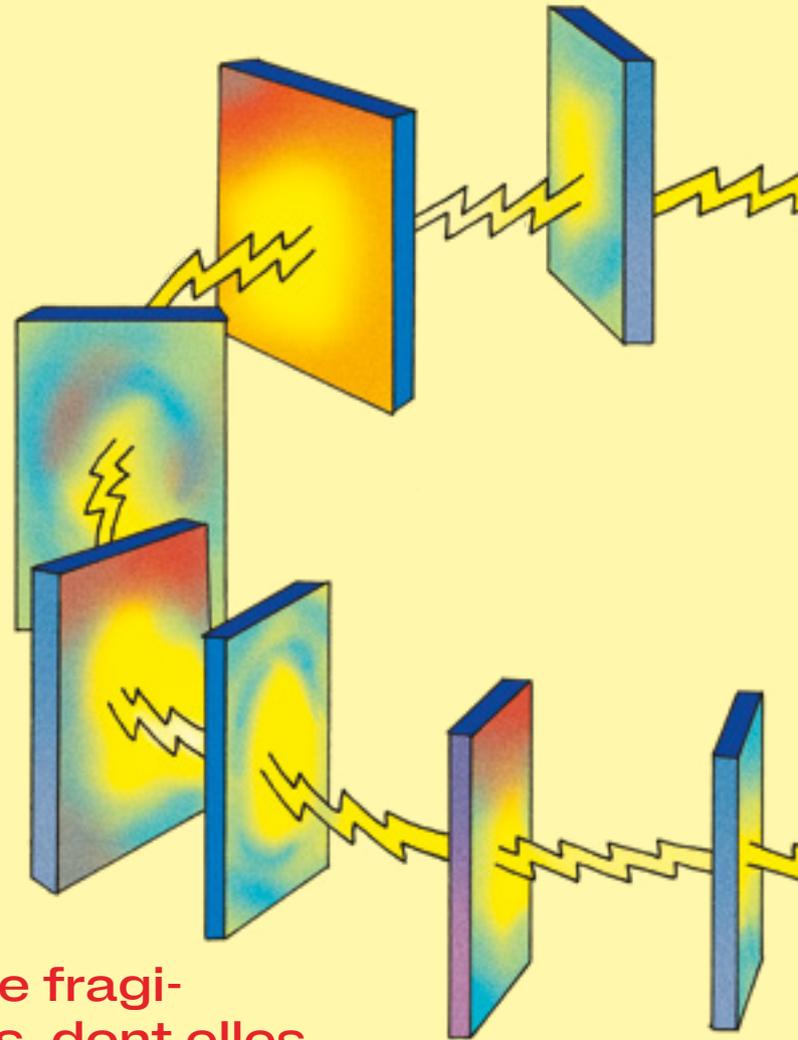
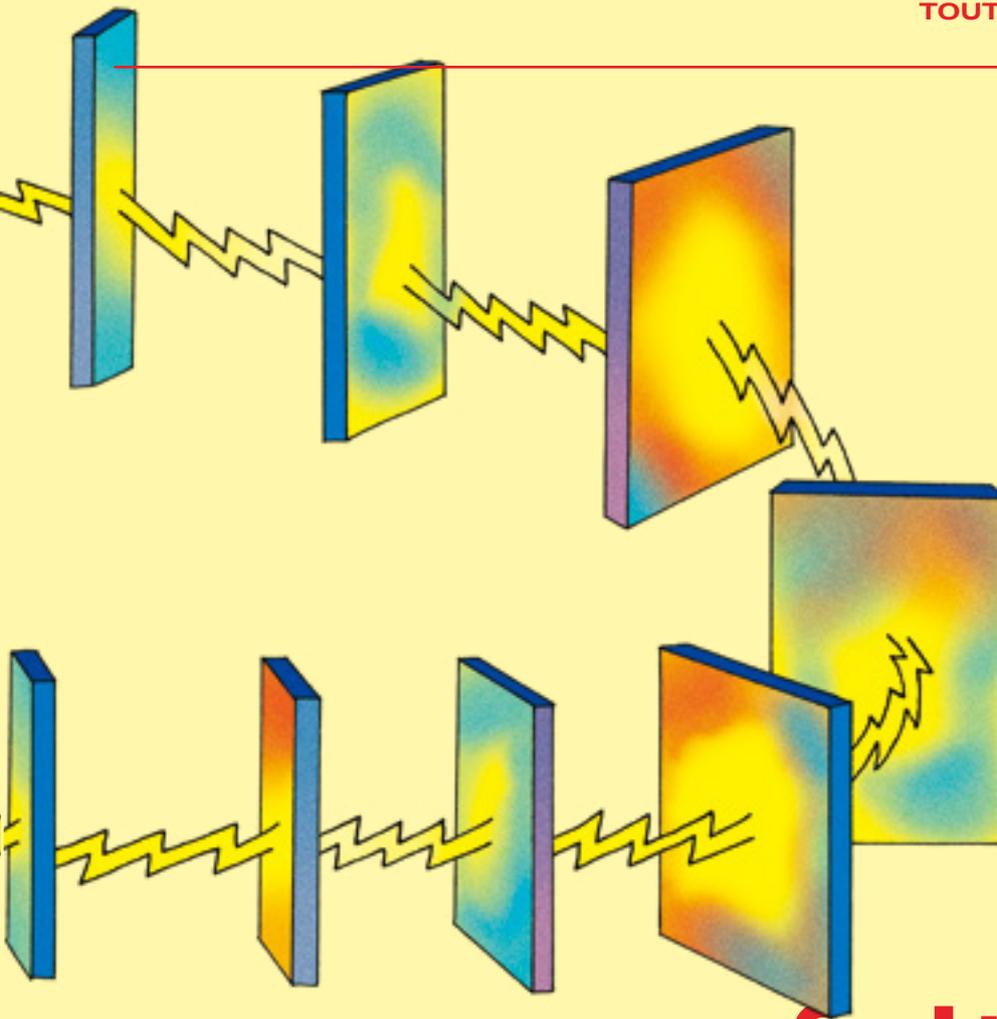


# Tout le mal que la viralité

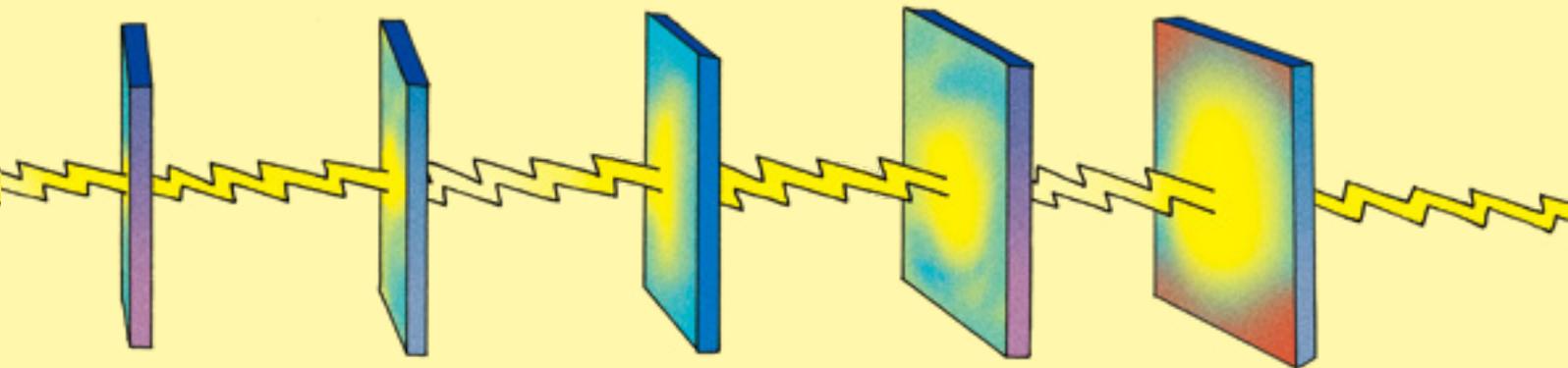
Par Dominique Boullier

Régulièrement dénoncées comme responsables de la dégradation de l'espace public, voire de la déstabilisation des régimes démocratiques, les plateformes de réseaux sociaux agissent comme des médias à part entière bien qu'elles continuent d'échapper largement aux encadrements juridiques. Elles sont aussi accusées de fragiliser les médias traditionnels, dont elles ont ponctionné les recettes publicitaires, et surtout d'encourager une viralité appauvrissante des contenus via des chaînes de propagation dont Dominique Boullier décortique ici les mécanismes.





# fait à la démocratie



**Dominique Boullier est professeur émérite des universités, chercheur au Centre d'études européennes et de politique comparée (CEE). Sociologue et linguiste, il explore les comportements induits par le numérique. Ses travaux actuels portent sur les conditions techniques et institutionnelles de survie dans les univers numériques.**



Aucun régime démocratique ne semble désormais à l'abri d'une déstabilisation. Ce constat pourrait inciter à chercher des responsables parmi des ennemis communs déjà bien identifiés, au nombre desquels les mouvances terroristes et les régimes dictatoriaux de tout type. Il est plus difficile de déceler les failles internes communes à toutes les démocraties tant ces dernières diffèrent par leurs conditions politiques et économiques. On peut néanmoins noter que, depuis le milieu des années 2010, un seul et même phénomène a affecté tous les régimes politiques : le changement profond de système médiatique. La prolifération des médias, leur dépendance accrue vis-à-vis de grands groupes financiers, leur internationalisation et la domination massive de l'image animée sur l'écrit ont modifié le statut de ces garants de l'ordre démocratique qu'étaient jusqu'alors la presse et les journalistes.

La remise en cause la plus radicale de ce statut est sans nul doute venue de la pression exercée par les réseaux sociaux (*social media*). Bien qu'ils publient des contenus, les réseaux sociaux ne subissent aucune des charges de responsabilité qui incombent aux médias classiques, quant à eux strictement encadrés. Ils doivent cet avantage aux différentes législations qui, depuis la *Communications Decency Act* (CDA) de 1996 mis en place par l'administration étasunienne Clinton-Gore à l'époque de la mutation du web public vers un web marchand, ont accordé un statut d'hébergeur à ceux que l'on appelait alors les fournisseurs d'accès. Grâce à ce statut, conforté notamment par la directive européenne de 2000 sur le commerce électronique et par la loi française de 2004 pour la confiance dans l'économie numérique (LCEN), les réseaux

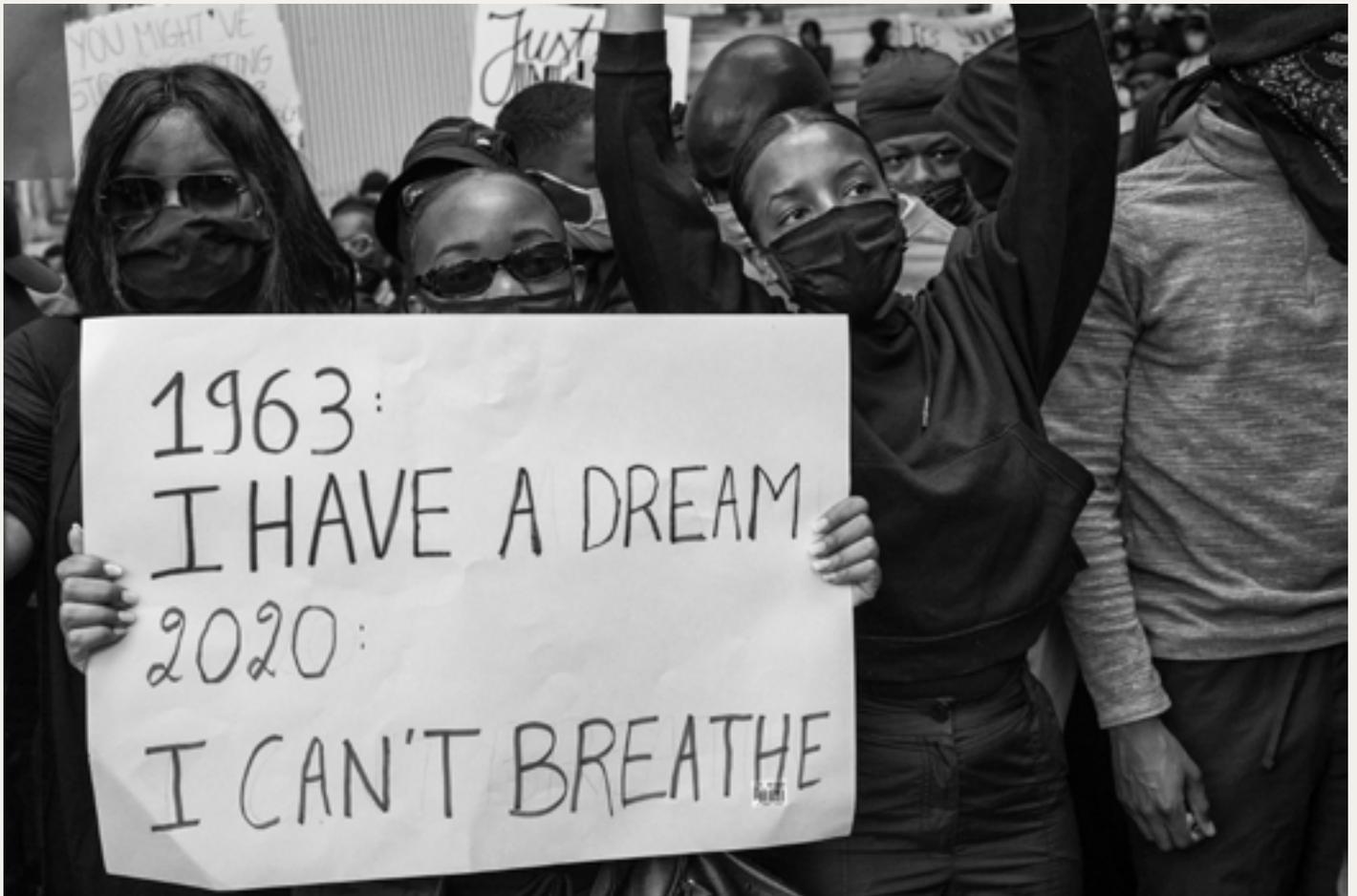
sociaux conservent une responsabilité limitée quant aux contenus que publient leurs utilisateurs. Si une telle protection pouvait se justifier durant la phase d'expansion d'internet, elle a favorisé une évolution insoutenable des réseaux sociaux, qui pratiquent désormais une politique éditoriale consistant à supprimer des contenus et des comptes. De ce vice de forme hérité du CDA découlent la plupart des problèmes contemporains ainsi que les difficultés à lutter contre un système tout-puissant, profitable et incontrôlable.

## Les effets délétères de la monétisation publicitaire

Une autre mutation a eu lieu en 2008-2009, lorsque les réseaux sociaux, alors essentiellement de convivialité, ont entrepris, au nom de la culture du web participatif, dit 2.0, de monétiser l'activité de leurs membres en vendant des placements publicitaires aux annonceurs. En 2016, la captation de la publicité par Google (essentiellement via YouTube) et Facebook a atteint 75% du marché publicitaire, mettant en péril son équilibre économique. Non soumis aux lois qui s'imposent aux autres médias, devenus des vitrines publicitaires majeures et des sources de revenus inédits, les réseaux sociaux ont acquis une position qui les rend omnipotents financièrement dans les choix d'innovation et culturellement dans la vie publique. Ce Far-West numérique ne résulte pas seulement de l'idéologie libérale du business à tout prix et des bénéfices systémiques qu'il assure aux marques et aux investisseurs, mais aussi du soutien des publics, dont ceux qui agissent comme des vecteurs de communication politique.

### L'AFFAIRE FACEBOOK-CAMBRIDGE ANALYTICA

En 2018, des lanceurs d'alerte révèlent qu'entre 2014 et 2016 les données personnelles de 87 millions d'utilisateurs de Facebook ont été exploitées sans le consentement de ces derniers par une société de marketing politique britannique, Cambridge Analytica. Agrégées à d'autres données, ces informations ont permis d'adresser, sur les réseaux sociaux, des messages ciblés selon les profils psychologiques des internautes, à partir d'un modèle élaboré à l'origine par un universitaire en exploitant des tests passés sur Facebook. En 2016, ce pillage aurait joué en faveur du vote pour le Brexit au Royaume-Uni et de l'adoubement de Donald Trump lors des primaires républicaines aux États-Unis. Le scandale a donné lieu à plusieurs enquêtes parlementaires, aux États-Unis et en Europe, à l'issue desquelles Facebook a été condamné à une amende de 5 milliards de dollars pour n'avoir pas protégé la vie privée de ses utilisateurs et Cambridge Analytica a dû cesser ses activités. L'affaire a également incité les États et l'Europe à renforcer la régulation sur la protection des données.



Une première désillusion est venue en 2018 avec l'usage détourné et abusif de données personnelles lors de l'affaire Facebook-Cambridge Analytica, qui a révélé l'absence de protection des données personnelles (voir page de gauche). Les gouvernements ont beaucoup tardé à réagir contre de tels risques. Outre le Règlement général de l'Union européenne sur la protection des données (RGPD), qui date de 2016 et a été mis en œuvre en 2018, il a fallu attendre 2022 pour que soient adoptés les *Digital Services Act* (DSA) et *Digital Markets Act* (DMA) européens. Surtout, les différents modèles de régulation persistent à ignorer la question publicitaire et ce qu'elle engendre comme dérives dans la captation de l'attention des publics. En réalité, en Europe comme aux États-Unis et dans le reste du monde, aucune régulation ne traite les phénomènes de viralité encapsulés dans le modèle économique des plateformes et les algorithmes qui les régissent.

### La propagation d'une culture de la viralité

Manifestation de soutien au mouvement Black Lives Matter, Bruxelles, 7 juin 2020.

La maladie qui ronge les démocraties est virale. La quête de visibilité est devenue un enjeu permanent de réputation, qui se manifeste à toutes les échelles. La viralité s'est notamment installée dans le monde de la finance, où elle est devenue maximale. Certains acteurs y jouent un jeu spéculatif délétère consistant à manipuler les perceptions des autres investisseurs dans le but de faire réagir le marché et d'observer les liquidités disponibles. Ils vont parfois jusqu'à l'infester de fausses promesses d'achat, qu'ils retirent avant la fin de la journée sur les marchés des transactions à haute fréquence (*high frequency trading*). Désormais, les investisseurs ne sont plus à l'affût des fondamentaux d'un titre ou d'une firme, mais plutôt des signaux propagateurs d'une croyance, d'attentes, d'anticipations sur un titre donné qui leur permettra de spéculer à la hausse ou à la baisse. Ainsi, les marques ne dépendent plus tant de leurs résultats et de leurs chiffres de vente que des réactions des investisseurs. Les statistiques de vue et de partage des

contenus (*analytics*) des plateformes alimentent la fiction de la relation de la marque à ses clients et de sa valeur potentielle (*goodwill*).

Tout cela nécessite de mobiliser les traces d'activité ou d'engagement des utilisateurs. À cette fin, les algorithmes privilégient des contenus à réactivité toujours plus élevée. Le « score de nouveauté » des contenus, indicateur mis au point par Soroush Vosoughi, Deb Roy et Sinan Aral, devient essentiel pour capter l'attention, ce qu'assurent la plupart des *fake news* en puisant au registre de l'incroyable, du clivant, de l'inédit, du drôle ou du sinistre. La réactivité engendre la viralité, stimulée par les algorithmes et les propriétés des plateformes, à l'exemple des *trending topics* que X/Twitter fournit à tous ses utilisateurs et qui agissent en tant qu'indicateurs non pas de la popularité globale d'un hashtag, mais de l'accélération de l'activité autour de lui (quitte à ce que cela retombe dans l'heure qui suit). Tous les autres médias s'alignent sur ces indicateurs de viralité pour ne pas risquer de passer à côté d'un thème qui attire de l'audience. Les journalistes subissent en cela un syndrome s'apparentant au FOMO (*fear of missing out*, peur d'avoir manqué quelque chose). Dès lors, tout thème plus complexe, moins choquant, demandant plus de temps d'explication ou susceptible d'alimenter une controverse plus documentée est relégué au bas de la liste, sa réputation et sa visibilité diminuent, cependant que s'accroît encore la viralité relative des autres thèmes ou *posts*.

Les internautes eux-mêmes, robots ou non, apprennent vite à formater leurs propres publications pour gagner en viralité. L'apprentissage collectif de ces formats de publication capteurs d'attention se fait désormais à un âge très précoce, sur des plateformes attractives pour les jeunes générations comme Snapchat et TikTok. L'ensemble des acteurs du débat public se sentent obligés de s'y adapter, pensant toucher des publics différents alors qu'ils sont en réalité condamnés à dénaturer leur propre message et à perdre leur singularité pour devenir une goutte d'eau dans les flux viraux propagés par les plateformes selon leurs propres critères. L'abrutissement général et la polarisation des contenus politiques dans le débat public sont le résultat d'un phénomène qui touche tous les médias. On pense notamment aux émissions d'*infotainment* diffusées par les mass médias, comme « Quotidien » sur TMC, « Touche pas à mon poste » sur C8 et aux

chaînes d'information répétitive, aussi appelées d'information continue, comme BFM TV, CNEWS, France info.

## Un écosystème médiatique autorépliquant

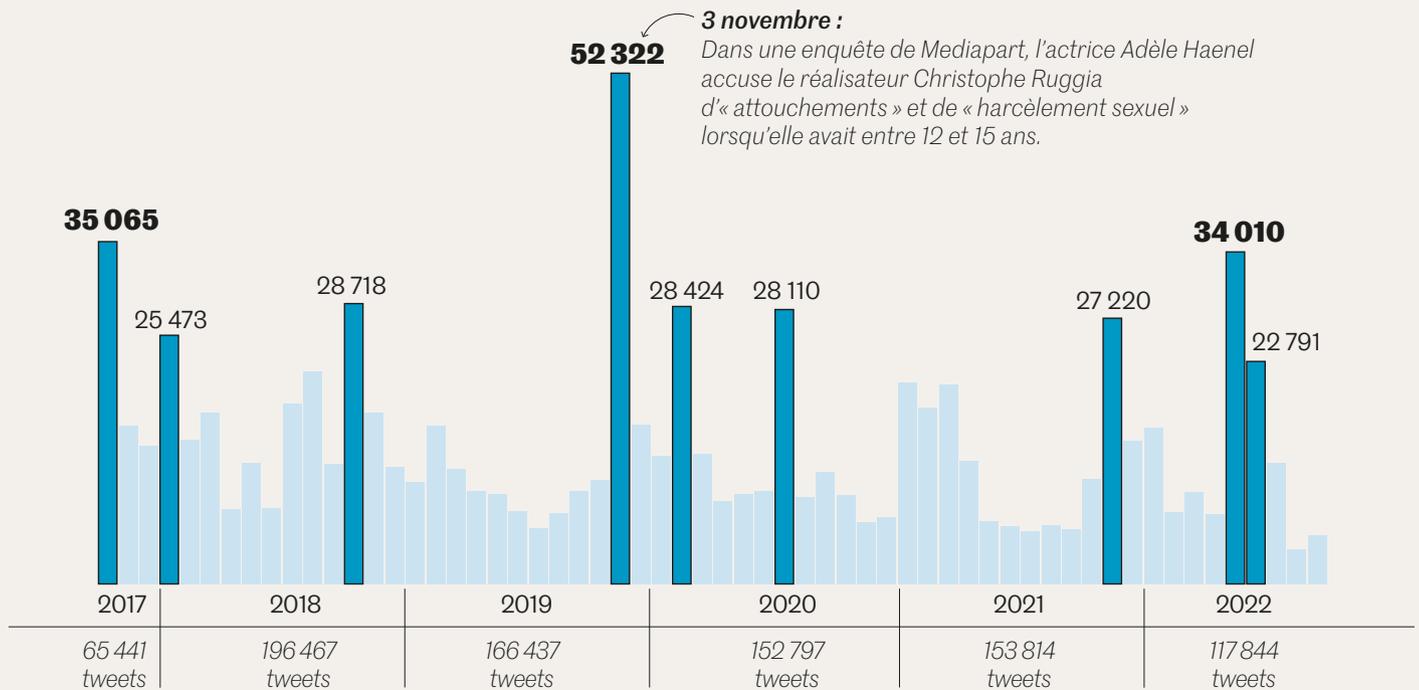
Si l'on veut bien comprendre la dégénérescence en cours de l'espace public, il est impératif de traiter l'ensemble de cette chaîne de propagation historique et fonctionnelle. Détaillons-la ici pour mettre en évidence tous les leviers qui auraient pu être actionnés, et qui peuvent l'être encore si l'on veut lutter contre cette dégradation. Chaque élément constitue une condition de possibilité (et non une cause) pour le suivant, et leur succession forme un ensemble très robuste et difficile à remettre en cause si l'on ne s'attaque qu'aux derniers maillons de la chaîne de propagation ainsi décrite :

- Plateformes, hébergeurs et non-médias —
- Publicité et monétisation des contenus eux-mêmes —
- Design des algorithmes en faveur du taux d'engagement, par captation de l'attention, pour une viralité maximale —
- Adoption du même rythme et des mêmes formats par les mass médias —
- Uniformisation des contenus conçus pour la viralité (format court, vidéo, effet de choc, catch phrases, culture du clash) —
- Captation des données personnelles et apprentissage des profils par les systèmes d'intelligence artificielle (IA) des plateformes pour mieux capter l'attention —
- Refus de modération réelle et faiblesse de la régulation structurelle — « Milieu autorépliquant » amplifié par les IA génératives —
- Triomphe du fake à tous les étages —
- Société de défiance généralisée.

Il est pourtant possible d'agir sur chacun de ces maillons, à condition de prendre en compte le milieu médiatique autorépliquant qui s'est ainsi constitué et qui a balayé tout le pouvoir de filtre des mass médias et des journalistes. Il convient en outre de ne pas oublier les producteurs de contenu à vocation propagandiste, qui sont à juste titre l'objet d'une vigilance accrue. On ne peut envisager de contrer ces derniers si l'on ne comprend pas qu'ils se sont eux-mêmes adaptés à ce milieu autorépliquant. Prenons l'exemple de l'Internet Research Agency (IRA), agence russe créée en 2013 à Saint Pétersbourg par l'oligarque Evgueni Prigojine. Plutôt que diffuser

## Le hashtag #metoo sur Twitter en France, 2017-2022

Nombre de tweets avec le hashtag #metoo



Infographie Le Monde

Source : Visibrain

des messages à visée de propagande univoques et bien ciblés, elle s'est voulue dès ses débuts comme une agence de « propagationnistes », c'est-à-dire de *geeks* capables de mobiliser des armées de trolls et de robots pour lancer simultanément sur le web, à des fins de déstabilisation des démocraties, des variants de messages très différents les uns des autres et ayant pour seul point commun de provoquer la viralité par effet de choc. L'un ou l'autre de ces variants pouvait prendre l'avantage selon des propriétés différentes, une conjoncture particulière, et le coût en restait minime puisque, par définition, ce sont les internautes eux-mêmes qui produisent la viralité par simple partage, *post*, *like* ou commentaire, sans aucun geste barrière.

Dès lors, toute stratégie pour contrer la désinformation doit être associée à une stratégie de coupure des chaînes de propagation. Les propagations intentionnelles n'ont de succès que parce que certains messages possèdent une puissance d'agir qui leur est propre – souvent issue de variations sémiotiques (voir l'exemple ci-contre) liées au hasard, comme dans toute évolution – et parce que les architectures algorithmiques des plateformes sont conçues pour favoriser cette dissémination, cette viralité. Le milieu des réseaux



### GRENOUILLE POLYSÉMIQUE

La capacité d'un signe à se propager en se transformant pour atteindre très vite de larges publics est le propre de la viralité, comme l'illustre Pepe the Frog, alias Pépé la Grenouille, héros de la bande dessinée *Boys' Club* né en 2005. Depuis 2008, ce personnage, qui incarne à l'origine un adolescent fan de jeux vidéo, a fait l'objet de détournements très variés sur internet. L'un de ses mêmes, parfois affublé d'une chevelure trumpienne, a été utilisé par l'ultra droite étatsunienne pour défendre les thèses suprémacistes, tandis qu'en Chine, une version larmoyante a servi de ralliement aux opposants au régime, avant de devenir l'un des symboles de la révolution des parapluies de 2014 puis de 2019 à Hong Kong.

sociaux, devenus marchands en accueillant la publicité et en monnayant leurs données, est économiquement et techniquement (et parfois moralement et politiquement, si l'on songe à Elon Musk) conçu pour la propagation à haute fréquence. Certes, tel n'est pas le cas de tous les dispositifs numériques contemporains, et des plateformes comme Mastodon, Twitch et surtout Wikipédia offrent des exemples d'une autre politique des rythmes, qui tend à freiner la viralité avec différentes solutions techniques, collectivement contrôlées. Dans tous les cas, les prétentions à bâtir des champions européens, comme celles qu'ont portées les fondateurs du moteur de recherche français Qwant et celles qui concernent aujourd'hui l'IA, sont condamnées à l'impuissance face aux forces de frappe économique et aux bases de clients dont disposent les plateformes étatsuniennes et chinoises. Ces ambitions sont en outre contre-productives, et l'on devrait plutôt encourager d'autres architectures techniques et économiques, plus alignées sur les valeurs européennes du droit, de la confidentialité (*privacy*), de la science, du contradictoire, de l'information vérifiée et du débat argumenté des démocraties.

## Casser les chaînes de contagion et se protéger de la viralité

En attendant que de telles initiatives, déjà en germe, se mettent en place et que des subventions publiques les soutiennent sans pour autant les contrôler, comment peut-on casser ces chaînes de propagation ? Tant que les réseaux sociaux ne seront pas considérés comme de vrais médias et qu'on ne leur interdira pas la publicité, on ne détruira pas la viralisation qui est au cœur de leur modèle. Et même si l'on doit continuer de s'accommoder de leur statut d'hébergeur et de la publicité, de nombreuses mesures de régulation restent envisageables.

Dans un *policy brief* intitulé « Puissance des plateformes numériques, territoires et souverainetés », publié en 2022 par la chaire Digital, gouvernance et souveraineté de Sciences Po, j'ai proposé vingt-neuf mesures destinées à rétablir un milieu médiatique vivable. L'une d'elles préconise que les activités en ligne de chaque internaute soient affichées en permanence sur son écran par le biais d'un tableau de bord non désactivable. Analogie à tout instrument de pilotage d'un véhicule que l'on doit pouvoir contrôler (et semblable aux panneaux de quantité de jeux vidéo), ce tableau de bord indiquerait les éléments suivants : durée de connexion, nombre de posts, nombre de likes, nombre de partages, nombre de commentaires, chacun de ces scores par durée de 24 heures. Un seuil pourrait être fixé qui déclencherait des alertes en cas d'excès de propagation. Ces signaux d'alerte, qui aideraient à passer d'une posture cognitive spontanée à une posture d'autocontrôle, de réflexivité, d'explicitation, perturberaient les comportements automatiques que les plateformes ne cessent de valoriser (achat en un clic au mépris du consentement éclairé). Le dispositif ne s'en prendrait qu'au *reach*, soit le mécanisme de viralité que chacun active par pur réflexe ; il ne concernerait pas les contenus, ne remettrait pas en cause la liberté d'expression (*free speech*). Il ne ferait que différer le partage des contenus, et rien n'empêcherait de bloquer, voire de supprimer, les comptes qui sont de très gros répliqueurs et, bien évidemment, de traquer les contenus illégaux. Il s'agirait plutôt d'enclencher un dispositif d'autocontrôle responsable chez les internautes, avant de mettre en place des régulateurs de vitesse mentale collective dès lors que certaines conditions porteuses de risques sont réunies.

*L'espace numérique relève du même type d'« action à plusieurs » que la circulation automobile, où l'on agit séparément et simultanément dans un même environnement.*

### LE DISPOSITIF DE RÉGULATION EUROPÉENNE

● **Règlement général sur la protection des données, RGPD (General Data Protection Regulation, GDPR)** Adopté en 2016,

en application depuis 2018  
Renforce la protection des données personnelles des citoyens européens et harmonise les règles relatives à la gestion de ces données par les organisations publiques et privées.

● **Règlement sur les marchés numériques (Digital Markets Act, DMA)** Adopté en 2022,

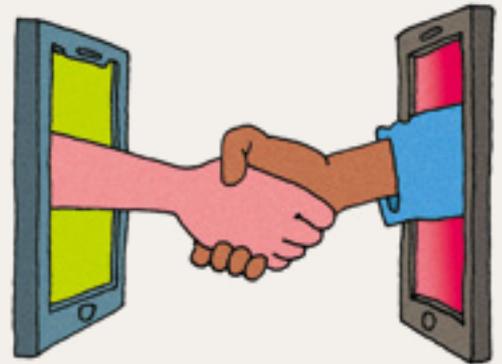
en application depuis 2023  
Impose des règles aux plateformes numériques pour lutter contre les pratiques anti-concurrentielles.

● **Règlement sur les services numériques (Digital Services Act, DSA)** Adopté en 2022,

en application depuis 2024  
Impose des règles aux plateformes numériques pour les responsabiliser, pour contrer la diffusion de contenus illicites ou préjudiciables ainsi que la vente de produits illégaux.

● **Règlement européen sur l'intelligence artificielle (Artificial Intelligence Act)** Adopté en 2024

Définit trois catégories d'applications d'IA :  
les applications à risque inacceptable (par exemple les systèmes de notation sociale gérés par le gouvernement, comme ceux utilisés en Chine) ;  
les applications à haut risque (par exemple un outil de classification des CV) ;  
les applications à risque élevé.



L'espace numérique relève du même type d'« action à plusieurs » que celle de la circulation automobile telle que l'ont décrite Pierre Livet et Laurent Thévenot. Ce n'est ni une « action collective » ni une « action ensemble », mais une situation où l'on agit séparément et simultanément dans un même environnement. De façon inhérente à ce type de situation, aucun des acteurs ne perçoit l'ampleur des interactions qui s'y déroulent, et chacun tend à penser que ses propres actions ne prêtent pas à conséquence, alors même que, dans le cas des réseaux numériques, chaque internaute contribue à activer la viralité et à mettre en alerte tout le système d'attention collective. Les espaces de circulation numérique d'informations ont beaucoup à apprendre de la régulation routière s'ils veulent redevenir vivables. Comme sur la route, il sera sans doute nécessaire, à certains moments ou pour certains comptes, de prendre des mesures contraignantes

fixant un seuil maximal de réactivité. On y gagnera un moyen d'éveiller l'attention du public à ses propres comportements et à leurs effets agrégés de même qu'un dispositif de ralentissement qui ne pourra que bénéficier à la vie publique.

### Pour une instance indépendante de collecte et de mesure des traces des plateformes

Pour engager ce type de politique, il est indispensable d'organiser l'accès aux données, lesquelles se réduisent généralement, pour chaque plateforme, à des traces de comportement. Cette obligation devrait figurer dans le cahier des charges de tous les opérateurs de plateforme sur un territoire donné. Un tiers de confiance comme il en existe pour les mass médias (par exemple, Médiamétrie en France), c'est-à-dire un organisme qui ne soit ni juge ni partie, devrait être mis en place, à charge pour lui de récupérer les traces à la source, de les analyser et de les partager avec les chercheurs (ce sont des informations d'utilité publique) et avec les parties prenantes à tous les revenus publicitaires engendrés par les plateformes, à savoir les annonceurs. Ces derniers sont actuellement facturés de façon totalement opaque (par un système d'enchères) pour des placements dont on ne mesure jamais les effets réels sur les comportements d'achat (puisque ce sont les réputations des marques qui comptent pour les investisseurs) et qui les empêchent de comprendre en détail l'engagement de leur propre communauté de consommateurs. Un tel marché de dupes, que Tim Hwang, ancien directeur du fonds philanthropique de recherche Harvard-MIT Ethics and Governance of AI Initiative, appelle aussi « la bulle de la publicité en ligne », doit être régulé. Seule une compréhension profonde des mécanismes de viralité, telle que je les ai formalisés dans un ouvrage paru en 2023, *Propagations, un nouveau paradigme pour les sciences sociales*, associée à une volonté politique forte, permettra de rétablir un espace public vivable.



Fresque murale « MeToo » sur le mur d'une école de filles, Parmer, Rajasthan, Inde, août 2024.

#### ■ RÉFÉRENCES

→ Boullier D., *Comment sortir de l'emprise des réseaux sociaux. Le réchauffement médiatique*, Paris, Le Passeur éditeur, 2020.

→ Boullier D., « Puissance des plateformes numériques, territoires et souverainetés », *Research Paper*, Sciences Po, chaire Digital, gouvernance et souveraineté, avril 2022.

→ Boullier D., *Propagations. Un nouveau paradigme pour les sciences sociales*, Paris, Armand Colin, 2023.

→ Boullier D., *Social Media Reset. « Réinventer l'infrastructure numérique des propagations pour couper les chaînes de contagion »*, Policy Brief, Sciences Po, chaire Digital, gouvernance et souveraineté, juin 2024.

→ Livet P. et Thevenot L., « Les catégories de l'action collective », dans André Orléan (dir.), *Analyse économique des conventions*, Paris, PUF, 1994.

→ Vosoughi S., Roy D. et Aral S., « The Spread of True and False News Online », *Science*, 359 (6380), 2018, p. 1146-1151.