

# Au cœur de la guerre des données de Facebook

Le 14 juillet 2021 par Kevin Roose <https://www.nytimes.com/2021/07/14/technology/facebook-data.html>



**Les cadres de Facebook se sont affrontés en interne au sujet de la transparence. Crédit... Josh Edelson/Agence France-Presse – Getty Images**

**Les dirigeants du réseau social se sont déchirés au sujet de CrowdTangle, un outil de données appartenant à Facebook qui a révélé les niveaux élevés de fréquentation des sources médiatiques de droite.**

Un jour d'avril, les personnes derrière CrowdTangle, un outil d'analyse de données appartenant à Facebook, ont appris que la transparence avait des limites. Brandon Silverman, co-fondateur et directeur général de CrowdTangle, a réuni des dizaines d'employés lors d'une visio-conférence pour leur annoncer qu'ils allaient être dissous. CrowdTangle, qui fonctionnait de manière quasi-indépendante au sein de Facebook depuis son rachat en 2016, allait être placé sous la coupe de la division intégrité du réseau social, autrement dit, le service chargé de débarrasser la plate-forme de toute désinformation et discours haineux. Certains employés de CrowdTangle ont été réaffectés vers d'autres secteurs, et M.Silverman ne sera plus en charge au quotidien de l'équipe.

L'annonce, qui a plongé les employés de CrowdTangle dans un mutisme ébahi, est la conclusion d'une bataille d'un an entre les dirigeants de Facebook concernant la transparence des données et la quantité de données que le réseau social doit révéler quant à son fonctionnement interne.

D'un côté, des cadres, dont M.Silverman et Brian Boland, vice-président de Facebook chargé de la stratégie des partenariats, ont fait valoir que Facebook devrait partager publiquement autant d'informations que possible sur ce qui se passe sur sa plate-forme – que ce soit bon, mauvais ou ignoble.

De l'autre côté, des cadres, dont le directeur du marketing et le vice-président chef de l'analytique, Alex Schultz, craignaient que Facebook n'en dévoile déjà trop.

Ils ont fait valoir que les journalistes et les chercheurs utilisaient CrowdTangle, une sorte de moteur de recherche turbo qui permet aux utilisateurs d'analyser les tendances sur Facebook et de mesurer les performances des publications, pour déterrer des informations qu'ils considéraient comme gênantes – par exemple, en montrant que les commentateurs de droite comme Ben Shapiro et Dan Bongino obtenaient beaucoup plus de visites sur leurs pages Facebook que les médias traditionnels.

Pour ces cadres, Facebook devrait divulguer de manière sélective ses propres données dans des rapports élaborés avec rigueur, plutôt que de fournir aux personnes extérieures les outils leur permettant de les découvrir.

Les tenants de la divulgation sélective ont gagné, et CrowdTangle et ses défenseurs ont perdu.



### ***The Thinking Conservative***

Une bataille interne sur la transparence des données peut sembler secondaire sur la liste des enquêtes à mener sur Facebook. Et c'est une chronique que depuis des mois j'hésite à écrire, en partie parce que je suis bien trop proche de l'action. (Nous y reviendrons).

Mais l'histoire de CrowdTangle est importante, car elle illustre la manière dont l'obsession de Facebook pour la gestion de sa réputation entrave souvent ses tentatives pour faire le ménage sur la plate-forme. Et elle est au cœur des tensions majeures auxquelles Facebook doit faire face dans l'ère post-Trump. L'entreprise, accusée de tout, depuis l'ingérence dans les élections jusqu'aux réticences à la vaccination, veut absolument rétablir la confiance d'un public sceptique. Mais plus elle partage ce qui se passe sur sa plate-forme, plus elle risque d'exposer des vérités très gênantes qui pourraient nuire davantage à son image.

Selon des entretiens avec plus d'une douzaine d'employés actuels et anciens de Facebook, ainsi que des courriels et des messages internes, la question de savoir ce qu'il convient de faire avec CrowdTangle préoccupe certains des principaux dirigeants de Facebook depuis des mois.

Ces personnes, dont la plupart n'ont accepté de parler que sous couvert d'anonymat parce qu'elles n'étaient pas autorisées à divulguer des conversations internes, ont déclaré que le management de Facebook était plus soucieux de corriger la perception voulant que Facebook amplifie les contenus à caractère préjudiciable que de déterminer si ces contenus étaient réellement amplifiés. La transparence, ont-ils précisé, a finalement été reléguée au second plan par rapport à la gestion de l'image.

Facebook conteste cette interprétation. Selon la direction de Facebook, la réorganisation de CrowdTangle visait à associer cette fonction à ses autres outils de transparence, et non à l'affaiblir, et les cadres supérieurs sont toujours déterminés à accroître la transparence. « CrowdTangle fait partie d'une gamme croissante de technologies de transparence que nous avons mises à la disposition des gens, y compris des universitaires et des journalistes, a déclaré Joe Osborne, porte-parole de Facebook. Avec le transfert de CrowdTangle dans notre équipe d'intégrité, nous développons une stratégie plus complète sur la façon dont nous allons développer certains aspects de la transparence à l'avenir. »

Mais les cadres qui ont exercé les plus fortes pressions en faveur de la transparence semblent avoir été mis sur la touche. M. Silverman, cofondateur et directeur général de CrowdTangle, est en congé et son rôle au sein de l'entreprise n'est plus clairement défini, selon plusieurs personnes au fait de la situation. (M. Silverman a refusé de commenter sa situation). Et M. Boland, qui a passé 11 ans chez Facebook, a quitté la société en novembre.

« L'une des principales raisons pour lesquelles j'ai quitté Facebook est que les plus hauts dirigeants de l'entreprise ne veulent pas s'investir pour comprendre l'impact de ses principaux produits, a déclaré M. Boland, dans la première interview qu'il a accordé depuis son départ. Et ils ne veulent pas mettre les données à la disposition d'autres gens afin

que ceux-ci puissent faire le travail difficile et leur demandent des comptes. »

M. Boland, qui a dirigé CrowdTangle, tout comme il a supervisé d'autres efforts quant à la transparence de Facebook, a déclaré que l'outil est tombé en disgrâce auprès des cadres influents de Facebook à peu près au moment de l'élection présidentielle de l'année dernière, lorsque les journalistes et les chercheurs l'ont utilisé pour montrer que les commentateurs pro-Trump diffusaient des fausses informations et des commentaires hyper partisans et que ceux-ci rencontraient un succès époustouflant.

« Les gens étaient enthousiasmés par la transparence que CrowdTangle offrait jusqu'à ce que cela devienne un problème et engendre des séries d'articles de presse que Facebook n'aimait pas, a-t-il déclaré. Ensuite, le ton a changé au niveau de la direction. »

### Le compte Twitter qui a initié 1 000 réunions

C'est ici que j'interviens, un peu à contrecœur. J'ai commencé à utiliser CrowdTangle il y a quelques années. Je cherchais un moyen de savoir quels articles d'informations avaient le plus de succès sur Facebook, et CrowdTangle – un outil utilisé principalement par les équipes de gestion de l'audience chez les éditeurs de presse et les spécialistes du marketing qui souhaitent suivre les performances de leurs publications – répondait à mes attentes.

J'ai découvert que, grâce à un bricolage astucieux, je pouvais utiliser sa fonction de recherche pour classer les publications de liens Facebook – c'est-à-dire les publications contenant un lien vers un site extérieur à Facebook – en fonction du nombre de réactions, de partages et de commentaires qu'elles avaient suscité. Les publications de liens n'étaient pas une mesure parfaite de l'actualité, le niveau de fréquentation n'était pas une mesure parfaite de la popularité et les données de CrowdTangle souffraient d'autres carences, mais c'était ce qui se rapprochait le plus d'une sorte de palmarès de l'actualité sur Facebook, alors je l'ai utilisé.



**Facebook's Top 10**  
@FacebooksTop10



**The top-performing link posts by U.S. Facebook pages in the last 24 hours are from:**

1. Dan Bongino
2. Dan Bongino
3. Fox News
4. Dan Bongino
5. Dan Bongino
6. Dan Bongino
7. Dan Bongino
8. Dan Bongino
9. CNN
10. Dan Bongino

8:47 AM · Oct 23, 2020 · Twitter Web App

**82** Retweets   **44** Quote Tweets   **353** Likes

*Quels articles ont été les plus performants dans les dernières 24h sur les pages US de Facebook ?*

Au début, Facebook a adoré que tant moi que d'autres journalistes ayons trouvé que son outil était utile. Avec seulement 25 000 utilisateurs, CrowdTangle est l'un des plus petits produits de Facebook, mais il est devenu une ressource précieuse pour les utilisateurs expérimentés, notamment les organisations mondiales de santé, les responsables électoraux et les spécialistes du marketing numérique, et il a donné à Facebook une image de transparence par rapport à des plate-formes concurrentes comme YouTube et TikTok, qui ne publient pas autant de données.

Mais l'ambiance a changé l'année dernière lorsque j'ai ouvert un compte Twitter appelé @FacebooksTop10, sur lequel j'ai publié, sur la base des données de CrowdTangle, un classement quotidien indiquant les sources des liens les plus actifs publiés par les pages américaines.

L'automne dernier, le classement était rempli de messages de M. Trump et de personnalités médiatiques pro-Trump. Depuis que M. Trump a été exclu de Facebook en janvier, la plate-forme est dominée par une poignée de polémistes de droite comme Messieurs Shapiro, Bongino et Sean Hannity, avec de temps en temps un article d'actualité grand public, une adorable histoire d'animal ou un blog de fan de K-pop [K-pop est un terme désignant plusieurs genres musicaux originaires de Corée du Sud, NdT].

Le compte est devenu quasi-viral, comptant plus de 35 000 abonnés (followers). Des milliers de personnes ont retweeté les listes, y compris des conservateurs heureux de voir les polémistes pro-Trump damer le pion aux grands médias et des libéraux qui les ont partagées avec des blagues du type « Admirez toute cette censure conservatrice ! » (Au cas où vous auriez été été sur une autre planète ces deux dernières années, les conservateurs aux États-Unis se plaignent fréquemment que Facebook les censure).

Les listes ont également attiré de nombreux détracteurs (haters) de Facebook. Les libéraux les ont partagées pour prouver que l'entreprise était un marécage toxique qui devait être démantelé ; les annonceurs progressistes se sont offusqués à l'idée que leur message apparaisse à côté de la propagande pro-Trump.

Le compte a même été cité par le représentant Jamie Raskin, Démocrate du Maryland, lors d'une audition au Congrès sur la technologie et l'antitrust, il a déclaré que cela prouvait que « si Facebook est là pour éliminer le discours conservateur, leur travail est vraiment nul. »

En interne, ce compte a fait bondir les dirigeants de Facebook. Certains pensaient que les données étaient mal analysées et s'inquiétaient de voir Facebook dépeint comme une chambre d'écho d'extrême droite.

D'autres craignaient que les listes ne fassent peur aux investisseurs en laissant entendre que les utilisateurs américains de Facebook vieillissaient et devenaient plus conservateurs.

Chaque fois qu'un tweet devenait viral, je recevais des appels de cadres de Facebook excédés par la disparité entre ce que Facebook était à leurs yeux – une place publique propre et bien éclairée où règnent la civilité et la tolérance – et l'image qu'ils en voyaient reflétée dans les listes Twitter.

À l'approche des élections l'année dernière, selon trois personnes qui y ont participé, les cadres de Facebook ont tenu des réunions pour déterminer ce qu'il convenait de faire. Ils ont cherché à déterminer si les informations de @FacebooksTop10 étaient exactes (elles l'étaient) et ont envisagé de créer un compte Twitter concurrent qui publierait des listes plus équilibrées basées sur les données internes de Facebook.

Ils ne l'ont jamais fait, mais plusieurs cadres – dont John Hegeman, responsable du fil d'actualité de Facebook – ont été envoyés pour débattre avec moi sur Twitter. Ces cadres ont fait valoir que les listes de mon Top 10 étaient trompeuses.

Ils ont déclaré que CrowdTangle ne mesurait que « l'intérêt » (engagement), alors que le véritable indicateur de popularité sur Facebook devrait être basé sur la « fréquentation » (reach), c'est-à-dire le nombre de personnes qui voient effectivement une publication donnée. (À l'exception des vues de vidéos, les données sur la « fréquentation » ne sont pas publiques, et seuls les employés de Facebook et les propriétaires de pages y ont accès).

En septembre dernier, Mark Zuckerberg, directeur général de Facebook, a déclaré à *Axios* que si les contenus de droite suscitaient beaucoup d'intérêt, l'idée que Facebook était une chambre d'écho de droite était « tout simplement fausse. »

« Je pense qu'il est important de faire la différence entre ça et, plus globalement, ce que les gens voient, lisent et apprennent sur notre plate-forme », a déclaré M. Zuckerberg.

Mais M. Boland, ancien vice-président de Facebook, a déclaré qu'il s'agissait d'une diversion bien commode. Il a déclaré que, lors de discussions internes, les dirigeants de Facebook étaient moins préoccupés par l'exactitude des données que par l'image de Facebook qu'elles donnaient.

« Cela racontait une histoire qu'ils n'aimaient pas, a-t-il dit à propos du compte Twitter, et franchement, ils ne voulaient pas admettre que c'était vrai. »



**Sur cette photo d'archives du 1er mai 2018, le PDG de Facebook, Mark Zuckerberg, prononce le discours d'ouverture du F8, la conférence des développeurs de Facebook, à San Jose, en Californie. (AP Photo/Marcio Jose Sanchez, dossier)**

### **Le problème avec CrowdTangle**

À peu près au même moment que celui où M.Zuckerberg a fait ses commentaires à *Axios*, les tensions ont atteint leur paroxysme. *The Economist* venait de publier un article affirmant que Facebook « offre une vision déformée des informations américaines. » L'article, qui citait des données de CrowdTangle, montrait que les sites d'information américains les plus consultés sur Facebook étaient *Fox News* et *Breitbart*, et affirmait que l'écosystème d'information global de Facebook penchait à droite.

Par courriel, John Pinette, vice-président de la communication mondiale de Facebook, a envoyé un lien vers l'article à un groupe de cadres avec pour objet « Le problème avec CrowdTangle. » « *The Economist* prend le train de Kevin Roose en marche », écrit M. Pinette. (Vous voyez ? Je vous avais bien dit que c'était inconfortablement proche).

Nick Clegg, vice-président des affaires mondiales de Facebook, a répondu en déplorant que « nos propres outils aident les journalistes à donner crédit à la mauvaise histoire. »

D'autres cadres sont intervenus, s'inquiétant de voir les données de CrowdTangle être utilisées pour présenter Facebook comme une chambre d'écho de droite.

David Ginsberg, vice-président de Facebook chargé des choix et de la concurrence, a écrit que si M.Trump était réélu en novembre, « les médias et nos détracteurs ne tarderaient pas à désigner cette « chambre d'écho » comme l'un des principaux vecteurs du résultat. »

Fidji Simo, responsable de l'application Facebook à l'époque, a approuvé. « Je crains vraiment que ce soit l'un des pires messages nous concernant », a-t-elle écrit.

Plusieurs cadres ont proposé de rendre publiques les données relatives à la fréquentation sur CrowdTangle, dans l'espoir que les journalistes citent ces données au lieu des données relatives à l'intérêt qui, selon eux, donnent une mauvaise image de Facebook.

Mais M. Silverman, directeur général de CrowdTangle, a répondu dans un courriel que l'équipe de CrowdTangle avait déjà testé une fonction à cet effet et qu'elle avait rencontré des problèmes. L'un des problèmes étant que les nouvelles

fausses et trompeuses se retrouvaient également en tête de ces listes. « Le classement par fréquentation n'est pas une victoire totale du point de vue de la communication », a écrit M. Silverman.

M. Schultz, directeur du marketing de Facebook, était celui qui avait la vision la plus sombre de CrowdTangle. Il a écrit qu'il pensait que « la seule façon d'éviter ce genre d'histoires » serait que Facebook publie ses propres statistiques sur les contenus les plus populaires de sa plate-forme, plutôt que de diffuser des données par le biais de CrowdTangle. « Si nous nous contentons de fournir davantage de données en libre-service, nous aurons droit à des histoires différentes, passionnantes et négatives à mon avis », a-t-il écrit.

M. Osborne, porte-parole de Facebook, a déclaré que M. Schultz et les autres dirigeants réfléchissaient à la façon de corriger les déformations des données produites par CrowdTangle, et non de la stratégie à adopter pour supprimer l'outil.

Quelques jours après l'élection de novembre, M. Schultz a écrit un article pour le blog de l'entreprise, intitulé « Que regardent réellement les gens sur Facebook aux États-Unis ? ». Il a expliqué que si on classait les publications sur Facebook en fonction de leur fréquentation plutôt que de leur degré d'intérêt (sa méthode préférée pour fractionner les données), on obtiendrait une liste de sources plus générale et moins nettement partisane. « Nous pensons que cela donne une image plus complète que les seules données de CrowdTangle », a-t-il écrit.

Il se pourrait que ce soit vrai, mais il y a un problème avec les données de fréquentation : la plupart d'entre elles sont inaccessibles et ne peuvent pas être vérifiées ou contrôlées par des tiers. Il nous faut simplement croire que les propres données privées de Facebook racontent une histoire bien différente des données partagées avec le public.

### **Bidouillage des variables**

M. Zuckerberg a raison sur un point : Facebook n'est pas une gigantesque chambre d'écho de droite.

Mais il contient cependant une chambre d'écho géante de droite – une sorte de station de radio AM intégrée au cœur même de l'écosystème d'information de Facebook, avec un public hyper-engagé de partisans loyaux qui aiment aimer (like), partager (share) et cliquer sur les publications des pages de droite, dont beaucoup sont devenues expertes dans l'art de servir des contenus racoleurs à scandale optimisés par Facebook à un rythme régulier.

Les données de CrowdTangle ont rendu cette chambre d'écho plus facile à voir et à quantifier pour les personnes extérieures. Mais ce n'est pas ce qui l'a créée, ni ne lui a donné les outils dont elle avait besoin pour se développer – c'est Facebook qui l'a fait – et blâmer un outil de données pour ces révélations n'a pas plus de sens que de blâmer un thermomètre pour le mauvais temps.

Il convient de noter que ces efforts de transparence relèvent du consentement et peuvent cesser à tout moment. Il n'existe aucune réglementation obligeant Facebook ou toute autre société de médias sociaux à révéler quels sont les contenus qui se vendent bien sur leurs plate-formes, et les politiciens américains semblent plus intéressés par la lutte contre les allégations de censure que par une amélioration de la qualité des données.

Il convient également de noter que Facebook peut à tout moment choisir, désactiver les cadrans de l'indignation et diffuser à ses utilisateurs des informations plus apaisées et moins clivantes. (En fait, il l'a brièvement fait après l'élection de 2020, lorsqu'il s'est inquiété de voir la désinformation liée à l'élection éventuellement conduire à une spirale de violence de masse). Et il y a des preuves que des changements plus permanents sont au moins envisagés.

Selon deux personnes au courant du projet, M. Hegeman, responsable du fil d'actualité de Facebook, a, cette année, demandé à une équipe de déterminer comment le fait de modifier certaines variables de l'algorithme de base du classement du fil d'actualité changerait les listes du Top 10.

Le projet, que certains employés appellent le projet « Top 10 », est toujours en cours, ont indiqué ces personnes, et il est difficile de savoir si ses conclusions ont été mises en œuvre. M. Osborne, porte-parole de Facebook, a déclaré que l'équipe examinait divers changements pour le classement et que cette expérimentation n'était pas motivée par le désir de modifier les listes du Top 10.

Quant à CrowdTangle, l'outil est toujours disponible, et à court terme, Facebook ne devrait pas en couper l'accès aux journalistes et aux chercheurs, selon deux personnes au courant des projets de l'entreprise.

M. Boland a toutefois déclaré qu'il ne serait pas surpris que les dirigeants de Facebook décident de supprimer complètement CrowdTangle ou de le priver de ressources, plutôt que de s'occuper des maux de tête causés par ses données.

**« Facebook aimerait une transparence totale s'il y avait une garantie que tant les histoires que les résultats soit positifs, a déclaré M. Boland. Mais lorsque la transparence crée des moments inconfortables, la réaction est**

# Facebook Algorithm Ranking Signals in 2021



## Relationship

Who a user typically interacts with



## Content Type

The type of media in the post (e.g., video, link, image, etc.)



## Popularity

How many likes/engagements the post gets



## Recency

Newer posts are shown first

