/cm/pages/result_details.aspx?ObjectId=0900001680a5ddd1

équitable des recettes de marketing et de publicité entre les organisations de médias qui produisent du contenu et les grandes plateformes en ligne et autres intermédiaires internet pertinents qui bénéficient de manière significative de sa distribution, ainsi que la transparence de la publicité. En outre, l'accent est mis sur le rôle du journalisme dans le processus de vérification des faits et dans le rétablissement et le maintien de la confiance et des relations saines avec le public et les contributeurs de contenu médiatique. Il est reconnu que la désinformation sape la confiance dans les médias et menace la fiabilité de l'information qui alimente le débat public et la démocratie, de sorte que les Etats membres du Conseil de l'Europe sont appelés à soutenir pleinement les efforts nationaux et/ou transnationaux concertés pour lutter contre la désinformation et la propagande sur une base continue, et pas seulement pendant les campagnes électorales : « *Alors que la manipulation de l'information alimente les divisions et les tensions, le renforcement de la résilience et de la cohésion des sociétés devrait être un objectif européen à long terme. Une société bien informée et éduquée aux médias (y compris les journalistes, les médias, les plateformes en ligne, les organisations non gouvernementales et les particuliers) est un élément essentiel de la défense contre la manipulation de l'information dans les sociétés démocratiques ».*

2.3.4.2. Union européenne

- **Code de bonnes pratiques sur la désinformation (2018)**

L'Union européenne a commencé à accorder plus d'attention à la question de la désinformation en 2018, lorsque le premier instrument d'autorégulation sur ce sujet au monde – le Code de bonnes pratiques contre la désinformation[61] – a été approuvé et signé par des représentants de l'industrie. Le Code de bonnes pratiques a été signé en octobre 2018 par les plateformes en ligne Facebook, Google, Twitter et Mozilla, ainsi que par des annonceurs et d'autres acteurs de l'industrie publicitaire. Microsoft a rejoint en mai 2019, tandis que TikTok a signé le Code en juin 2020. Cette initiative a été prise dans le contexte d'un éventail plus large d'efforts déployés par l'Union européenne pour lutter contre la désinformation en ligne, y compris la communication de la Commission européenne susmentionnée et le plan d'action contre la désinformation[62].

Le code comprend un ensemble de 15 engagements organisés en cinq domaines:
1. Examen minutieux des placements publicitaires (visant à démonétiser les fournisseurs de désinformation en ligne).
2. Publicité politique et publicité thématique (visant à s'assurer que les publicités politiques sont clairement identifiées par les utilisateurs).
3. Intégrité des services (visant à identifier et à fermer les faux comptes et à utiliser des mécanismes appropriés pour signaler les interactions pilotées par des robots).

[61] https://digital-strategy.ec.europa.eu/en/library/2018-code-practice-disinformation
[62] https://ec.europa.eu/info/publications/action-plan-disinformation-commission-contribution-european-council-13-14-december-2018_en

4. Responsabiliser les consommateurs (visant à diluer la visibilité de la désinformation en améliorant la repérabilité des contenus fiables et en facilitant la découverte et l'accès des utilisateurs à différentes sources d'information représentant des points de vue différents).
5. Responsabiliser la communauté des chercheurs (visant à donner aux chercheurs l'accès aux données des plateformes qui sont nécessaires pour surveiller en permanence la désinformation en ligne).

Le suivi de la mise en œuvre du code s'est déroulé en deux phases: la première phase entre janvier et mai 2019 visait à suivre la mise en œuvre des engagements du Code qui revêtent une importance particulière pour l'intégrité des élections européennes. Cette surveillance a révélé que toutes les publicités politiques diffusées sur les plateformes n'étaient pas correctement étiquetées comme telles. Au cours de la deuxième phase, une évaluation complète de la mise en œuvre des engagements des cinq piliers du Code a été réalisée, avec l'aide de l'ERGA. Bien que certaines plateformes aient clairement déployé des efforts pour se conformer à ses mesures, l'ERGA a néanmoins identifié des faiblesses importantes qui doivent être corrigées pour que le Code puisse atteindre ses objectifs à l'avenir. À cet égard, l'ERGA a identifié des recommandations concrètes comme moyen d'aller de l'avant[63] :

- *« Le modèle actuel d'autorégulation s'est avéré être une première étape importante et nécessaire, mais il est nécessaire d'être plus efficace pour lutter contre la désinformation en ligne.*
- *Le nombre de signataires du code est limité et n'inclut pas toutes les plateformes importantes, les services d'information et de communication et les acteurs du secteur de la publicité actifs dans l'UE.*
- *Il est nécessaire d'accroître la transparence, y compris des données beaucoup plus détaillées (en particulier des données spécifiques à chaque pays) sur la manière dont les signataires appliquent le Code; de plus, certaines des mesures prévues par le Code sont de nature trop générale et ne sont pas adoptées uniformément par tous les signataires ».*

Dans cette optique, il a été proposé de passer d'une approche d'autorégulation souple à une approche de corégulation, dans laquelle les autorités de régulation et la Commission européenne disposeraient de pouvoirs de surveillance et d'exécution. Le processus de renforcement du Code a été entrepris dans le contexte du nouveau cadre juridique complet de l'Union européen pour les services intermédiaires en ligne: le Règlement sur les services numériques (Digital Services Act – DSA)[64].

- **Règlement sur les services numériques**

La législation sur les services numériques (DSA) est un nouvel acte législatif de l'Unuin européenne qui établit des normes pour les plateformes en ligne et protège mieux les

[63] ERGA (2020).
[64] https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:32022R2065

utilisateurs en ligne, y compris une structure de responsabilité garantissant que les fournisseurs de services en ligne seront tenus responsables de leurs pratiques de modération de contenu dans l'espace numérique. Il est entré en vigueur le 16 novembre 2022 et s'appliquera à partir du 17 février 2024, bien que les règles spécifiques aux « très grandes plateformes en ligne » entreront en vigueur beaucoup plus tôt.

Le DSA s'intéresse principalement aux contenus illicites et limite les mécanismes de suppression de contenus aux seuls contenus illicites, mais il prévoit des mécanismes spécifiques visant à assurer la protection des utilisateurs contre les contenus qui ne sont pas illégaux en soi, mais qui présentent des risques pour les droits fondamentaux, y compris la désinformation. Elle introduit une nouvelle approche de la modération des contenus fondée sur les risques, en imposant des obligations de diligence raisonnable spéciales aux très grandes plateformes en ligne (VLOP) et aux très grands moteurs de recherche (VLSE), qui sont ceux qui fournissent des services à 45 millions d'utilisateurs actifs ou plus dans l'Union, ou couvrant 10 % de la population de l'Union, compte tenu de leur large portée et donc de leur impact potentiel.

Ces plateformes auront l'obligation de prendre des mesures pour identifier, analyser et évaluer, au moins une fois par an, la probabilité et la gravité de tout risque systémique important découlant de leur conception, de leurs systèmes algorithmiques, de leurs caractéristiques intrinsèques, de leur fonctionnement et de l'utilisation de leurs services dans l'Union. En ce qui concerne la désinformation, les risques systémiques comprennent *« tout dysfonctionnement ou manipulation intentionnelle de leur service, y compris au moyen d'une utilisation non authentique ou d'une exploitation automatisée du service ou les risques inhérents au fonctionnement prévu du service, y compris l'amplification de contenus illicites, de contenus contraires à leurs conditions générales ou de tout autre contenu ayant un effet négatif réel ou prévisible sur la protection des mineurs et d'autres groupes vulnérables destinataires du service, sur les valeurs démocratiques, la liberté des médias, la liberté d'expression et de discours civique, ou les effets réels ou prévisibles liés aux processus électoraux et à la sécurité publique » ainsi que « tout effet négatif réel et prévisible sur la protection de la santé publique ainsi que les dépendances comportementales ou autres conséquences négatives graves sur l'intégrité physique de la personne, le bien-être mental, social et financier »*.

Même si la DSA ne fait pas spécifiquement référence à la désinformation dans ce contexte, la formulation ci-dessus vise clairement à répondre aux principales préoccupations du désordre de l'information telles que l'infodémie et la manipulation électorale. Il est important de noter que les plateformes devront tenir compte en particulier de la question de savoir si et comment leurs systèmes de modération de contenu, leurs conditions générales, leurs lignes directrices de la communauté, leurs systèmes algorithmiques, leurs systèmes de recommandation et leurs systèmes de sélection et d'affichage de publicités, ainsi que la collecte, le traitement et le profilage des données sous-jacents, influencent les risques systémiques susmentionnés.

Si de tels risques systémiques sont identifiés, les plateformes doivent prendre des mesures d'atténuation, notamment « *adapter la modération de contenu, les systèmes algorithmiques ou les systèmes de recommandation et les interfaces en ligne, leurs processus décisionnels, la conception, les caractéristiques ou le fonctionnement de leurs services, leur modèle publicitaire ou leurs conditions générales* ». Là encore, la modération automatisée et le modèle commercial sont considérés comme essentiels pour prévenir les abus.

Contrairement à l'approche d'autorégulation adoptée auparavant, le DSA va encore plus loin en imposant des exigences en matière d'atténuation des risques, mais garantit également des obligations de surveillance et de transparence. L'obligation d'identifier et d'atténuer les risques est renforcée par l'obligation pour les VLOP et les VLSE de faire l'objet d'audits annuels indépendants afin d'évaluer le respect, entre autres, de leurs obligations d'atténuer les risques systémiques, ainsi que de proposer des mesures en cas de non-respect. En outre, le DSA prévoit des obligations de déclaration de transparence pour ces fournisseurs, notamment pour fournir un aperçu de la boîte noire des algorithmes et expliquer leur conception, leur logique et leur fonctionnement, si cela est demandé.

Enfin, le DSA introduit un mécanisme de surveillance, puisque la Commission européenne évaluera la mise en œuvre et l'efficacité des mesures d'atténuation et émettra des recommandations lorsque les mesures mises en œuvre sont jugées inappropriées ou inefficaces pour faire face au risque systémique en jeu.

Il reste à voir comment ces mécanismes fonctionneront dans la pratique, compte tenu de la complexité des enjeux. En outre, les limites du DSA dans la lutte contre le phénomène mondial de la désinformation ont été reconnues puisqu'il n'applique son approche fondée sur les risques que sur de très grandes plateformes, alors que les petits réseaux ne sont pas couverts par cette obligation – et ces petits réseaux, dans leurs chambres d'écho, en particulier impliquant des théoriciens du complot, peuvent causer toute une série de préjudices[65].

- **Code de bonnes pratiques renforcé sur la désinformation (2022)**

Comme nous l'avons mentionné, c'est dans le contexte de l'adoption du DSA que le nouveau Code renforcé a été élaboré et élevé au rang d'instrument de corégulation. Non limité aux très grandes plateformes mais, au contraire, visant à inclure le plus large éventail possible de signataires, il a été signée en juin 2022 par un éventail d'acteurs beaucoup plus large que sa version précédente, tels que les plateformes en ligne, les acteurs de l'écosystème publicitaire, les vérificateurs de faits, la société civile, la

[65] McGonagle, T. et Katie Pentney K. (2021), p. 52.

recherche et d'autres organisations[66]. Il contient 44 engagements et 128 mesures spécifiques, dont[67] :

Démonétisation de la désinformation. Le Code renforcé vise à faire en sorte que les pourvoyeurs de désinformation ne bénéficient pas des recettes publicitaires. Les signataires s'engagent à prendre des mesures plus strictes pour éviter de placer de la publicité à côté de la désinformation, ainsi que la diffusion de publicité contenant de la désinformation.

Transparence de la publicité politique. Reconnaissant l'importance de la publicité politique dans l'orientation de la vie publique, le Code renforcé engage les signataires à mettre en place des mesures de transparence plus strictes, permettant aux utilisateurs de reconnaître facilement les publicités politiques en fournissant un étiquetage plus efficace, en s'engageant à révéler le sponsor, les dépenses publicitaires et la période d'affichage. De plus, les signataires s'engagent à mettre en place des bibliothèques publicitaires efficaces et consultables pour la publicité politique.

Assurer l'intégrité des services. Le Code renforcera les mesures visant à réduire les comportements manipulateurs utilisés pour diffuser de la désinformation (par exemple, les faux comptes, l'amplification pilotée par des robots, l'usurpation d'identité, les deep fakes malveillants) et établit une coopération plus étroite entre les signataires pour lutter contre les défis liés à ces techniques.

Responsabiliser les utilisateurs. Les utilisateurs seront mieux protégés contre la désinformation grâce à des outils améliorés pour reconnaître, comprendre et signaler la désinformation, pour accéder à des sources faisant autorité et par des initiatives d'éducation aux médias. En particulier, le Code veillera à ce que des pratiques de conception sûres soient mises en place pour limiter la propagation de la désinformation et assurer une plus grande transparence de leurs systèmes de recommandation, en les adaptant pour limiter la propagation de la désinformation.

Responsabiliser les chercheurs. Le Code prévoit que les plateformes en ligne soutiennent mieux la recherche sur la désinformation. Les chercheurs auront un accès meilleur et plus large aux données des plateformes. Cela implique de garantir un accès automatisé aux données non personnelles, anonymisées, agrégées ou manifestement rendues publiques, et de s'employer à mettre en place une structure de gouvernance pour simplifier l'accès aux données nécessitant un examen supplémentaire.

Responsabiliser la communauté de la vérification des faits. Le nouveau code étendra la couverture de la vérification des faits à tous les États membres et dans toutes les langues de l'Union et garantira que les plateformes feront un usage plus cohérent de la vérification des faits sur leurs services. En outre, le Code vise à garantir des contributions financières équitables pour le travail des vérificateurs de faits et un meilleur accès aux informations facilitant leur travail quotidien.

[66] https://digital-strategy.ec.europa.eu/en/library/signatories-2022-strengthened-code-practice-disinformation

[67] https://digital-strategy.ec.europa.eu/en/policies/code-practice-disinformation

Sa nature de corégulation découle du fait que la Commission européenne surveillera et évaluera régulièrement la réalisation des objectifs du code, avec l'aide de l'ERGA et de l'EDMO.

- **Le règlement européenn sur la liberté des médias**

Le dernier membre de la famille nombreuse des récentes propositions législatives de l'Union européenne concernant les médias est arrivé à la mi-septembre 2022, lorsque la Commission européenne a publié une proposition de règlement sur la liberté des médias (European Media Freedom Act - EMFA), dont l'objectif global est de renforcer la liberté, l'indépendance et le pluralisme des médias.

Des médias indépendants et pluralistes, libres de toute ingérence extérieure, sont des piliers essentiels des systèmes démocratiques. Ils permettent aux citoyens de se forger une opinion libre et éclairée sur des questions d'intérêt public, de demander des comptes aux personnes au pouvoir et de favoriser le débat démocratique. Cependant, il y a des tendances de plus en plus inquiétantes qui sapent cette liberté, de l'ingérence politique et commerciale dans les décisions éditoriales, des financements instables, de la surveillance et des menaces contre les journalistes... Entre autres choses, cela influence fortement la qualité du journalisme et, en fin de compte, la mesure dans laquelle les médias eux-mêmes contribuent à la propagation de la désinformation.

Il est envisagé que le régime EMFA s'applique à tous les services de médias, publics et privés, quels que soient les moyens utilisés pour leur diffusion. Cela inclut à la fois les médias hors ligne et en ligne, par exemple les éditions en ligne de journaux. Les mesures proposées comprennent[68] :

- La protection de la liberté éditoriale et de l'indépendance des médias (interdiction de l'ingérence directe ou indirecte dans les décisions éditoriales; protection des sources journalistiques et des journalistes, y compris les garanties contre l'utilisation de logiciels espions contre les médias, les journalistes et leurs familles).
- Des garanties pour le fonctionnement indépendant des médias de service public (financement adéquat et stable; procédures de nomination transparentes).
- La transparence de la propriété des médias (divulgation publique des informations sur la propriété).
- L'obligation pour les États membres d'évaluer l'incidence des concentrations du marché des médias sur le pluralisme des médias et l'indépendance éditoriale.
- L'allocation transparente et équitable des ressources économiques (allocation de la publicité d'État telle que la publicité, qui a été utilisée par les politiciens pour subventionner et récompenser des médias amis dans le passé; transparence et objectivité des systèmes de mesure d'audience, qui ont un impact sur les revenus publicitaires des médias, en particulier en ligne).

[68] https://ec.europa.eu/commission/presscorner/detail/en/ip_22_5504

- Des obligations supplémentaires sur les plateformes en ligne concernant les contenus médiatiques professionnels. À cet égard, la proposition de règlement EMFA s'appuie sur le DSA, qui n'exemptait pas les contenus des médias professionnels de l'application de ses dispositions. Toutefois, l'EMFA, en tant que *lex specialis*, comprend des garanties supplémentaires contre la suppression injustifiée de contenus médiatiques professionnels, ainsi que l'obligation d'étiqueter les services de médias et de protéger leur distribution, reconnaissant l'impact considérable que les systèmes de recommandation des plateformes ont sur le public des médias.

Les réformes proposées comprennent un nouveau Conseil européen des services de médias composé d'autorités de régulation des médias, pour remplacer l'ERGA. Le comité encouragera l'application efficace et cohérente du cadre législatif de l'Union européenne sur les médias et aidera la Commission européenne à élaborer des lignes directrices sur les questions de réglementation des médias. Ses tâches consisteront notamment à organiser un dialogue structuré entre les très grandes plateformes en ligne et le secteur des médias afin de promouvoir l'accès à diverses offres médiatiques et de contrôler le respect par les plateformes des initiatives d'autorégulation, telles que le Code de bonnes pratiques contre la désinformation.

2.3.5. Éducation aux médias et à l'information

La nécessité de renforcer l'éducation aux médias et à l'information (EMI) comme moyen de limiter la diffusion et l'impact de la désinformation en renforçant la résilience du public face aux campagnes de désinformation est reconnue dans les politiques publiques du monde entier. L'EMI est un concept complexe introduit par l'UNESCO en 2007, englobant un ensemble de connaissances, d'aptitudes, d'attitudes, de compétences et de pratiques qui permettent aux personnes d'accéder, d'analyser, d'évaluer de manière critique, d'interpréter, d'utiliser, de créer et de diffuser efficacement des produits d'information et des médias sur une base créative, juridique et éthique en utilisant les ressources et outils existants. Il représente un ensemble interdépendant de compétences qui aident les gens à maximiser les avantages et à minimiser les dommages dans les nouveaux paysages de l'information, du numérique et de la communication[69].

Comprendre les médias et le contenu médiatique et être habilité à l'évaluer est devenu l'une des compétences les plus importantes pour vivre à l'ère numérique. Par-dessus tout, les compétences en EMI ont été largement reconnues comme l'un des outils les plus importants dans la lutte contre la désinformation, car elles permettent aux citoyens d'évaluer de manière critique et de prendre des décisions éclairées sur le contenu des médias, ainsi que de créer du contenu médiatique de manière responsable et sûre.

[69] UNESCO, About Media and Information Literacy, disponible à l'adresse : www.unesco.org/en/communication-information/media-information-literacy/about

L'EMI touche à divers domaines: comprendre les médias et leur contenu et être habilité à les évaluer, protection des mineurs et des utilisateurs (par exemple, leur vie privée), ainsi que les chances éducatives, la participation numérique et l'égalité des chances. De plus en plus, l'EMI consiste à faire valoir ses droits en tant que citoyen numérique dans une démocratie[70].

Comme indiqué dans la Recommandation du Conseil de l'Europe susmentionnée, l'EMI implique le développement de compétences et de capacités cognitives, techniques et sociales qui permettent aux personnes de :

- accéder efficacement au contenu médiatique et analyser de manière critique l'information, leur permettant ainsi de comprendre comment le contenu médiatique est produit, financé et réglementé, ainsi que d'avoir la confiance et la compétence nécessaires pour prendre des décisions éclairées sur les médias qu'ils utilisent et la manière dont ils les utilisent ;
- comprendre les implications éthiques des médias et de la technologie ;
- communiquer efficacement, notamment en interprétant, en créant et en publiant du contenu.

Étant donné que les médias numériques sont en constante évolution, le développement des compétences EMI est un apprentissage tout au long de la vie. L'éducation aux médias ne peut donc pas être limitée aux jeunes, mais doit également englober les adultes – y compris les personnes âgées – qui ne peuvent souvent pas suivre le rythme de la transformation numérique induite par l'évolution rapide des technologies des médias.

[70] https://erga-online.eu/wp-content/uploads/2021/12/ERGA-AG3-2021-Report-on-Media-Literacy.pdf