

# Économie de l'attention, économie de la désinformation ?

## I. Les modèles économiques des plateformes numériques favorisent les contenus toxiques

### A. L'économie de l'attention

**Les plateformes numériques sont fondées sur modèle économique reposant sur la captation de l'attention des utilisateurs à des fins de valorisation pécuniaire.** Les plateformes sont des marchés dis “bifaces” ou “multifaces”, elles agissent intermédiaires entre deux catégories d'acteurs, ou plus, comme les utilisateurs ou les annonceurs par exemple. Ces marchés multifaces se caractérisent par leurs effets de réseau : chaque utilisateur apporte une valeur à l'ensemble du réseau et la valeur du service pour un groupe d'acteurs dépend de la quantité de participants sur l'autre face du marché. Par exemple, les annonceurs auront tendance à proposer leurs publicités sur des plateformes comptant déjà beaucoup d'utilisateurs. De leur côté, les utilisateurs choisiront la plateforme avec le plus de leurs pairs. Ces effets de réseau tendent à la concentration des marchés : les marchés numériques se caractérisent ainsi souvent par un petit nombre d'acteurs entre lesquels la concurrence est extrêmement forte. **À cela s'ajoute une quantité exponentielle de contenus disponibles qui rend d'autant plus compliquée de maintenir l'attention des utilisateurs,** créant un cercle vicieux résumé par Emmanuel Kessous, Kevin Mellet et Moustafa Zouinar : *“Pour les entreprises, ces facteurs élèvent le seuil minimal d'intensité du signal requis pour obtenir l'attention des consommateurs. Paradoxalement, cette incitation individuelle à augmenter l'intensité du signal se traduit au niveau agrégé par une augmentation de l'abondance informationnelle, ce qui en retour incite les entreprises à augmenter de nouveau l'intensité de leurs signaux”*<sup>1</sup>.

**Pour garder les utilisateurs sur leur interface et continuer à engranger des revenus, les plateformes ont donc intérêt à leur proposer les contenus les plus susceptibles de leur plaire en se basant sur leurs usages passés.** Ainsi, l'économie de l'attention peut se définir comme *“un ensemble de dispositifs mis en œuvre afin d'extraire une valeur marchande à partir de la captation de l'attention des utilisateurs.”*<sup>2</sup>

Ce modèle économique n'est pas né avec les technologies numériques, mais ces dernières l'ont particulièrement exacerbé. **Le Conseil national du numérique identifie quatre ruptures majeures liées aux médias numériques**<sup>3</sup> :

- **L'immersion et l'ubiquité** : plus que des outils, les technologies numériques fondent un véritable mode de vie omniprésent ;
- **L'immédiateté et le “temps réel”** : l'information arrive en flux continu et rend difficile toute prise de recul ;
- **Le calcul, le câblage et les prédictions** : la collecte permanente de données affine la connaissance des préférences individuelles et permet de plus en plus de prédire ceux à venir, influençant en retour les choix des individus ;
- **Le modèle économique** : les individus ne sont plus seulement utilisateurs mais aussi producteurs en contribuant directement à l'optimisation des systèmes en place par le données et traces laissées en ligne.

Si elle est souvent associée à la publicité, **l'économie de l'attention recouvre plusieurs modèles économiques**, résumés ainsi par le Conseil national du numérique :

---

<sup>1</sup> KESSOUS, Emmanuel, MELLET, Kevin, et ZOUINAR, Moustafa. [L'économie de l'attention: entre protection des ressources cognitives et extraction de la valeur](#). *Sociologie du travail*, 2010, vol. 52, no 3, p. 359-373.

<sup>2</sup> <https://cnnumerique.fr/files/uploads/2022/Livres/CNNum-Votre-attention-sil-vous-plait.pdf>

<sup>3</sup> *Ibid.*

EXEMPLES	TYPES DE PLATEFORME	OBJECTIF
Publicité (services sans contrepartie monétaire)*		
Google, Yahoo, Bing	Moteurs de recherche	
Facebook, Instagram, Twitter, Snapchat	Réseaux sociaux	
Dailymotion	Plateformes de partage de vidéos	
YouTube		
Freemium*		
YouTube	Plateformes de partage de vidéos	Prolonger le temps passé sur le service
Candy Crush, Angry Birds	Jeux sur mobile gratuits avec achats intégrés	
Spotify, Deezer	Répertoires audiovisuels dématérialisés	
Abonnement payant*		
Netflix, Disney +	Répertoires audiovisuels dématérialisés	
Commission prélevée sur la vente de biens ou de services ou vente directe de biens ou de services*		
Uber, Deliveroo, Airbnb	Plateformes collaboratives	Aboutir à un acte d'achat, peu importe le temps passé sur la plateforme
Amazon, Booking	Places de marché	

\* Modèle économique principal

Au regard du périmètre de l'espace informationnel, la présente note se concentrera sur les trois premiers modèles économiques partageant l'objectif de prolonger le temps passé sur le service.

Ce modèle économique est servi par le design des interfaces des plateformes, leur modération et leurs systèmes de recommandation algorithmique (voir fiche dédiée)

## B. Une prépondérance des contenus toxiques, dont les fausses informations

**Le modèle économique de la captation de l'attention n'est pas neutre pour les utilisateurs. Au contraire, il tend à valoriser les contenus toxiques pour l'utilisateur.** Karen Hao, journaliste pour le MIT Technology Review, l'explique parfaitement : *“Les modèles qui maximisent l'engagement favorisent également la controverse, la désinformation et l'extrémisme : pour faire simple, les gens aiment les choses scandaleuses.”*<sup>4</sup> Dans son ouvrage *Toxic Data*, David Chavalarias explique ce phénomène par la conjonction du biais de confirmation et du biais de négativité : les utilisateurs préfèrent être confrontés à des contenus qui confirment leurs opinions et sont davantage marqués par les contenus négatifs que par les contenus positifs : *“il y a donc à nouveau ici une boucle de rétroaction positive entre biais cognitifs et algorithmes de personnalisation : d'un côté, les algorithmes de personnalisation sont biaisés par une surreprésentation de vos peurs dans les contenus avec lesquels vous choisirez d'interagir ; de l'autre, ils amplifient la proportion de contenus anxigènes parmi ce à quoi ils vous exposent.”*<sup>5</sup> Les études empiriques illustrent ce constat. Des chercheurs des universités de Chicago et Columbia<sup>6</sup> ont ainsi montré que l'engagement des utilisateurs et l'exposition à la publicité est inversement proportionnel à l'intensité de la modération d'une plateforme. En masquant 6 % de contenus responsables de 73 % des contenus toxiques, l'extension de navigateur proposées à l'échantillon a également réduit de 23 % l'exposition à la publicité sur Facebook et de 9 % sur Twitter. Pour le dire autrement : **plus les contenus sont modérés, moins l'utilisateur interagit avec eux et**

<sup>4</sup> Nous traduisons.

<sup>5</sup> CHAVALARIAS, David. *Toxic data*. Flammarion, 2023.

<sup>6</sup> BEKNAZAR-YUZBASHEV, George, JIMÉNEZ DURÁN, Rafael, MCCROSKY, Jesse, et al. Toxic Content and User Engagement on Social Media: Evidence from a Field Experiment. *Available at SSRN*, 2022.

**moins il voit de publicités.** Ce fonctionnement affecte aussi la production informationnelle des médias et des partis politiques sur les réseaux sociaux. David Chavalarias rapporte ainsi les propos de la lanceuse d'alerte Frances Haugen à l'origine des Facebook Files dans une interview pour CBS News :

*“L’une des informations les plus choquantes que j’ai obtenues de Facebook et qui, selon moi, est essentielle à cette divulgation, est que les partis politiques ont été cités, dans les propres recherches de Facebook, disant : “Nous savons que vous avez changé la façon dont vous sélectionnez le contenu qui va dans le fil d’actualité. Et maintenant, si nous ne publions pas de contenu négatif, haineux, polarisant ou qui divise, c’est silence radio. Nous n’obtenons rien. Et nous pas ça. Nous savons que nos électeurs n’aiment pas ça, mais, si nous ne publions pas ce type de messages, ils ne sont pas diffusés. Avant, nous en faisons très peu, mais maintenant nous devons en faire beaucoup, parce que c’est notre travail, et si nous n’obtenons pas de trafic et d’engagement, nous perdrons nos emplois.”*”

**Les contenus contenant des fausses informations sont particulièrement présents parmi ces contenus toxiques.** Un utilisateur sur deux a déjà croisé une information fausse ou peu fiable en ligne<sup>7</sup>. 45 % des Français ont le sentiment d’être confronté plusieurs fois par semaine sur les réseaux sociaux à des informations qui déforment la réalité ou qui sont même fausses, dont 26 % tous les jours ou presque<sup>8</sup>. En 2018, une étude conduite par des chercheurs du MIT sur les contenus de Twitter entre 2006 et 2017 avait également conclu que **“les fausses informations se diffusent significativement plus loin, plus rapidement, plus profondément et plus largement que la vérité dans toutes les catégories d’informations, et les effets sont plus prononcés pour les fausses nouvelles politiques que pour les fausses nouvelles sur le terrorisme, les catastrophes naturelles, la science, les légendes urbaines ou les informations financières”**<sup>9</sup>. De même, des chercheurs de l’Université de New York et de l’Université Grenoble Alpes ont montré que, sur Facebook, entre août 2020 et janvier 2021 (période autour de l’élection présidentielle américaine), **les fausses informations ont généré six fois plus d’engagement**<sup>10</sup> **que les contenus provenant de sources considérées comme fiables** comme l’Organisation mondiale de la Santé ou le média CNN<sup>11</sup>. Cette période de l’élection américaine était également visée par les révélations de la lanceuse d’alerte Frances Haugen à l’origine des “Facebook Files”. D’après les documents révélés, durant la campagne, Facebook avait modifié ses algorithmes pour réduire la propagation de fausses informations. Cependant, d’après Frances Haugen elle, « dès que l’élection a été terminée », la plateforme les a reconfigurés comme auparavant « pour donner la priorité à la croissance plutôt qu’à la sécurité »<sup>12</sup>, allant jusqu’à faire le lien avec l’assaut du Capitole du 6 janvier 2021 qui s’est en partie organisé sur ce réseau social. En ce qui concerne la plateforme YouTube, le chercheur Jonathan Albright a mis en lumière que **les 50 vidéos complotistes les plus vues concernant les tueries de masse aux États-Unis cumulent plus de 50 millions de vue**<sup>13</sup>. Les utilisateurs sont conscients de cette propension : **70 % des Français estiment que le partage d’informations sur les réseaux sociaux participe à la diffusion de fausses informations et aux théories du complot et 66 % que cela nuit à la qualité de l’information**<sup>14</sup>. Toutefois, 50 % estiment

<sup>7</sup> VIARD-GUILLOT, Louise. “82 % des internautes protègent leurs données personnelles en ligne”. *Insee Focus* n° 27. 21 juillet 2022.

<sup>8</sup> <https://www.kantarpublic.com/fr/barometres/barometre-de-la-confiance-des-francais-dans-les-media/barometre-2023-de-la-confiance-des-francais-dans-les-media>

<sup>9</sup> VOSOUGHI, Soroush, ROY, Deb, et ARAL, Sinan. The spread of true and false news online. *science*, 2018, vol. 359, no 6380, p. 1146-1151.

<sup>10</sup> L’engagement regroupe les likes, partages, commentaires et interactions avec les contenus.

<sup>11</sup> Laura Edelson, Minh-Kha Nguyen, Ian Goldstein, Oana Goga, Damon McCoy, and Tobias Lauinger. 2021. [Understanding engagement with U.S. \(mis\)information news sources on Facebook](https://doi.org/10.1145/3487552.3487859). In *Proceedings of the 21st ACM Internet Measurement Conference (IMC '21)*. Association for Computing Machinery, New York, NY, USA, 444–463. <https://doi.org/10.1145/3487552.3487859>

<sup>12</sup> « Une lanceuse d’alerte accuse Facebook de choisir le profit plutôt que la sécurité de ses utilisateurs », *Le Monde avec AFP*, 18 octobre 2021.

<sup>13</sup> ALBRIGHT, Jonathan. Untrue-Tube: Monetizing misery and disinformation. *February*). <https://d1gi.medium.com/untrue-tube-monetizing-misery-and-disinformation-388c4786cc3d>, 2018.

<sup>14</sup> <https://www.kantarpublic.com/fr/barometres/barometre-de-la-confiance-des-francais-dans-les-media/barometre-2023-de-la-confiance-des-francais-dans-les-media>

aussi que cela permet d'intéresser à l'actualité des personnes qui ne s'informerait pas autrement et 40 % que cela apporte plus de transparence et de diversité dans l'information.

Cette abondance des fausses informations sur les grandes plateformes de médias sociaux est à relier avec la théorie de "l'enshittification" ("merdification") posée par Cory Doctorow<sup>15</sup>. Ce qu'il décrit à travers ce terme, c'est un **déclin constant de la qualité des réseaux sociaux qui est directement lié à leur modèle économique. Ce phénomène suit un cycle en 3 étapes** :

1. Les plateformes sont **bonnes pour les utilisateurs**, agréables à utiliser, leur proposent du contenu pertinent
2. Les plateformes **abusent de leurs utilisateurs** au profit de leurs clients professionnels : les feeds sont progressivement inondés de pubs, de contenus sponsorisés non choisis, de fausses informations...
3. Les plateformes **abusent de leurs clients professionnels** pour récupérer toute la valeur pour elles-mêmes, **créant une insatisfaction générale et conduisant *in fine* à leur disparition**.

Toute l'analyse de Cory Doctorow tient au fait que les réseaux sociaux ont réussi à bâtir de solides **écosystèmes fermés**, qui circonscrivent tant les utilisateurs que les professionnels.

Les données transmises par les plateformes sur ce sujet sont rares. Dans son bilan annuel 2022 des moyens et mesures mis en œuvre par les opérateurs de plateformes en ligne en 2021 pour lutter contre la manipulation de l'information, l'Arcom rapporte avoir demandé aux opérateurs de fournir des données chiffrées sur les fausses informations identifiées sur leur services. Sur les 10 opérateurs soumis à cette revue annuelle, seuls 3 ont transmis des données chiffrées<sup>16</sup>. Les données transmises sont par ailleurs souvent hétérogènes, lacunaires et le plus fréquemment au niveau mondial et non de la France. Le régulateur en tire la remarque suivante :

*"Tout d'abord, lorsque l'opérateur n'est pas en mesure de communiquer l'ensemble des données chiffrées demandées, la transparence devrait s'effectuer a minima sur des données chiffrées générales permettant une évaluation globale de la pertinence des moyens et mesures mis en œuvre. L'absence de communication d'informations telles que le nombre de fausses informations (ou notion équivalente) identifiées en France, le nombre d'émetteurs de ces contenus ou leur audience moyenne témoigne non seulement d'un manque de transparence à l'égard du public, mais soulève en outre la question de la capacité des opérateurs à identifier ces données et à évaluer leurs propres politiques en matière de lutte contre la manipulation."*<sup>17</sup>

Malgré ce constat, **il est important de rappeler que les fausses informations restent, en valeur absolue, minoritaires en ligne**<sup>18</sup>. Une étude de 2021 rappelle que les Français passent en moyenne moins de 5 minutes par jour à consulter des informations, soit 3 % de leur temps passé en ligne. Sur ce temps, les fausses informations représentent 0,16 % des contenus<sup>19</sup>. Au sein de ceux-ci, la distribution est très hétérogène : 61 % des Français interrogés ne consultent aucune source non fiable. En d'autres termes, une minorité d'utilisateurs représentent la majorité de la consommation et du partage de fausses informations en ligne<sup>20</sup>. Quelques exemples précis illustrent ce constat général. Sacha Altay, Manon Berriche et Alberto Acerbi rappellent qu'en 2016, un article de BuzzFeed avait

---

<sup>15</sup> <https://pluralistic.net/2023/01/21/potemkin-ai/#hey-guys>

<sup>16</sup> <https://www.arcom.fr/nos-ressources/etudes-et-donnees/mediatheque/lutte-contre-la-manipulation-de-linformation-sur-les-plateformes-en-ligne-bilan-2021>

<sup>17</sup> *Ibid.*

<sup>18</sup> ALTAY, Sacha, BERRICHE, Manon, et ACERBI, Alberto. [Misinformation on misinformation: Conceptual and methodological challenges](#). *Social Media+ Society*, 2023, vol. 9, no 1, p. 20563051221150412.

<sup>19</sup> CORDONIER, Laurent et BREST, Aurélien. *How do the French inform themselves on the Internet? Analysis of online information and disinformation behaviors*. 2021. Thèse de doctorat. Fondation Descartes. Cité par ALTAY, Sacha, BERRICHE, Manon, et ACERBI, Alberto, *op. cit.*, 2023.

<sup>20</sup> GRINBERG, Nir, JOSEPH, Kenneth, FRIEDLAND, Lisa, *et al.* Fake news on Twitter during the 2016 US presidential election. *Science*, 2019, vol. 363, no 6425, p. 374-378. Cité par ALTAY, Sacha, BERRICHE, Manon, et ACERBI, Alberto, *op. cit.*, 2023.

souligné que les 20 *fake news* les plus vues sur Facebook totalisaient plus de 9 millions d'interactions durant la campagne présidentielle américaine. Toutefois, ce nombre ne représente que 0,042 % des interactions totales sur la période étudiée. De même, une étude avait rapporté qu'aux États-Unis, entre 2019 et 2020, le trafic vers les sites non fiables avait augmenté de 70 % contre seulement 47 % pour les sites fiables. Toutefois, en valeur totale, cette hausse représente 30 millions de vues supplémentaires pour les sites non fiables contre 2 milliards de vues supplémentaires pour les sources fiables. Par ailleurs, **les sites des sources fiables restent les plus consultées pour s'informer<sup>21</sup> et Internet reste derrière la télévision<sup>22</sup>**. Il est néanmoins à rappeler que **les données pour clarifier cette prééminence ou non des fausses informations en ligne sont rares et que la recherche sur ce sujet mérite d'être approfondie en ce sens**.

## C. Un changement actuel de modèle économique qui ne résout pas le problème

**La période actuelle est marquée par un changement dans les modèles économiques des services en ligne.** Alors qu'ils étaient essentiellement gratuits ou se revendiquaient du moins comme tels, de plus de plus de réseaux sociaux se tournent désormais vers des modèles payants fondés sur l'abonnement ou le subventionnement. Les cas de X (ancien Twitter) ou de Facebook et Instagram sont probablement les plus emblématiques. Mais bien d'autres réseaux sociaux fonctionnent en partie sur des contributions pécuniaires des utilisateurs. Il est à noter que **ce changement de modèle répond, dans certains cas** (comme celui du groupe Meta) **à un enjeu de données personnelles**, le groupe ayant proposé cette solution pour se conformer au droit européen sur la protection des données et ne plus proposer de publicité ciblée en l'absence du consentement explicite des utilisateurs<sup>23</sup>, sans mettre en péril ses revenus.

Cependant, dans d'autres cas, comme X, la raison est autre. Après l'arrivée d'Elon Musk à la tête du réseau social, le système de certification a changé. D'une certification basée sur l'authenticité d'un compte d'intérêt public, le réseau social est passé à une certification sur la base d'un abonnement mensuel à X Premium allant de 3,6 € par mois à 19,2 € par mois en fonction des options choisies. Publiquement, le choix de ce nouveau mode de certification a beaucoup été défendu comme une façon de mieux vérifier l'authenticité des comptes et de lutter contre les faux comptes (*bots*). Pour cela, les utilisateurs qui souhaitent obtenir la certification doivent avoir un profil complet et actif, associé à un numéro de téléphone. En conséquence, ces comptes sont considérés par X comme plus fiables et sont priorisés par la plateforme par rapport aux contenus des autres comptes. **Si ce nouveau modèle d'affaires peut a priori laisser penser que les réseaux sociaux s'éloigneraient de l'économie de l'attention, il semble en réalité que cela ne résolve pas les très nombreux problèmes posés par certains réseaux sociaux, dont les enjeux de désinformation.** Au contraire, NewsGuard a rapporté que **sur les 250 affirmations fausses les plus virales sur le conflit entre Israël et le Hamas sur X, 74 % émanaient d'utilisateurs vérifiés** et abonnés à X<sup>24</sup>, et donc promus par l'algorithme de recommandation<sup>25</sup>. Une analyse de Bloomberg confirme ce constat : 83 comptes devenus "premium"

---

<sup>21</sup> FLETCHER, Richard, CORNIA, Alessio, GRAVES, Lucas, *et al.* Measuring the reach of "fake news" and online disinformation in Europe. *Australasian Policing*, 2018, vol. 10, no 2 ; GUESS, Andy, ASLETT, Kevin, TUCKER, Joshua, *et al.* Cracking open the news feed: Exploring what us Facebook users see and share with large-scale platform data. *Journal of Quantitative Description: Digital Media*, 2021, vol. 1. Cité par ALTAY, Sacha, BERRICHE, Manon, et ACERBI, Alberto, *op. cit.*, 2023.

<sup>22</sup> ALLEN, Jennifer, HOWLAND, Baird, MOBIUS, Markus, *et al.* Evaluating the fake news problem at the scale of the information ecosystem. *Science advances*, 2020, vol. 6, no 14, p. eaay3539. Cité par ALTAY, Sacha, BERRICHE, Manon, et ACERBI, Alberto, *op. cit.*, 2023.

<sup>23</sup> Cette disposition est contenue dans le règlement européen sur les marchés numériques (RMN - DMA).

<sup>24</sup> NewsGuard, [Les utilisateurs "vérifiés" sur X produisent 74% des affirmations fausses ou sans fondement les plus virales liées à la guerre Israël-Hamas sur la plateforme](#), *Misinformation Monitor*, Octobre 2023.

<sup>25</sup> <https://www.newsguardtech.com/fr/misinformation-monitor/octobre-2023/>



entre juin et novembre 2023 ont ensuite partagé de fausses informations sur le conflit entre Israël et le Hamas<sup>26</sup>. Certains d'entre eux se présentent même à tort comme journalistes.

De son côté, TikTok expérimente aussi auprès d'un petit groupe d'utilisateurs un abonnement de 4,99 \$ par mois pour un usage sans publicité. Sur ce cas, ainsi que concernant Facebook et Instagram, nous n'avons pas encore suffisamment de recul pour juger des effets de ce changement. **Il sera particulièrement intéressant d'analyser si le choix de l'abonnement payant se traduit, au-delà de la moindre captation de ses données personnelles et du changement du modèle publicitaire, par une exposition différente aux contenus toxiques.**

## II. Le modèle économique des plateformes est servi par leurs interfaces

Pour maximiser l'engagement des utilisateurs, les plateformes ont recours à des **mécanismes de designs exploitant les mécanismes cognitifs** de ces derniers pour influencer voire modifier leurs comportements sans qu'ils en aient forcément conscience. Ces stratégies sont parfois qualifiées de **technologies persuasives**, designs persuasifs ou encore "**captologie**", acronyme de *Computer as a Persuasive Technologies*. Ces stratégies ont notamment été étudiées par B. J. Fogg dans son livre *Persuasive Technology : Using Computers to Change What We Think and Do* (2002). Ces interfaces jouent en particulier sur les **biais cognitifs des individus** particulièrement bien décrits par Albert Moukheiber<sup>27</sup> et Grégoire Borst<sup>28</sup>. Dans un projet mené conjointement avec les équipes de sciences comportementales de la DITP, la DGCCRF résume très bien la prise de décision humaine selon la théorie de la rationalité limitée :

"Ce phénomène s'explique, entre autres, par le fait que nos ressources cognitives sont limitées et que notre comportement en ligne est fortement motivé par un but précis (acheter quelque chose, trouver une information, etc.). Pour économiser nos ressources, et arriver rapidement à notre but, nous utilisons alors des raccourcis mentaux. Ceux-ci représentent un gain d'efficacité essentiel pour nos cerveaux, mais nous exposent aussi à des décisions biaisées. C'est ainsi qu'on a tendance à préférer les options requérant moins de clics, à suivre ce que font les autres, à accepter les options pré-sélectionnées, à cliquer sur ce qui est le plus facilement visible, etc. - autant de biais systématiques documentés par les sciences comportementales et qui peuvent permettre de prédire le comportement des consommateurs."<sup>29</sup>

Les interfaces particulièrement trompeuses sont souvent appelées "**dark patterns**". Ce terme créé par Harry Brignull en 2010 désigne « *une interface utilisateur soigneusement conçue pour inciter les utilisateurs à faire des choses qu'ils ne feraient pas autrement* »

En 2018, Woodrow Harzog raffine ce terme en trois catégories de pratiques, comme le rappelle le LINC de la CNIL dans son cahier IP *La forme des choix* paru en 2019<sup>30</sup> :

- Le **design abusif** qui « utilise les limites et les biais cognitifs des individus pour les amener à effectuer des actions sur lesquels ils n'ont pas de contrôle » ;
- Le **design trompeur** qui « désigne des pratiques visant à représenter des éléments de telle manière qu'ils pourraient induire l'individu en erreur.[...] Il désigne également la tromperie par omission, lorsqu'une application collecte certaines données sans que l'utilisateur en ait conscience » ;
- Le **design dangereux** qui « [correspond] ici à des méthodes qui vont nous rendre vulnérables, soit directement, soit indirectement ».

<sup>26</sup> <https://www.bloomberg.com/graphics/2023-israel-hamas-war-misinformation-twitter-community-notes/>

<sup>27</sup> Albert Moukheiber, « [L'IA est-elle digne de confiance ?](#) », 22 novembre 2023.

<sup>28</sup> Grégoire Borst, « [Apprendre à penser par et contre soi-même, une compétence critique pour relever les défis du 21ème siècle](#) », 19 février 2021.

<sup>29</sup> <https://www.modernisation.gouv.fr/files/2023-06/Rapport%20de%20diagnostic%20Lutter%20contre%20les%20pratiques%20commerciales%20d%C3%A9loyales%20en%20ligne.PDF>

<sup>30</sup> LINC. [La forme des choix. Données personnelles, design et frictions désirables](#). *Cahiers IP Innovation & Prospective* no 06. CNIL.

Ce terme se retrouve aujourd'hui dans **le règlement sur les services numériques qui interdit à son article 25 les interfaces de lignes conçues, organisées ou exploitées pour “tromper ou à manipuler les destinataires de leur service ou de toute autre façon propre à altérer ou à entraver substantiellement la capacité des destinataires de leur service à prendre des décisions libres et éclairées.”**<sup>31</sup>

En terme d'espace informationnel, plusieurs *dark pattern* peuvent être identifiés :

- **La difficulté à signaler un contenu contenant une information erronée ou trompeuse.** Sur une thématique connexe, le Conseil national de la Refondation Numérique dédié à l'Apaisement de l'espace numérique et à la lutte contre les violences en ligne rappelle les résultats du Baromètre du numérique<sup>32</sup> réalisé par le Crédoc et piloté par le CGE, l'ANCT, l'Arcep et l'Arcom :

“l'édition 2022 a fait apparaître que 87 % des sondés jugeaient les interfaces de signalement des plateformes facilement accessibles et que 85 % des participants les jugeaient simples d'utilisation. Pourtant, seuls 4 répondants sur 10 affirment avoir déjà utilisé un dispositif de signalement, ce qui permet de contextualiser le sentiment d'accessibilité des dispositifs de signalement des répondants. Il apparaît également que les pratiques de signalement de comptes et de contenus inappropriés varient grandement en fonction du profil du signalant : alors que dans les populations les plus jeunes (entre 12 et 39 ans), la majorité a déjà signalé un compte ou un contenu, la proportion de signalants décroît fortement avec l'âge.”<sup>33</sup>

**Le règlement européen sur les services numériques impose désormais aux plateformes de mettre en place des dispositifs de signalement faciles d'accès et d'utilisation.** Les plateformes devront informer les utilisateurs de la décision retenue à la suite de leur signalement et les orienter vers des voies de recours internes et externes.

En complément de cette mesure, **la feuille de route du CNR remise en septembre 2023 invite à intégrer la question du signalement au Baromètre du numérique** pour “sonder les usagers sur la pertinence des dispositifs existants” et ainsi “assurer que la mise en œuvre des nouvelles obligations va dans la bonne direction”. **Il est également proposé de générer un tableau de bord des actions portées** pour “rendre compte des actions conduites par les acteurs concernés autour de l'apaisement de l'espace numérique.” Ce tableau pourrait intégrer des éléments de reddition de comptes sur le signalement opéré par les utilisateurs et leurs suites.

- **La non-transparence des contenus**, notamment :
  - **Ne pas afficher de façon claire et compréhensible qu'un contenu est sponsorisé voire est un contenu publicitaire.** En France, deux textes législatifs sont venus renforcer l'encadrement de ces contenus. Dès 2018, **la loi du 22 décembre 2018 relative à la lutte contre la manipulation de l'information** impose de faire preuve de transparence quant à l'identité de la personne physique ou la raison sociale, le siège social et l'objet social des personnes morales leur versant des rémunérations en contrepartie de la promotion de contenus d'information se rattachant à un débat d'intérêt général. Plus récemment, **la loi du 9 juin 2023 visant à encadrer l'influence commerciale et à lutter contre les dérives des influenceurs sur les réseaux sociaux** impose de communiquer clairement le caractère commercial des publications selon une taxonomie définie en fonction du lien qui unit l'influenceur à la marque.
  - **Ne pas afficher de façon claire et compréhensible qu'un compte ou un contenu est peu fiable ou erroné ou a été signalé parce qu'il contient des informations erronées ou trompeuses.** Cet étiquetage du contenu est particulièrement important premièrement en termes de transparence vis-à-vis des utilisateurs et deuxièmement parce qu'il s'agit d'un moyen intéressant pour contextualiser l'information sans pour autant la supprimer, dans une logique d'équilibre entre la lutte contre les manipulations de l'information et la préservation de la liberté d'expression. Il s'agit également d'un moyen potentiellement intéressant pour réduire la propagation d'information comme l'a

<sup>31</sup> <https://eur-lex.europa.eu/legal-content/FR/TXT/PDF/?uri=CELEX:32022R2065>

<sup>32</sup> CREDOC, [Baromètre du numérique](#), credoc.fr, 2023.

<sup>33</sup> <https://cnnumerique.fr/nos-travaux/cnr-apaisement-de-lespace-numerique-et-lutte-contre-les-violences-en-ligne>

montré Pennycook : les utilisateurs sont moins enclins à partager une information étiquetée comme fausse ou controversée<sup>34</sup>.

- Une **signalétique trompeuse**, comme par exemple sur X où la certification de comptes qui était auparavant signe de fiabilité et d'authenticité d'un compte (signification que le logo de certification revêt dans la majeure partie des réseaux sociaux par ailleurs) et qui désigne désormais l'abonnement à une formule premium.
- La **mise en avant de contenus ou commentaires contenant des fausses informations** (voir fiche sur les algorithmes de recommandation).
- **Donner l'illusion de choix aux utilisateurs, non suivi d'effet**. À ce sujet, en septembre 2021, la *Panoptykon Foundation* a publié une étude montrant que les options de paramétrage des publicités sur Facebook sont inefficaces. Malgré la possibilité offerte aux utilisateurs de désactiver certains critères pour cibler la publicité - par exemple "voir moins de publicités sur la parentalité", le choix de cette option restreint la capacité de ciblage des annonceurs mais pas celle de Facebook<sup>35</sup>.

Face à l'importance de ces interfaces, **le Parlement européen a voté le 12 décembre une résolution sur les interfaces addictives des très grandes plateformes**. Dans ce texte, l'institution insiste sur la nécessité, inscrite dans le règlement européen, d'**examiner si les interfaces des plateformes sont un facteur de propagation des risques systémiques** dont les très grandes plateformes peuvent être porteuses et le cas échéant, d'aller jusqu'à imposer des mesures correctives sur les interfaces elles-mêmes. Il encourage ainsi la Commission européenne à prendre en compte cette dimension du problème dans la mise en œuvre du règlement sur les services numériques. Dans ce texte, le Parlement invite notamment la Commission à explorer la l'opportunité d'**ouvrir l'infrastructure des réseaux dans le but d'offrir la possibilité à l'utilisateur de configurer son expérience**.

### III. La désinformation est source de gains économiques sur les plateformes

#### A. Les *super-propagateurs*

Au-delà des revenus publicitaires générés pour les plateformes par le temps passé par les utilisateurs sur leurs services et nourri - entre autres - par la propagation d'informations trompeuses ou erronées, **les manipulations de l'information sont aussi une manne financière importante pour les créateurs de contenus**.

Il faut tout d'abord rappeler que **la production de contenu en ligne est extrêmement concentrée**. Dans la très grande majorité des cas, les utilisateurs des réseaux sociaux sont des "*lurkers*" (rodeurs en français), c'est-à-dire des utilisateurs passifs qui consomment le contenu sans jamais ou très peu en produire. Parmi eux, les travaux de Pedro Ramaciotti Morales, Manon Berriche et Jean-Philippe Cointet<sup>36</sup> rappellent que **le partage de fausses informations est lui aussi marginal : sur X, les relayeurs de fausses informations représentent 0,8 % des internautes**. Au sein de ce groupe, 90 % des relayeurs ont partagé moins de 5 fausses informations et seuls 0,1 % en ont partagé plus de 50. En février 2022, *The Atlantic* publie une étude menée sur les 500 pages Facebook sur lesquelles l'engagement des utilisateurs est le plus fort aux États-Unis, durant l'été 2020<sup>37</sup>. **Ces 500 pages totalisent environ 50 % de l'engagement des utilisateurs américains sur le réseau social**. De

<sup>34</sup> PENNYCOOK, Gordon, BEAR, Adam, COLLINS, Evan T., *et al.* The implied truth effect: Attaching warnings to a subset of fake news headlines increases perceived accuracy of headlines without warnings. *Management science*, 2020, vol. 66, no 11, p. 4944-4957.

<sup>35</sup> <https://en.panoptykon.org/algorithms-of-trauma>

<sup>36</sup> MORALES, Pedro Ramaciotti, BERRICHE, Manon, et COINET, Jean-Philippe. The geometry of misinformation: embedding Twitter networks of users who spread fake news in geometrical opinion spaces. In : *Proceedings of the International AAAI Conference on Web and Social Media*. 2023. p. 730-741.

<sup>37</sup> <https://www.theatlantic.com/technology/archive/2022/02/facebook-hate-speech-misinformation-superusers/621617/>



surcroît, les chercheurs ont observé qu'il existe des **utilisateurs "élites"** qui génèrent plus de *likes*, partages, réactions, commentaires etc que 99 % du reste des utilisateurs. Et comme Facebook récompense l'engagement, c'est un cercle vertueux pour la visibilité de ces "super-utilisateurs". Or, **ces super-utilisateurs sont aussi ceux qui propagent le plus de contenus haineux et controversés**. 68 % de ces comptes répandent des fausses informations, publient du contenu raciste, sexiste, antisémite, homophobe ou des appels à la violence. En termes de profil, ces comptes sont principalement détenus par des hommes, blancs, généralement plus âgés que la moyenne des utilisateurs.

En d'autres termes, derrière l'idéal d'horizontalité et de partage de la prise de parole, **il semble que les réseaux sociaux reproduisent en leur sein le modèle de "winner takes all"** dont les marchés numériques sont l'emblème. **La production de fausses informations est concentrée entre les mains de quelques super-propagateurs.**

Les études du *Center for Countering Digital Hate* en la matière sont extrêmement éclairantes. En novembre 2021, le centre publie une première étude consacrée à l'influence climato-sceptique sur Facebook<sup>38</sup>. Il en ressort que **10 comptes sont responsables de 69 % des contenus climato-sceptiques sur le réseau social, totalisant 186 millions d'abonnés**. En mai 2022, une nouvelle étude dédiée cette fois à la désinformation sur la pandémie de Covid-19 et la vaccination sur Facebook et Twitter révèle que **65 % des fausses informations en la matière proviennent de 12 comptes, cumulant 59 millions d'abonnés**<sup>39</sup>. Ces deux études révèlent l'inefficacité des actions des plateformes à l'encontre de ces comptes. Dans le premier cas, au moment de la publication de l'étude, 99 % des posts climato-sceptiques n'avaient pas été labellisés par Facebook comme contenant des informations erronées. Dans le second cas, seuls 3 des comptes propageant de fausses informations sur le covid avaient été supprimés au moment de la publication. La première enquête montre même que la plateforme Facebook bénéficie directement de cette influence toxique puisque 8 des "*toxic ten*" ont payé pour promouvoir leurs contenus auprès des utilisateurs du réseaux social. Dans le cas du conflit entre Israël et le Hamas, Newsguard rapporte que 30 posts propagés par les 10 "pires colporteurs d'infox sur la guerre entre Israël et le Hamas" cumulent plus de 92 millions de vues<sup>40</sup>.

**Pourquoi ces comptes particulièrement influents propagent-ils de fausses informations ? L'une des clés d'explication réside dans l'intérêt économique et financier qu'ils en retirent.** Pour reprendre l'exemple des 10 comptes partageant la majorité du contenu climato-sceptique sur Facebook, ces comptes ont généré **5,3 millions de dollars de revenus publicitaires** sur Google sur les 6 derniers mois via 1,1 milliards de visites de leurs sites web sur les six mois précédant la publication de l'étude, dont 1,7 millions de dollars au profit de Google.

**La campagne de l'élection présidentielle américaine en 2016** est aussi illustrative à ce niveau : des étudiants macédoniens de la ville de Veles avaient lancé 140 sites web sur la politique américaine sur lesquels étaient publiés des contenus pro-Trump dans un but, non pas politique, mais purement lucratif selon leurs affirmations<sup>41</sup>. **Ces sites leur ont permis de gagner jusqu'à 2 500 euros par jour**. De même, lors des "convois de la liberté" durant la pandémie de Covid-19, une étude de l'université de Harvard a montré que de nombreux groupes Facebook promouvant ces convois sont contrôlés par de faux comptes alimentés dans des fermes à contenus basés au Vietnam, au Bangladesh et en Roumanie notamment, certainement dans un objectif financier puisque ces groupes redirigent fréquemment leurs utilisateurs vers des sites internet vendant des produits dérivés pro-Trump et anti-vaccination<sup>42</sup>.

Toutefois, **la création de fausses informations en ligne ne peut pas être seulement expliquée par l'intérêt économique**. Mehdi Moussaïd, chercheur en sciences cognitives à l'institut Max-Planck de Berlin, qui distingue trois profils principaux de créateurs de fausses nouvelles : les arnaqueurs dont les motivations sont purement financières comme nous venons de l'évoquer, mais également les créateurs de fausses informations qui répondent à des **motivations politiques**, afin de faire progresser un camp

<sup>38</sup> <https://counterhate.com/wp-content/uploads/2021/11/211101-Toxic-Ten-Report-FINAL-V2.5.pdf>

<sup>39</sup> <https://counterhate.com/wp-content/uploads/2022/05/210324-The-Disinformation-Dozen.pdf>

<sup>40</sup> <https://www.newsguardtech.com/fr/misinformation-monitor/novembre-2023/>

<sup>41</sup> Silverman, C. & Alexander, L. « [Comment la Macédoine s'est retrouvée au cœur de la désinformation pro-Trump](#) ». *Buzzfeed News*. 4 novembre 2016.

<sup>42</sup> <https://www.nbcnews.com/tech/internet/us-trucker-convoy-picks-momentum-foreign-meddling-adds-fray-rcna15932>

ou un candidat en particulier et, enfin, celles et ceux qui créent de fausses informations pour **faire de l'humour**, rendant la frontière entre satire et *fake news* parfois floue<sup>43</sup>.

## B. La désinformation nourrie par la publicité

**Les contenus propageant de fausses informations en ligne sont donc notamment financés directement ou indirectement par la publicité.** Directement dans le cas de contenus sponsorisés où le créateur de contenus est rémunéré pour partager une information, recommander un produit ou un service, et indirectement lorsque le contenu est apposé à côté de publicités génériques, comme par exemple en amont et durant les vidéos YouTube, à côté des posts sur Facebook ou X ou encore sur les sites internet. Par exemple, en octobre 2021, Google avait annoncé arrêter de monétiser les contenus niant le consensus scientifique sur l'existence du changement climatique et ses causes sur ses plateformes (dont YouTube). Pourtant, une étude menée par le *Climate Action Against Disinformation* en partenariat avec le *Center for Countering Digital Hate* et parue en mai 2023 montre que le climato-scepticisme continue d'être extrêmement profitable sur la plateforme<sup>44</sup>. L'étude identifie **100 vidéos contrevenant à cette politique, totalisant 18,8 millions de vues en avril 2023**. Au-delà de YouTube, l'étude révèle que **63 % des articles populaires en ligne niant le changement climatique comportent des publicités Google**.

Ces publicités programmatiques sont diffusées par le biais d'algorithmes ciblant les utilisateurs. De fait, **les marques ne choisissent en général pas l'endroit où leurs publicités sont affichées et ignorent là où elles le sont *in fine***. Pour reprendre l'exemple de l'analyse de NewsGuard sur le conflit entre Israël et le Hamas, sur les 30 posts contenant des informations fausses ou manifestement trompeuses à ce sujet, 24 étaient affichées à côté de publicités. Au total, ce sont 200 publicités de 86 grandes marques : *"De grandes sociétés, établissements éducatifs, gouvernements et ONG figureraient parmi les annonceurs, qui partageraient pour certains les revenus générés par les annonces affichées à côté de faux contenus avec leurs créateurs"*. De la même façon, l'ONG Media Matters for America a révélé en novembre 2023 que les publicités de grandes marques étaient placées à côté de contenus extrémistes, comme des contenus faisant la promotion du nazisme, de l'antisémitisme, de la LGBT-phobie ou encore du suprémacisme blanc<sup>45</sup>. À la suite de ces révélations et du soutien par Elon Musk de messages antisémites, **plusieurs grandes entreprises ont annoncé retirer leurs publicités de X**, comme Microsoft, Netflix, IBM, Disney, Paramount, Comcast, Airbnb, Uber, Coca-Cola ou encore Amazon<sup>46</sup>. **Au total, ce sont ainsi 75 millions de dollars de revenus publicitaires qui seraient menacés par ces retraits**<sup>47</sup>.

## IV. Quels leviers face aux modèles économiques favorisant les fausses informations ?

### A. Les dispositions en vigueur

Plusieurs dispositions encadrent d'ores et déjà la propagation de fausses informations en ligne. **Au niveau national, la France s'est dotée en 2018 de la loi visant à lutter contre les manipulations de l'information**. Cette loi concerne principalement la période de campagne électorale (trois mois avant un scrutin national) :

- Elle crée une **nouvelle voie de référé civil** visant à faire cesser la diffusion de fausses informations durant les trois mois précédant un scrutin national. Quand il est saisi, le juge des référés doit apprécier, sous 48 heures, si ces fausses informations sont diffusées "de manière

<sup>43</sup> Cité par Paulic, M. "Voyage au pays des images truquées. Fake news, mensonges et vidéos". *Le Un hebdo*. 2 juin 2021.

<sup>44</sup> <https://foe.org/wp-content/uploads/2023/04/YouTubes-Climate-Denial-Dollars.pdf>

<sup>45</sup> <https://www.mediamatters.org/twitter/x-placing-ads-amazon-nba-mexico-nbcuniversal-and-others-next-content-white-nationalist>

<sup>46</sup> <https://www.washingtonpost.com/technology/2023/11/17/elon-musk-x-companies-pulling-ads-anti-semitism/>

<sup>47</sup> <https://www.nytimes.com/2023/11/24/business/x-elon-musk-advertisers.html>

artificielle ou automatisée" et "massive". Dans sa décision du 20 décembre 2018, le Conseil constitutionnel a précisé que le juge ne pouvait faire cesser la diffusion d'une information que si le caractère inexact ou trompeur de l'information était manifeste et que le risque d'altération de la sincérité du scrutin était également manifeste.

- Les plates-formes numériques (Facebook, Twitter, etc.) sont soumises à des **obligations de transparence lorsqu'elles diffusent des contenus contre rémunération**. Celles qui dépassent un certain volume de connexions par jour doivent avoir un représentant légal en France et rendre publics leurs algorithmes.
- L'Arcom peut aussi **empêcher, suspendre ou interrompre la diffusion de services de télévision contrôlés par un État étranger** ou sous l'influence de cet État, et portant atteinte aux intérêts fondamentaux de la nation.

En dehors de ces périodes électorales, **la loi impose un devoir de coopération de la part des plateformes qui doivent mettre en œuvre des mesures en vue de lutter contre la diffusion de fausses informations susceptibles de troubler l'ordre public**. Elles doivent mettre en place un dispositif facilement accessible et visible permettant aux internautes de signaler de fausses informations. Les opérateurs de plateformes doivent également mettre en œuvre des mesures complémentaires pouvant notamment porter sur :

- La transparence de leurs algorithmes ;
- La promotion des contenus issus d'entreprises et d'agences de presse et de services de communication audiovisuelle ;
- La lutte contre les comptes propageant massivement de fausses informations ;
- L'information des utilisateurs sur l'identité de la personne physique ou la raison sociale, le siège social et l'objet social des personnes morales leur versant des rémunérations en contrepartie de la promotion de contenus d'information se rattachant à un débat d'intérêt général ;
- L'information des utilisateurs sur la nature, l'origine et les modalités de diffusion des contenus ;
- L'éducation aux médias et à l'information.

Ils rendent des comptes annuellement auprès de l'Arcom à cet égard qui publie un bilan chaque année.

Au niveau européen, l'encadrement de la propagation de fausses informations en ligne a d'abord pris la forme du **Code européen contre la désinformation**<sup>48</sup>. Initialement signé en 2018, le Code a été refondu en 2022 dans la perspective de l'entrée en vigueur du règlement sur les services numériques (RSN). Signé par 44 plateformes, acteurs de la publicité en ligne, associations et acteurs spécialisés, **ce texte a principalement pour vocation d'assécher le financement de la désinformation, de freiner sa diffusion, d'aider les utilisateurs à s'en prémunir et de soutenir les chercheurs et les fact-checkers**. Anticipant le RSN, ce texte est contraignant pour les très grandes plateformes totalisant plus de 45 millions d'utilisateurs dans l'Union européenne. Toutefois, Apple et Amazon ne font pas partie des signataires de ce Code et X (à l'époque encore Twitter) s'est retiré du Code au printemps 2023 à la suite du rachat par Elon Musk<sup>49</sup>.

- Les dispositions du Code en matière de **démonétisation** sont particulièrement intéressantes : *"Le code renforcé vise à faire en sorte que les fournisseurs de désinformation ne bénéficient pas des recettes publicitaires. Les signataires s'engagent à prendre des mesures plus énergiques pour éviter la mise en place de publicités à côté de la désinformation, ainsi que la diffusion de publicités contenant de la désinformation. Le code met également en place une coopération plus efficace entre les acteurs du secteur de la publicité, ce qui permet de renforcer l'action commune."* Le Code impose aussi une transparence renforcée en matière de **publicité politique** par le biais d'un étiquetage clair, affichant précisément le sponsor, les dépenses associées et la période d'affichage. Ces publicités politiques seront classées dans des bibliothèques d'annonces aisément consultables.
- En ce qui concerne les **utilisateurs**, le Code vise à mieux les protéger grâce à des outils permettant de reconnaître, comprendre et signaler la désinformation, en signalant les sources faisant autorité et en mettant en place des initiatives d'éducation aux médias et à l'information.

<sup>48</sup> <https://digital-strategy.ec.europa.eu/fr/policies/code-practice-disinformation>

<sup>49</sup> [https://www.lemonde.fr/pixels/article/2023/05/27/twitter-se-retire-du-code-europeen-contre-la-desinformation-en-ligne\\_6175078\\_4408996.html](https://www.lemonde.fr/pixels/article/2023/05/27/twitter-se-retire-du-code-europeen-contre-la-desinformation-en-ligne_6175078_4408996.html)

- Le Code prévoit que les signataires apportent un soutien accru à la **recherche et aux vérificateurs de faits** (*fact checkers*) sur la désinformation, par un accès plus large aux données notamment

Certaines des dispositions de ce Code font désormais partie du **RSN** qui complète l'arsenal réglementaire en matière de lutte contre les fausses informations en ligne.

- Les articles 34 et 35 imposent aux fournisseurs de très grandes plateformes en lignes et de très grands moteurs de recherche en ligne d'évaluer les **risques systémiques** liés au fonctionnement de leurs services ou de leurs systèmes connexes. Certains des risques d'ores et déjà identifiés dans le texte (liste non exhaustive) sont à rapprocher de la lutte contre les fausses informations en ligne, notamment **la mise en péril des droits fondamentaux, des discours civiques et des processus électoraux**. Les plateformes doivent mettre en place des **mesures d'atténuation raisonnables, proportionnées et efficaces**, adaptées aux risques systémiques spécifiques recensés.
- L'**article 40** dispose que les fournisseurs de très grandes plateformes en ligne et de très grands moteurs de recherche en ligne devront **faciliter et fournir l'accès à leurs données à des chercheurs agréés à des fins de recherche** contribuant à la détection, au recensement et à la compréhension des risques systémiques dans l'Union et d'évaluation des mesures d'atténuation des risques.

## B. Autre levier : l'ajout de fonctionnalités tierces aux réseaux sociaux

La propagation de fausses informations en ligne est un problème de modèle économique. Pour envisager des réseaux sociaux plus sains, au service des utilisateurs et non des intérêts financiers, plusieurs chercheurs et collectifs proposent d'**ouvrir ces plateformes à des fonctionnalités tierces, allant pour certains jusqu'à parler de "dégroupage des réseaux sociaux"**. Cette piste est notamment défendue par Maria-Luisa Stasi, directrice "Law & Policy des marchés numériques" chez Article 19, une ONG qui défend la liberté d'expression<sup>50</sup>. Pour elle, le dégroupage des réseaux sociaux offre l'un des moyens les plus intéressants pour concilier liberté d'expression et droit de la concurrence dans la régulation des réseaux sociaux.

Actuellement, les plateformes et les services qu'elles proposent (messaging, stockage de photos et vidéos, classement du contenu, modération...) se présentent comme un bloc monolithique: "*En proposant tous ces services ensemble, elles parviennent à la fois à se protéger de la pression concurrentielle et à priver les utilisateurs d'alternatives.*" Or, l'absence de concurrence fait écran à l'innovation : "**Sans concurrence, ces plateformes ne sont pas incitées à améliorer la qualité de leurs services** : elles ne sont pas obligées d'arrêter la surveillance des utilisateurs, la collecte excessive de données personnelles, l'utilisation de designs addictifs ou de faire taire certaines voix et d'en amplifier d'autres, sur la base de leurs seuls intérêts économiques."<sup>51</sup> Cet enfermement a aussi un impact sur les droits fondamentaux des utilisateurs, notamment la vie privée et la liberté d'expression : "*En l'absence d'alternative, les conditions d'utilisation des grandes plateformes de réseaux sociaux deviennent le standard sur lequel reposent les droits des utilisateurs. Les utilisateurs sont contraints d'accepter ces conditions, même si elles impliquent la collecte de données sans consentement, le retrait excessif de contenus légaux ou la sous-exposition à la diversité.*"

Face à cette situation, le dégroupage se présente comme une réponse à ce statu quo pour ouvrir le marché à la concurrence et redonner aux utilisateurs la maîtrise de leurs expériences en ligne :

<sup>50</sup> <https://cnnumerique.fr/paroles-de/reseaux-sociaux-explorer-lopportunite-du-degroupage-un-echange-avec-maria-luisa-stasi>

<sup>51</sup> *Ibid.*

“Pour les utilisateurs, **le dégroupage des services** signifierait que, lorsqu'ils créent ou possèdent un profil sur une grande plateforme (par exemple Instagram), **il leur serait demandé s'ils souhaitent qu'Instagram lui-même ou un autre acteur (à choisir librement) fournissent le service de recommandation de contenus**. De cette manière, **les utilisateurs pourraient choisir des services de recommandation de contenus qui répondent à leurs préoccupations ou à leurs préférences.**”

En d'autres termes, il s'agirait de **permettre la création de fonctionnalités développées par d'autres**. À la manière des plug-ins, nos outils numériques - notamment les réseaux sociaux - pourraient être enrichis par de multiples fonctionnalités développées par la communauté utilisatrices, nous permettant de personnaliser encore davantage notre expérience.

En termes réglementaires, l'**article 6(12) du règlement sur les marchés numériques** (RMN, DMA) offre une piste intéressante en ce qu'il oblige ces contrôleurs d'accès à **fournir un accès équitable, raisonnable et non discriminatoire aux acteurs économiques désireux de fournir des services concurrents ou complémentaires**. Toutefois, pour Maria Luisa Stasi, *“cette voie ne permettra pas d'atteindre le dégroupage des réseaux sociaux aussi immédiatement et directement qu'il aurait pu l'être par l'imposition d'une obligation explicite et adaptée.”* Elle appelle donc à aller au-delà de ce que les textes européens prévoient aujourd'hui et insiste sur le fait que cette solution du dégroupage a déjà été employée dans d'autres secteurs, notamment celui des télécommunications

De façon très concrète, **la Panoptikon Foundation et le collectif People vs Big Tech ont publié en novembre 2023 une note explorant concrètement le prototypage d'un réseau social favorisant le choix des utilisateurs**<sup>52</sup>. 9 fonctionnalités sont ainsi proposées, articulées autour de 3 thématiques :

- **Le feedback**
  1. Pouvoir donner un feedback instantané sur le contenu : est-il désiré ou non ?
  2. Etablir un mécanisme de plainte pour signaler ou filtrer un contenu
- **La recommandation de contenu**
  3. Pouvoir choisir un fil de contenus non-basés sur le profilage
  4. Prévoir un processus d'onboarding pour en savoir plus sur les envies de l'utilisateur
  5. Prévoir des curseurs pour choisir ce que l'on souhaite davantage voir ou moins voir
  6. Disposer d'un tableau de bord détaillant ce qui détermine de quoi est composé notre fil
  7. Disposer d'un rapport quotidien/hebdomadaire sur les évolutions de notre fil
- **Le bien-être**
  8. Disposer de fonctionnalités et de rapports sur son temps d'écran
  9. Disposer de filtres protecteurs pour éviter d'être exposé à certains contenus

Pour chacune de ces fonctionnalités, la note explore très précisément pourquoi il s'agit d'une fonctionnalité importante à instaurer, ce que cela changerait, ses limites et comment celle-ci pourraient être mises en place, par exemple :

---

<sup>52</sup>Au final, la perte pourrait être supérieure à un milliard d'euros entre 2022 et 2030. (voir tableau plus bas)



What matters more to you?

Expert content

Most popular content

More informative content

More entertaining content

Content you've liked

Learn new things

Local relevance

Global relevance

John Doe

Select signals that impact your recommended content

Your implicit feedback

Your explicit feedback ('show me more of this', 'subscriptions...')

Signals from your social network

Your profile data (age, gender...)

Location-related data

Device-related data

John Doe

Set your tolerance for borderline content

Nudity

Low High

Violence

Low High

Click-bait

Low High