

La démocratie sous l'œil des algorithmes

Description

Interview de [David Chavalarias](#) – Propos recueillis par [Françoise Laugé](#)

Mathématicien, directeur de recherche CNRS à l'Institut des systèmes complexes de Paris-Œle-de-France, vous avez créé le Politoscope, un « macroscopie politique » dites-vous, pour analyser les flux d'informations sur Twitter ([voir La rem n°64, p.67](#)). Pourquoi avoir choisi ce réseau social ?

Twitter est un réseau social mondialisé dont les utilisateurs peuvent publier des messages courts et publics (tout internaute peut les lire). Chaque utilisateur peut s'abonner aux productions des autres et être ainsi informé instantanément de ce qu'ils produisent. Vers 2009, Twitter a ajouté la possibilité de retweeter un message d'un simple clic, c'est-à-dire de republier à l'identique auprès de ses propres abonnés le message d'un autre utilisateur, ce qui a introduit une très forte viralité dans la circulation d'informations. Certains messages atteignent ainsi plusieurs millions de retweets et ont un retentissement planétaire.

Dans les années 2010, Twitter a été pris en compte des journalistes et est devenu un des principaux canaux pour la communication politique en période électorale. La très grande majorité des partis politiques utilisent pour relayer leurs messages ; certains ont des équipes dédiées à ce réseau et il est fréquent que des messages Twitter soient cités par d'autres médias tels que la radio ou la télévision. En intégrant ce réseau, le pari était que l'on pourrait saisir la plupart des actualités politiques, les grandes lignes de la communication des partis et les interactions entre militants avec une granularité temporelle inédite.

Twitter avait aussi un autre atout de taille : c'est un des rares réseaux sociaux offrant la possibilité de collecter en masse les messages *via* ce qu'on appelle des API (interfaces de programmation d'applications), contrairement aux autres réseaux sociaux comme Facebook, dont la plupart des contenus sont privés. De ce fait, un grand nombre de travaux académiques dans le domaine des sciences sociales computationnelles ont porté sur Twitter.

Grâce à l'analyse des réseaux complexes, le macroscopie dessine en temps réel une cartographie des changements sur Twitter. Quels événements ou mouvements d'opinion les plus frappants avez-vous ainsi identifiés ?

Nous avons plusieurs terrains d'observation : la politique française, le changement climatique, les

pandémies, etc. En commençant à travailler sur Twitter en 2015, je ne savais pas dans quelle mesure les données de ce réseau permettraient de décrire les structures sociales à une échelle nationale. J'ai eu la bonne surprise de découvrir que, pour autant que Twitter soit central à une activité sociale donnée (par exemple, la politique), non seulement ces données permettent de cartographier précisément différentes structures de militantisme numérique, mais l'évolution de ces structures est également très riche de phénomènes se produisant hors ligne, qui nécessitent généralement de grandes enquêtes sociologiques pour être quantifiés. Par exemple, nous avons observé une très forte recrudescence du «climato-dérialisme» au cours de l'été 2022, alors que des enquêtes sociologiques réalisées depuis ont confirmé que cette attitude a gagné 8 % en France entre 2021 et 2022 ([voir La rem n°64, p.67](#)).

CES DONNÉES PERMETTENT DE CARTOGRAPHIER PRÉCISÉMENT DIFFÉRENTES STRUCTURES DE MILITANTISME NUMÉRIQUE

Un autre phénomène marquant est la rapidité avec laquelle peuvent se reconfigurer les réseaux de militantisme suite à des événements majeurs tels qu'un vote, la découverte d'un vaccin, ou même un débat télévisé. C'est politique, en six ans, nous avons vu moins de deux grandes organisations du paysage dans le militantisme : l'émergence de la communauté anti-«dictature sanitaire», suite à la découverte du vaccin contre le Covid-19, et la cristallisation d'une très importante communauté «Reconquête !» autour de Nicolas Zemmour, suite à son débat télévisé avec Jean-Luc Mélenchon en septembre 2021. Nous avons vraiment vu de quelle manière il a pu fédérer autour de lui une partie des militants du Rassemblement national et des Républicains, entamant très fortement en quelques jours le capital social numérique de chacun de ces deux partis.

UN AUTRE PHÉNOMÈNE MARQUANT EST LA RAPIDITÉ AVEC LAQUELLE PEUVENT SE RECONFIGURER LES RÉSEAUX DE MILITANTISME

Vous constatez une dislocation du tissu social sous l'influence des réseaux sociaux. En quels termes ?

Via nos macroscopes, nous analysons les prises de parole et le militantisme politique sur Twitter (16 millions de comptes uniques observés), ce qui n'est pas nécessairement représentatif de la société française. Néanmoins, sur cet espace, nous observons une forte fragmentation du paysage de l'opinion en communautés d'utilisateurs qui entretiennent des rapports de plus en plus conflictuels. Sous l'effet des partis situés aux extrémités du spectre politique, il y a également

une multiplication de mises en scène de débats très fortement bipolarisés (tels que le grand remplacement, l'islamisme-gauchisme, le port du voile), qui pourraient favoriser l'émergence de deux camps irréconciliables au sein de la société française.

NOUS OBSERVONS UNE FORTE FRAGMENTATION DU PAYSAGE DE L'OPINION EN COMMUNAUTES D'UTILISATEURS QUI ENTRETIENNENT DES RAPPORTS DE PLUS EN PLUS CONFLICTUELS

C'est ce qui est arrivé aux États-Unis, où une étude de 2021 indique que « 52 % des lecteurs de Donald Trump, ainsi que 41 % des lecteurs de Joe Biden, sont plutôt d'accord ou tout à fait d'accord pour dire qu'il est temps de couper le pays en deux ». Quand un pays arrive à ce stade de dislocation, c'est extrêmement grave et cela semble être, peu ou prou, le chemin que nous prenons.

Enfin, il faut noter que certains pays comme la Russie se sont faits maîtres dans l'art d'affaiblir leurs adversaires en utilisant les réseaux sociaux pour semer du dissensus à grande échelle. Cette stratégie a été documentée, par exemple, dans le cas des États-Unis lors de l'élection présidentielle de 2016 et tout porte à croire qu'elle est également d'application en Europe.

Dans l'avant-propos de votre ouvrage *Toxic Data. Comment les réseaux manipulent nos opinions* (Flammarion, 2022), vous présumez que « le modèle économique actuel des Big Tech, fondé sur la marchandisation de l'influence sociale, est incompatible avec la pérennité de nos démocraties ». Le Politoscope suffit-il à démontrer ce capitalisme d'influence ?

Je définis le capitalisme d'influence comme une forme de capitalisme qui tire profit de sa capacité à organiser la manière dont nous nous influençons les uns les autres et dont nous nous constituons en groupes sociaux. C'est d'abord le thèse de Dupuy-von Foerster, un résultat théorique confirmé empiriquement, qui permet de poser cette notion. Celui-ci démontre le lien entre prédictibilité des comportements individuels et niveau d'information sociale au sein d'un réseau : plus les actions des membres d'un réseau social sont corrélées entre elles (par exemple parce qu'elles s'appuient sur la même information sociale), plus les comportements collectifs seront prévisibles à partir de l'observation d'une petite partie des comportements.

LE THÈME DE DUPUY-VON FOERSTER DÉMONTRÉ LE LIEN ENTRE PRÉDICTIBILITÉ DES COMPORTEMENTS INDIVIDUELS ET NIVEAU

 D'INFORMATION SOCIALE AU SEIN D'UN RÉSEAU

Les Big Tech qui exploitent les très grands réseaux sociaux numériques ont accès à l'observation de tous les comportements sur leur plateforme et ont par ailleurs pour principal modèle d'affaires la publicité ciblée, dont l'efficacité dépend de la capacité à prévoir le comportement des catégories d'acteurs. Ces Big Tech sont donc amenées à agir sur les réseaux sociaux de manière à les modeler pour les rendre compatibles avec ces pratiques de publicité ciblée. Ceci, d'après le thèse de von Foerster, est d'autant plus facile que le réseau sera saturé d'informations sociales (recommandations, trends, affichages du nombre de likes, etc.). Ce qui est effectivement le cas.

L'analyse des interactions sur Twitter apporte des illustrations de ce phénomène, de même que l'analyse de son algorithme de recommandation, récemment rendu public. On y découvre que Twitter a établi un peu plus de 130 000 catégories d'utilisateurs, au sein desquelles ses algorithmes tendent à homogénéiser l'information et donc les perceptions du monde. Dans *Toxic Data*, je détaille également d'autres expériences réalisées sur d'autres réseaux comme Facebook qui vont dans le même sens.

Quelles seraient les conditions nécessaires afin que convergent démocratie et mondes numériques ? Comment l'intelligence artificielle pourrait-elle servir le débat démocratique ?

Nous n'avons pas encore pris pleinement la mesure de l'importance des réseaux sociaux et de leurs algorithmes de recommandation dans la formation de l'opinion publique. Or, tels qu'ils sont mis en œuvre actuellement, ils sont soumis à des exigences de rentabilité en dehors de tout cadre éthique, dont les effets collatéraux menacent la stabilité des démocraties. Ils participent à la fragmentation du tissu social et ils les rendent vulnérables aux ingérences informationnelles étrangères. Mais ceci n'est pas une fatalité.

TWITTER A ÉTABLI UN PEU PLUS DE 130 000 CATEGORIES
 D'UTILISATEURS, AU SEIN DESQUELLES SES ALGORITHMES TENDENT À
 HOMOGENÉISER L'INFORMATION

Tout comme nous avons des routes publiques ou des places publiques, nous pourrions avoir des réseaux sociaux numériques publics affranchis de telles contraintes économiques ou du moins des réseaux privés sérieusement régulés de manière que leur impact sur la circulation de l'information et sur la formation de l'opinion publique ne porte pas atteinte au bien-être individuel et collectif. Outre le fait que cela rendrait les démocraties moins vulnérables, ce qui n'a pas de prix, ce serait

probablement avantageux, d'un point de vue économique, car le coût induit par des réseaux sociaux dysfonctionnels est très important. On peut penser par exemple à l'effet de désorganisation sociale et sanitaire due à la désinformation autour du Covid-19 pendant la pandémie.

Enfin, un des grands problèmes du déploiement des réseaux sociaux actuels est qu'ils sont systémiques : le même algorithme de recommandation est déployé à l'échelle de dizaines de millions d'individus (2,8 milliards pour Facebook). Le moindre dysfonctionnement engendre donc un risque systémique. Ce problème a cependant des solutions qui reposent notamment sur la déconcentration. Nous allons aujourd'hui vers une nouvelle génération de réseaux sociaux qui seront décentralisés et où chacun pourra choisir son algorithme de recommandation. Des technologies de réseaux sociaux s'appuyant sur un protocole commun permettant à un utilisateur de migrer de l'un à l'autre sans perdre son capital social sont aussi une voie à explorer.

NOUS POURRIONS AVOIR DES RÉSEAUX SOCIAUX NUMÉRIQUES PUBLICS AFFRANCHIS DE TELLES CONTRAINTES ÉCONOMIQUES

Quant à l'IA (intelligence artificielle), il y a plusieurs manières dont elle pourrait servir le débat démocratique, notamment en aidant à l'exploration et à la synthèse des débats sur un sujet donné. Mais elle peut aussi lui nuire fortement, en augmentant le degré d'intermédiation technologique entre les utilisateurs ou en donnant à des acteurs malveillants la capacité d'influencer l'opinion par la création d'une multitude de faux comptes plus vrais que nature. Elle est d'ailleurs déjà utilisée à cet effet, ce qui inquiète jusqu'aux concepteurs même de ces AI. En témoigne l'appel récent du PDG d'OpenAI, Sam Altman, à réguler l'IA alors qu'il se dit « nerveux » sur les perspectives de manipulations électorales¹.

Un Centre européen pour la transparence algorithmique (ECAT) a été inauguré, le 8 avril 2023, à Sville. Cette nouvelle organisation a pour mission de contrôler les bonnes pratiques mises en œuvre par les Big Tech dans le respect du Digital Services Act. L'ECAT parviendra-t-il à ouvrir la boîte noire des réseaux sociaux ?

Si elle est dotée d'un fort pouvoir judiciaire, oui. Sinon, elle dépendra du bon vouloir des Gafam pour accéder aux données et pour l'implémentation des recommandations.

Le robot conversationnel ChatGPT a envahi l'espace public. Que peut-on raisonnablement attendre de ce nouvel outil de recherche en ligne ?

Il y a tout d'abord des promesses de gain de productivité dans un grand nombre de domaines et d'accélération de la création de la connaissance. De ce point de vue, les perspectives sont effectivement immenses. Mais il y a aussi un réel risque de saturation de l'espace public par ces IA

conversationnelles et une opacification des relations interindividuelles. Tant que leur utilisation respectera un certain cadre éthique, les bénéfices supplanteront probablement les risques. Mais si ces IA tombent entre de mauvaises mains, ce qui n'a pas de raison de ne pas se produire, ce peut être catastrophique. Rappelons qu'Elon Musk est en train de créer une IA qu'il appellera TruthGPT au motif que ChatGPT serait trop contraint par le politiquement correct de ses créateurs et donc qu'il ne dirait pas la vérité (ChatGPT refuse, par exemple, de tenir des propos racistes). Mais l'appellation « Truth » (vérité) est trompeuse, car une IA conversationnelle ne fait que recombinaison de manières très sophistiquées ce qui était contenu dans son corpus d'entraînement. Si le corpus d'entraînement contient des propos faux ou racistes, ceux-ci n'en deviendront pas pour autant plus vrais ou plus éthiques en passant dans une moulinette comme ChatGPT. Une IA conversationnelle peut autant manipuler les opinions qu'elle donne le sentiment d'être intelligente et « objective » ; or, ne pas pouvoir éliminer les biais contenus dans son corpus d'apprentissage est un vrai problème, ce qui est quasiment toujours le cas pour les IA déployées par des entreprises privées.

Si Elon Musk met sa menace d'exécution de faire payer l'accessibilité aux données de Twitter, y compris aux chercheurs, quelle alternative avez-vous pour poursuivre vos travaux ?

Qu'il mette ou pas sa menace d'exécution, nous commençons à nous tourner vers les nouvelles générations de réseaux sociaux qui apporteront de nouveaux terrains de recherche. Ces réseaux ne sont pas encore déployés à grande échelle, mais nous sommes confiants sur le fait qu'ils le seront prochainement.

LE COÛT INDUIT PAR DES RÉSEAUX SOCIAUX DYSFUNCTIONNELS EST TRÈS IMPORTANT

Twitter, de son côté, est déjà en train de perdre de son intérêt pour observer le social car il perd en diversité au profit de populations aux idées extrêmes ou farfelues. Nous avons ainsi pu observer sur les échanges autour du réchauffement climatique que, depuis l'arrivée de Musk, l'équilibre entre les pro-climat et les climato-dénialistes était passé en quelques jours de 70/30 à 50/50, ce qui n'est plus du tout représentatif de la population. Dans le même temps, plusieurs études, dont la nôtre², ont montré un regain de conflictualité sur Twitter.

On sait, par ailleurs, qu'Elon Musk souhaite utiliser ces réseaux pour mener une « guerre culturelle » et favoriser ses idées, que l'on peut classer comme extrême droite avec un antisémitisme latent et un penchant pour le complotisme.

NE PAS POUVOIR DÉTERMINER LES BIAIS CONTENUS DANS SON CORPUS
D'APPRENTISSAGE EST UN VRAI PROBLÈME, CE QUI EST QUASIMENT
TOUJOURS LE CAS POUR LES IA DÉPLOYÉES PAR DES ENTREPRISES
PRIVÉES

Il a, par exemple, essayé d'interférer dans la vie politique américaine avec une mise en scène complotiste visant l'administration américaine et les Démocrates qu'il a appelé «*Twitter Files*», et dont les révélations ne comportaient en réalité rien de très compromettant. Il a également accédé à la demande du président turc Erdogan de fermer plusieurs comptes d'opposants politiques en pleine élection présidentielle. Au nom de la «*liberté d'expression*», il a établi plusieurs comptes bannis de Twitter pour propos nazi ou pratiques de harcèlement tandis que, dans le même temps, il bannissait des comptes parce qu'ils s'attaquaient à lui ou des comptes de journalistes qui avaient eu l'audace d'enquêter sur lui. On a donc ici la parfaite illustration du danger que constituent les réseaux sociaux centralisés. Après avoir accumulé des données et capital social, ils peuvent changer de main du jour au lendemain et être mis au service d'idéologies extrêmes, qui profiteront de l'aura de tous les influenceurs.

TWITTER EST DÉJÀ EN TRAIN DE PERDRE DE SON INTÉRÊT POUR
OBSERVER LE SOCIAL CAR IL PERD EN DIVERSITÉ

Ceux-ci n'auront sans doute pas le courage de quitter le réseau au risque de perdre leurs audiences, qui seront alors détournées pour offrir à ces idéologies une exposition sans précédent. C'est comme cela que commence le conditionnement des masses.

Sources :

1. Anthony Cuthbertson, «*ChatGPT creator Sam Altman «*nervous*» about AI election manipulation*», independent.co.uk, 17 May 2023.
2. Paul Bouchaud (ISC-PIF ; CAMS), David Chavalarias (ISC-PIF ; CAMS), Maziyar Panahi (ISC-PIF), «*Le système de recommandation de Twitter est-il biaisé ? Un audit*», hal.science/hal-04036232, 2023.

Categorie

1. Articles & chroniques

date cr  e

4 juillet 2023

Auteur

davidchavalarias