

Les réseaux sociaux, incompatibles avec la démocratie ?

Julien Schuh

À propos de : David Chavalarias, *Toxic Data. Comment les réseaux manipulent nos opinions*, Paris, Flammarion, « Champs actuel », 2025, 320 p.

Le constat de David Chavalarias dans Toxic Data est sans appel : les réseaux sociaux comme X (Twitter) et Facebook sont dangereux pour la démocratie. L'ouvrage détaille les mécanismes derrière ces technologies de manipulation politique à grande échelle.

Issu du projet Politoscope, qui analysait les dynamiques politiques sur Twitter (<https://politoscope.org/>), le petit livre *Toxic Data* de David Chavalarias, mathématicien, directeur de recherche au CNRS et directeur de l'Institut des systèmes complexes de Paris, défend une hypothèse centrale forte : « le modèle économique actuel de la *Big Tech*, fondé sur la marchandisation de l'influence sociale, est incompatible avec la pérennité de nos démocraties » (p. 13). L'ouvrage documente avec précision la manière dont les réseaux sociaux ont servi de terrain à des campagnes de désinformation dans une véritable offensive médiatique, impliquant aussi bien des États que des groupuscules et mouvances plus ou moins institutionnalisés, qui trouvent dans ces plateformes des outils de coordination, de mise en réseau, de discussion et surtout d'amplification de leur pouvoir potentiel sur l'espace public à l'échelle internationale.

Une technologie de l'influence à grande échelle

L'analyse met en lumière deux caractéristiques essentielles de ces médias, qui doivent être pensées en détail dans leur matérialité socio-technique : d'une part, « n'importe quel individu est désormais capable d'échanger avec un nombre arbitraire et potentiellement très large de ses congénères, indépendamment des frontières géographiques, et à coût quasi nul » ; d'autre part, « l'instantanéité des messages numériques et leur potentiel infini de duplication permettent à ces mouvements d'opinion de se développer à une vitesse foudroyante, bien plus rapidement que leurs équivalents du monde réel » (p. 32). Ces spécificités permettent à des acteurs d'agir sur des espaces distants, éventuellement dans d'autres langues que la leur, et d'aligner leur stratégie sur la spécificité d'un terrain qu'ils peuvent apprendre à connaître à travers la cartographie que leur offrent ces plateformes (avec plus ou moins de succès si les codes locaux leur échappent).

L'*alt-right* américaine¹ a ainsi utilisé ces réseaux pour organiser une campagne d'influence lors de l'élection présidentielle française de 2017, à coups de mêmes visant à provoquer

¹ Abréviation d'*alternative right*, l'*alt-right* désigne une nébuleuse de mouvements d'extrême droite américains, apparue au tournant des années 2010, qui se distingue du conservatisme traditionnel par son recours aux codes de la culture internet (mèmes, forums, réseaux sociaux) et par l'affirmation décomplexée de positions nationalistes, racistes et antiféministes.

l'abstention ou à modifier les votes au dernier moment, en ciblant différemment les électorats de François Fillon (sur l'islam et l'immigration) et de Jean-Luc Mélenchon (sur l'Europe et la finance), ou en jouant sur le rejet de François Hollande pour discréditer Emmanuel Macron par association. L'amplification de ces contenus est produite par des effets propres aux plateformes, comme le *hashtag hacking* consistant à repérer un mot-clef à la mode pour faire passer des messages qui seront vus ensuite par un large public. Ces outils favorisent des captations de l'attention qui n'avaient pas d'équivalent auparavant, avec une facilité déconcertante : alors que l'achat d'un espace publicitaire lors d'un grand événement sportif comme le *Super Bowl*, dont les tarifs atteignent plusieurs millions de dollars pour trente secondes d'antenne, exclut de fait la plupart des annonceurs potentiels, ici le coût d'entrée est quasi nul. Les « mèmes » font ainsi l'objet d'une véritable industrialisation au sein de ces communautés d'extrême droite, avec partage de modèles et de conseils permettant une efficacité maximale. On pourrait parler d'une fordisation des processus de génération d'objets culturels, au sens d'une production en chaîne standardisée avec division des tâches et gabarits réutilisables. L'essor de l'IA générative ne fait qu'accentuer cette tendance. Les stratégies de manipulation sont pensées jusque dans leur temporalité, comme de véritables escroqueries au long cours : des faux comptes sur des sujets tendances sont préparés des semaines à l'avance pour gagner en légitimité algorithmique avant la diffusion massive de messages par des armées de bots.

L'homophilie de valeur

Une question essentielle de l'ouvrage porte sur les plateformes comme dispositifs d'alignement idéologique massif, interrogeant leur fonction : s'agit-il de lieux d'exposition des opinions, ou d'outils de manipulation ? « Twitter favorise-t-il la convergence idéologique, ou bien le réseau social ne fait-il que matérialiser des alignements préexistants ? » (p. 44). L'analyse se fonde sur la théorie sociologique de l'homophilie (qui se ressemble, s'assemble), conceptualisée par Paul Lazarsfeld et Robert Merton (1954) : si dans le monde physique les groupes se constituent surtout par homophilie de statut (classe sociale, profession, âge, genre), dans l'espace virtuel, c'est l'homophilie de valeur qui devient la force principale de constitution des groupes. L'auteur soutient que les mécanismes d'influence sociale se sont renforcés avec la multiplication et l'accélération des interactions en ligne, ce qui expliquerait une plus grande sensibilité aux opérations de manipulation depuis une dizaine d'années. Une idée à nuancer sans doute : l'homophilie de statut n'a pas disparu, et le numérique n'est pas la totalité de l'espace social.

L'invention du fil d'actualité par Facebook en 2006 marque un tournant : le passage d'un fil chronologique à une sélection algorithmique provoque des effets de mise en valeur massive de contenus et de concentration de l'attention. C'est ici qu'on touche à la question de l'influence comme dispositif de transfert de crédit social : les réseaux sociaux apparaissent comme de nouveaux outils de gestion de ce crédit social, avec une externalisation et une industrialisation de processus qui étaient auparavant majoritairement communicationnels (construction de la réputation par circulation de discours, etc.). Désormais, c'est au sein des plateformes de réseaux sociaux, derrière des algorithmes largement opaques, que cette valeur est générée. La pratique du « filtrage collaboratif » transforme les relations sociales en indicateurs d'homophilie, permettant ensuite de cibler des messages publicitaires ou politiques avec davantage de chances de succès.

Des bulles de filtres...

L'ouvrage décrit la manière dont ces dispositifs fabriquent des « bulles de réalité » à travers deux mécanismes : le renforcement (une plateforme propose des contenus similaires à ceux que vous avez déjà appréciés) et la « contagion algorithmique » (ces contenus sont proposés de manière ciblée à tout votre réseau), créant une « bulle de filtres » à travers un « biais de confirmation » fondé sur la comparaison sociale. S'y ajoute un biais de négativité, les informations négatives ayant plus d'impact sur l'attention, et un biais idéologique, les algorithmes de recommandation amplifiant davantage la parole des partis conservateurs (en France, l'audience des messages des Républicains sur Twitter était amplifiée de plus de 2,5 fois par rapport à une diffusion neutre en 2020) que celle des partis progressistes (p. 88).

Les réseaux sociaux trouvent ainsi leur pouvoir en grande partie en détournant des mécanismes cognitifs qui se sont développés dans des environnements non numériques. Ces mécanismes ne sont certes pas propres aux réseaux sociaux ; pour prolonger l'analyse de Chavalarias, on pourrait parler à ce sujet de phénomènes de computation sociale ou culturelle, au sens où toute société produit, agrège et transforme de l'information à travers ses interactions, à la manière d'un calcul distribué. Mais c'est l'industrialisation et l'automatisation de ces pratiques qui les rend encore plus puissantes : ces renforcements par boucles itératives s'appuient désormais sur des dispositifs extrêmement bien rodés, dotés d'une mémoire et d'une capacité d'action autonome considérables avec l'intégration d'agents d'intelligence artificielle à toutes les étapes des processus numériques, de la génération de contenus à leur diffusion ciblée. On pourrait modéliser ceci en termes d'« attracteurs », ces états vers lesquels un système dynamique tend à converger : les bulles de filtre et les communautés polarisées fonctionnent comme des bassins d'attraction, des configurations stables vers lesquelles les opinions et les groupes sont progressivement aspirés, sous l'effet de boucles de rétroaction algorithmique. Ces phénomènes émergents fabriquent des espaces culturels distincts, dont la stabilité est entretenue par les mécanismes mêmes qui les ont fait naître.

... qui rapportent !

Chavalarias rejoint les analyses de Shoshana Zuboff dans *L'âge du capitalisme de surveillance* (2020) sur la capitalisation des données personnelles, qui permet de constituer des publics cibles à partir d'attributs signalant les individus plus réceptifs à certains messages. La structuration des réseaux sociaux permet ainsi de monétiser l'influence de manière bien plus ciblée qu'auparavant, avec des expérimentations en temps réel sur une partie du public, pour vérifier l'efficacité d'une campagne avant de la lancer. Le microciblage module les messages de manière très précise et la fragmentation de la communication rend sa contestation publique encore plus difficile. Des agences de manipulation étatiques, comme l'IRA (Internet Research Agency), « ferme à trolls » russe spécialisée dans les campagnes d'influence en ligne, s'appuient ainsi sur ces outils de ciblage et exploitent le « biais de répétition » (on croit davantage une information du seul fait de la voir à plusieurs reprises). La Chine, en mettant en avant une plateforme comme TikTok, « a bien compris qu'elle pouvait reprendre à sa sauce la recette utilisée par les entreprises de la *Big Tech* pour dominer l'espace informationnel mondial : il suffit de proposer des réseaux sociaux gratuits, pratiques et funs et d'y attirer le plus grand nombre d'utilisateurs possible ».

Vers des régimes totalitaires ?

L'ouvrage montre enfin comment ces algorithmes tendent à une polarisation des débats, avec la possibilité pour un groupe de poser les thèmes sur lesquels s'affrontent ensuite les acteurs

politiques, autrement dit de focaliser l'attention sur des questions clivantes. La création de « nouvelles catégories », comme celle d'« islamo-gauchisme », permet d'ancrer les débats (Hayat 2020) ; ce concept ne prend cependant dans la sphère publique large qu'après que Gérald Darmanin le reprend en tant que ministre de l'Intérieur en octobre 2020, par un transfert de légitimité. Ce processus « amène les politiques et l'opinion publique à problématiser plus fréquemment les enjeux politiques à partir des idées de l'extrême droite ». Cette polarisation est aussi recherchée et activement produite par les plateformes elles-mêmes, dans la mesure où « les environnements très polarisés sont beaucoup plus prévisibles et donc beaucoup plus rentables que les autres, puisque les lignes de démarcation entre les populations sont très nettes ».

Au terme de son analyse, Chavalarias formule une « conjecture » inquiétante : « Les infrastructures numériques mondialisées de gestion des données sociales et de marchandisation de l'influence sociale sont des éléments de stabilité des États totalitaires, et des facteurs d'instabilité des démocraties. En l'état actuel des choses, les régimes démocratiques ne peuvent donc qu'être des régimes transitoires vers des démocraties illibérales, voire des régimes totalitaires, seuls régimes politiques stables sous l'économie numérique de l'influence sociale » (p. 172).

Le contrôle de la confiance

Toxic Data offre une synthèse documentée et argumentée des mécanismes de manipulation à l'œuvre sur les réseaux sociaux. On peut certes discuter certains points. D'abord, la question du corpus : l'analyse repose en grande partie sur Twitter/X, mais cette plateforme constitue-t-elle une représentation fiable de l'opinion ? On connaît les critiques adressées aux médias, les journalistes étant accusés justement de voir l'espace discursif social depuis la lorgnette d'un réseau où ils sont massivement présents, contrairement à une bonne partie de la population. Il faudrait peut-être réduire l'applicabilité de certaines hypothèses à cette plateforme en particulier, d'autant que l'ouvrage n'examine pas les interactions avec d'autres médias. La question de la répliquabilité de l'expérience se pose également aujourd'hui, Twitter, transformé en X, ayant fermé les accès gratuits à ses données.

Un autre manque, lié à l'actualité récente, concerne l'instrumentalisation de ces plateformes par leurs propriétaires eux-mêmes. Le rachat de Twitter par Elon Musk et sa transformation en corpus d'entraînement pour modèles d'IA, mais surtout en terrain d'expérimentation pour un projet global de manipulation de l'opinion par *chatbot* interposé, illustre une évolution majeure. Grok, l'IA générative déployée sur X, devient un outil de médiation par rapport au passé et à l'actualité, que l'on peut modifier pour donner des réponses alignées sur un axe idéologique précis, comme l'ont montré les *fake news* propagées par Musk sur des « génocides de blancs » en Afrique du Sud ou des contenus négationnistes. On voit que c'est moins désormais l'influence sociale qui est vendue par ces plateformes que le contrôle des discours dans l'espace public.

Les institutions qui étaient les cadres principaux de construction des vérités (l'école, l'État, les médias) se voient aujourd'hui concurrencées par des outils d'IA générative facilement exploitables, dont le ton assertif et l'apparente omniscience induisent une confiance souvent injustifiée. Comme le note l'auteur, les réseaux court-circuitent les processus traditionnels de construction de la confiance, qu'ils soient cognitifs (vérification des sources, recoupement des informations, évaluation de la crédibilité d'un locuteur) ou politiques (délibération collective, validation par des institutions légitimes, confrontation des points de vue dans l'espace public). Il s'agit donc de penser ces technologies dans une histoire plus large de la production du crédit réputationnel dans l'économie discursive publique. Ceci relève aussi d'un problème

d'économie de l'attention : ces plateformes, dont TikTok est peut-être aujourd'hui le meilleur exemple, deviennent des outils à la fois d'apprentissage sur les goûts des publics et de captation et de transformation de leur attention, par le biais d'algorithmes maximisant la rétention et d'outils de génération de contenus. Ceci pose de manière essentielle la question de la production de la confiance dans nos écosystèmes médiatiques numériques et dans nos démocraties représentatives.

Bibliographie

Hayat, S. 2020. « L'islamo-gauchisme : comment (ne) naît (pas) une idéologie », *Le Nouvel Obs*, 27 octobre 2020, URL : <https://www.nouvelobs.com/idees/20201027.OBS35262/l-islamo-gauchisme-comment-ne-naît-pas-une-ideologie.html>.

Lazarsfeld, P. et Merton R. K. 1954. « Friendship as Social Process. A Substantive and Methodological Analysis », in M. Berger, T. Abel et C. H. Page (dir.), *Freedom and Control in Modern Society*, New York : Van Nostrand, p. 18-66.

Zuboff, S. 2020. *L'âge du capitalisme de surveillance : le combat pour un avenir humain face aux nouvelles frontières du pouvoir*, traduit par B. Formentelli et A.-S. Homassel, Paris : Éditions Zulma.

Julien Schuh est professeur de littérature française à l'Université Paris Nanterre. Il copilote le Consortium-HN pictoria (<https://pictoria.hypotheses.org/>) consacré à l'utilisation de l'IA dans la culture visuelle. Dernière publication : *Passés contestés et présents numériques* (Le Bord de l'eau, 2025), avec Daniel Foliard.

Pour citer cet article :

Julien Schuh, « Les réseaux sociaux, incompatibles avec la démocratie ? », *Métropolitiques*, 1^{er} juin 2026.

URL : <https://metropolitiques.eu/Les-reseaux-sociaux-incompatibles-avec-la-democratie.html>.

DOI : <https://doi.org/10.56698/metropolitiques.2300>.