



CONSEIL SUPÉRIEUR DE L'AUDIOVISUEL

# FOCUS

A large, light blue circle is centered behind the word 'FOCUS'. To the left of the circle, a thin red arc curves upwards. To the right, a yellow stepped line extends horizontally and then vertically. Below the circle, a dotted line of small white dots curves downwards and to the right. A thin green horizontal line is positioned below the letter 'O' in 'FOCUS'.

## LA PROPAGATION DES FAUSSES INFORMATIONS SUR LES RÉSEAUX SOCIAUX : *ÉTUDE DU SERVICE TWITTER*

Novembre 2020





## Sommaire

<b>1. Synthèse.....</b>	<b>5</b>
<b>2. Observations de Twitter.....</b>	<b>8</b>
<b>3. Introduction.....</b>	<b>11</b>
3.1. Éléments de définition et champ d'étude.....	12
3.2. L'enjeu fort posé par les réseaux sociaux en matière d'accès à l'information et aux fausses informations .....	13
3.3. Description du service Twitter .....	14
3.4. Les objectifs de l'étude.....	15
3.5. La collecte de données sur Twitter.....	16
3.6. Les précautions prises par le CSA afin de s'assurer du respect du règlement général sur la protection des données.....	16
<b>4. Analyse de l'activité des comptes d'information fiables, non fiables et des comptes     de fact-checkers .....</b>	<b>18</b>
4.1. Présentation du <i>Décodex</i> .....	18
4.2. Analyse statistique des différentes catégories de comptes.....	19
4.3. Analyse du contenu des Tweets des comptes connus comme étant propices à la diffusion de fausses informations .....	26
4.4. Les analyses ne font pas apparaître de phénomène de chambre d'écho sur Twitter ...	34
4.5. Analyses du graphe des abonnés.....	40
4.6. Examen de l'activité des comptes des vérificateurs d'information .....	42
<b>5. Analyse des Tweets liés à de fausses informations : chronologie de la propagation     et de la correction .....</b>	<b>54</b>
5.1. Description des fausses informations sélectionnées et de la méthodologie de collecte de données.....	54
5.2. Méthodologie de qualification des Tweets collectés .....	57
5.3. Analyse de la base de Tweets qualifiée .....	59
5.4. Analyse de l'audience potentielle de chaque catégorie de Tweet .....	61
5.5. Analyse des délais de correction .....	63
5.6. Analyse des messages les plus retweetés.....	64
<b>6. Conclusion .....</b>	<b>68</b>



<b>7. Annexes.....</b>	<b>70</b>
7.1. Chronologie de diffusion des Tweets.....	70
7.2. Évolution de la répartition des Tweets par catégorie .....	71
7.3. Évolution de la volumétrie des Tweets par catégorie .....	72
<b>8. Bibliographie .....</b>	<b>73</b>
<b>9. Remerciements .....</b>	<b>81</b>



## 1. Synthèse

Suite à la promulgation de la loi relative à la lutte contre la manipulation de l'information (2018), le Conseil supérieur de l'audiovisuel a souhaité renforcer sa connaissance des mécanismes de propagation des fausses informations en initiant une étude spécifique. La pandémie de Covid-19 a rendu plus aigüe encore la problématique des fausses informations à l'ère où l'accès à l'information passe par une pluralité de médias, dont les réseaux sociaux.

Cette étude s'intéresse plus précisément aux mécanismes de propagation et de correction des fausses informations sur le réseau social Twitter. Pour les besoins de cette étude, des données concernant les comptes d'information francophones actifs sur Twitter et des données relatives aux Tweets concernant certaines fausses informations ont été collectées à l'aide de l'API mise à disposition par le réseau social. Dans le cadre de cette étude, une information est considérée comme fautive si elle a été désignée comme telle par des journalistes spécialisés dans la vérification de l'information. De la même façon, des comptes Twitter seront considérés comme fiables ou susceptibles de partager des fausses informations en fonction de la classification opérée par certains vérificateurs d'information (en particulier les Décodeurs du Monde dans le cadre de l'initiative du *Décodex*).

La première partie de l'étude est consacrée à l'analyse des comptes d'information les moins fiables. Ces derniers possèdent un nombre d'abonnés sur Twitter significativement plus faible que la majorité des comptes d'information fiables. Cependant, les comptes connus pour partager de fausses informations font jeu égal, en termes de *ReTweets*, avec les comptes fiables. Les abonnés des comptes non fiables ont en effet une propension beaucoup plus élevée à contribuer à la diffusion des informations diffusées sur ces comptes que les abonnés des comptes fiables (avec un nombre de *ReTweets* par abonnés de 10 à 20 fois plus élevés selon l'indicateur retenu).

Afin de mieux appréhender les ressorts et l'impact de la viralité de ces comptes, l'étude propose une analyse qualitative de leurs Tweets les plus partagés. Il en ressort que les comptes identifiés comme non fiables privilégient largement des thématiques d'actualité et liées à des sujets clivants. Ainsi les Tweets traitant de politique, de l'immigration, de la santé, de la religion et du terrorisme représentent plus de la moitié du corpus étudié. L'analyse quantitative montre la surreprésentation de termes liés la délinquance, l'immigration, Israël et la Palestine, la pédophilie, l'Islam ou encore la franc-maçonnerie. La tonalité utilisée relève principalement de **l'information de leurs abonnés et de la critique des personnes ou institutions citées**. L'utilisation de contenus non textuels (images, vidéos, liens, etc.) semble par ailleurs caractériser une stratégie de construction d'une légitimité des contenus publiés, reposant majoritairement sur des sites liés à ces comptes mais également sur des sources issues de médias traditionnels.



Les individus exposés à un phénomène de chambre d'échos<sup>1</sup> sont minoritaires parmi les abonnés aux comptes les moins fiables de Twitter. En effet, moins de 20 % des abonnés à ces comptes ne suivent aucun compte fiable. Au contraire, ces individus suivent en moyenne une majorité de comptes fiables. Certains individus (minoritaires) présentent cependant des comportements d'abonnements qui ne leur permettent pas d'être confrontés à des sources d'information susceptibles de corriger les fausses informations reçues par ailleurs. Une analyse en termes de graphe fait par ailleurs ressortir que de nombreux comptes peu fiables possèdent une proportion significative d'abonnés communs ou se suivent mutuellement sur Twitter.

La première partie de l'étude se conclut avec une analyse des comptes Twitter de vérificateurs d'informations affiliés à des médias traditionnels. Les comptes de cellules journalistiques spécialisées dans la vérification de l'information sont en moyenne davantage suivis que les comptes catégorisés comme non fiables, mais génèrent moins d'interactions sur leurs publications que ces derniers. En termes de thématiques majoritairement traitées, seules la politique et la religion sont communes entre ces deux types de comptes. Les vérificateurs d'information ou *fact-checkers* s'intéressent davantage à des thématiques en lien avec leur métier, comme l'actualité des médias ou des contenus qui relèvent de l'éducation aux médias et à l'information.

L'étude présente ensuite de façon plus détaillée une analyse de Tweets collectés se rapportant à des fausses informations préalablement sélectionnées, que ces Tweets les aient propagées ou réfutées.

Toutes les fausses informations étudiées présentent une forte concentration de Tweets sur une période de temps très réduite. Cette observation conduit à souligner la difficulté du rôle des vérificateurs d'informations, qui doivent réagir de façon très rapide afin d'avoir un impact sur la propagation de ces fausses informations. Les fausses informations étudiées se partagent par ailleurs équitablement entre les thèmes pour lesquels la fausse information est partagée de façon majoritaire et ceux pour lesquels c'est au contraire l'information vérifiée qui est majoritairement partagée. Cependant, pour les thèmes qui présentent la volumétrie de Tweets la plus élevée, la fausse information est dans tous les cas majoritaire. L'étude de la chronologie de la diffusion des fausses informations montre que, contrairement à ce qui aurait pu être attendu ou espéré, **les « vraies » informations ne chassent pas les fausses**. Pour les sujets sur lesquels les fausses informations sont majoritaires, la discussion sur ces sujets s'éteint rapidement sur Twitter sans que l'information vérifiée n'ait jamais occupé une place majoritaire.

Cette analyse est complétée par une analyse du nombre d'abonnés aux comptes ayant partagé des fausses informations ou des informations avérées. Comme les informations vérifiées sont fréquemment partagées par des comptes ayant un nombre d'abonnés élevé (notamment lorsqu'il s'agit de comptes attachés à des médias traditionnels), l'audience des informations vérifiées est plus importante que ce qu'une analyse menée sur le seul volume de Tweets pourrait laisser paraître. Pour certains thèmes, la fausse information reste cependant plus visible que l'information vérifiée.

---

<sup>1</sup> Entendu dans le cadre de cette étude comme consistant à être exposé à des sources connues pour partager des fausses informations, sans dans le même temps suivre de comptes fiables.



L'analyse des Tweets ayant témoigné de la viralité la plus importante dans le corpus étudié permet enfin de faire ressortir plusieurs traits intéressants. Les fausses informations sont majoritairement le support d'une critique adressée aux autorités ou expriment une forme de panique, par exemple sur des sujets sanitaires sensibles. Elles s'appuient sur des informations non recoupées ou qualifiées de fausses par les journalistes, qui apportent de leur côté des correctifs empruntant les codes spécifiques aux réseaux sociaux. Ces correctifs récoltent néanmoins un faible degré d'engagement (ReTweet, commentaires) par rapport aux comptes susceptibles de diffuser de fausses informations. Il faut néanmoins relever que dans plusieurs des cas étudiés, le correctif a pu être directement apporté par des internautes, qui se sont appuyés sur les éléments de ces articles de médias.



## 2. Observations de Twitter

Dans une démarche de coopération avec les plateformes numériques, le Conseil a partagé avec Twitter les objectifs et résultats de ce rapport avant sa publication. Les représentants de la plateforme ont pu soumettre leurs remarques au Conseil. Celui-ci a pris en compte celles d'entre elles qu'il jugeait pertinentes dans la rédaction finale du rapport.

Il a par ailleurs proposé à Twitter de disposer d'un espace d'expression dans ce rapport afin d'apporter toutes précisions que la société jugeait utile, notamment sur les efforts qu'elle met en œuvre en matière de lutte contre la manipulation de l'information. Celui-ci est consultable ci-après :

### Contribution de Twitter



#### **Notre approche pour garantir la sérénité de la conversation publique**

En servant la conversation publique, notre objectif est de faciliter la recherche d'informations crédibles sur Twitter et de limiter la diffusion de contenus potentiellement préjudiciables et trompeurs.

Twitter reconnaît que la désinformation est un problème sociétal complexe qui mérite une solution d'ensemble. À cette fin, nous nous engageons à travailler avec la Commission européenne, les gouvernements, les régulateurs, la société civile et les institutions universitaires afin d'élaborer des solutions collaboratives et réalisables. Nous sommes signataires du Code de conduite de l'UE en matière de désinformation, nous avons établi un partenariat avec l'UNESCO sur l'éducation aux médias et nous publions des ensembles de données sur les campagnes de désinformation à l'étranger afin de promouvoir une compréhension commune des tactiques et des menaces des mauvais acteurs.

Nous ne nous reposons pas non plus uniquement sur l'interaction des titulaires de comptes. Nous prenons des mesures proactives pour tenter d'empêcher la diffusion de comptes et de Tweets malveillants, ce qui implique de tenter de rester en avance sur les tactiques en constante évolution des mauvais acteurs. Pour tenter d'atteindre cet objectif, nous avons apporté un certain nombre de changements, notamment des politiques de sécurité renforcées, de meilleurs outils et ressources pour détecter et arrêter les activités malveillantes, des règles publicitaires plus strictes et une transparence accrue pour promouvoir la compréhension du public dans tous ces domaines.

#### **Ce que nous faisons contre la désinformation sur la pandémie COVID-19**

Alors que la communauté mondiale fait face à la pandémie COVID-19, Twitter aide les gens à trouver des informations fiables, à se connecter avec d'autres et à suivre ce qui se passe en temps réel.



Tout au long de cette période sans précédent, Twitter a continué d'adapter et d'actualiser ses politiques et ses actions, ainsi que d'accroître la transparence et de partager davantage de données afin que les experts et le public puissent mieux analyser l'évolution du débat autour du COVID-19.

Nous avons élargi nos orientations politiques pour traiter des contenus qui vont directement à l'encontre des orientations sur la COVID-19 provenant de sources d'information sur la santé publique mondiale et locale qui font autorité. Nous avons introduit de nouvelles étiquettes et de nouveaux messages d'avertissement qui fournissent un contexte et des informations supplémentaires sur certains Tweets contenant des informations contestées ou trompeuses liées à la COVID-19. Il sera ainsi plus facile de trouver des faits et de prendre des décisions éclairées sur ce que les gens voient sur Twitter.

### **Notre travail pour protéger la conversation civique**

La conversation qui se déroule sur Twitter n'est jamais plus importante que pendant les élections. C'est sur Twitter que les gens viennent s'informer directement auprès des élus et des candidats, c'est là qu'ils trouvent les dernières nouvelles et, de plus en plus, c'est une source importante d'informations sur le moment et la manière de voter aux élections. Alors que de plus en plus de personnes cherchent des moyens de voter et d'exprimer leurs droits civils fondamentaux en toute sécurité pendant la pandémie COVID-19, le besoin de ce type d'information n'a fait que croître.

Notre politique d'intégrité civique cible donc les types de contenu les plus directement préjudiciables, à savoir ceux qui concernent les informations ou fausses déclarations sur la façon de participer à des processus civiques ; les contenus pouvant intimider ou supprimer la participation et la fausse affiliation. Compte tenu de l'évolution de la façon dont les gens voteront en 2020, et conformément à notre engagement à protéger l'intégrité de la conversation électorale, en septembre 2020, nous avons élargi ce cadre existant. L'objectif est de renforcer la protection contre les contenus susceptibles d'altérer le vote et de contribuer à mettre un terme à la diffusion d'informations erronées préjudiciables susceptibles de compromettre l'intégrité d'une élection ou d'un autre processus civique. Nous allons maintenant également étiqueter ou supprimer les informations fausses ou trompeuses destinées à saper la confiance du public dans une élection ou un autre processus civique. Conformément à notre approche actuelle en matière d'application des règles, les Tweets étiquetés dans le cadre de cette politique élargie auront une visibilité réduite dans l'ensemble du service. La réduction de la visibilité des Tweets signifie que nous n'amplifions pas les Tweets sur un certain nombre de surfaces à travers Twitter.

### **Transparence**

La transparence est au cœur du travail que nous faisons sur Twitter. La nature ouverte de notre service a entraîné des défis sans précédent en matière de protection de la liberté d'expression et du droit à la vie privée, alors que les gouvernements du monde entier tentent de plus en plus d'intervenir dans cet échange ouvert d'informations. Nous pensons que la transparence est un principe clé dans notre mission de protéger l'Internet ouvert et de faire progresser l'Internet en tant qu'outil du progrès.



La croyance fondamentale dans le pouvoir d'une conversation publique et ouverte a inspiré Twitter à lancer l'un des premiers rapports de transparence du secteur en 2012.

### **Accès aux données**

Depuis 2006, les API de Twitter ont donné aux chercheurs et aux développeurs la possibilité d'accéder à ce qui se passe dans le monde. Twitter croit fermement au libre accès aux données pour étudier, analyser et contribuer à la conversation publique ; c'est pourquoi nous continuons à maintenir une API grand public. Les chercheurs utilisent les données de Twitter pour fournir un retour d'information précieux sur la façon dont les conversations et les interactions en ligne évoluent sur et hors de Twitter. Nous continuons à fournir des moyens plus accessibles pour rendre les données et les informations accessibles au public et aux chercheurs.

Notre service est la plus grande source de données en temps réel sur les réseaux sociaux, et nous mettons ces données gratuitement à la disposition du public par le biais de notre API publique. Aucun autre service important ne le fait.

Pour soutenir davantage nos efforts en cours pour protéger la conversation publique et aider les gens à trouver des informations de santé faisant autorité autour du COVID-19, Twitter a publié un point d'accès spécifique de l'API COVID-19 dans les laboratoires de développeurs Twitter pour permettre aux développeurs et chercheurs approuvés d'étudier la conversation publique sur la COVID-19 en temps réel. Il s'agit d'un ensemble de données unique qui couvre plusieurs dizaines de millions de Tweets par jour et qui offre un aperçu de l'évolution de la conversation publique mondiale autour d'une crise sans précédent. Rendre cet accès disponible gratuitement est l'une des choses les plus uniques et les plus précieuses que Twitter puisse faire alors que le monde se rassemble pour protéger nos communautés et chercher des réponses aux défis urgents.

Le 12 août, nous avons également présenté la nouvelle API de Twitter. Reconstituée de fond en comble pour offrir de nouvelles fonctionnalités plus rapidement, cette version comprend le premier ensemble de nouveaux terminaux et de nouvelles fonctionnalités afin que les développeurs puissent aider le monde à se connecter à la conversation publique qui se déroule sur Twitter.

### **Opérations d'information**

Depuis 2019, nous publions tous les comptes et contenus associés aux opérations de désinformation potentielles que nous avons trouvés sur notre service depuis 2016. Les archives sont désormais les plus importantes de ce type dans le secteur. En 2019, de nouveaux corpus de données ont été ajoutés aux archives, permettant ainsi aux membres du public, aux gouvernements, aux universitaires et aux chercheurs d'étudier ces comportements, d'apprendre et de renforcer les capacités d'éducation aux médias pour l'avenir afin de lutter contre la manipulation et la désinformation. Les données comprennent plus de 160 millions de Tweets et plus de huit téraoctets de médias.



### 3. Introduction

Le Conseil supérieur de l'audiovisuel (CSA), autorité publique en charge de la régulation de la communication audiovisuelle, est le garant de **la liberté d'expression, dans l'intérêt du public et des professionnels du secteur**. La loi relative à la lutte contre la manipulation de l'information du 22 décembre 2018 lui confie de nouvelles compétences visant à répondre à certains des enjeux pour la démocratie qui sont apparus dans l'univers numérique.

Le texte instaure un devoir de coopération en matière de lutte contre la diffusion de fausses informations à la charge des opérateurs de plateformes en ligne<sup>2</sup>. Sont plus précisément concernées les plateformes dont l'activité dépasse cinq millions de visiteurs uniques par mois<sup>3</sup>. Le CSA leur a adressé des recommandations<sup>4</sup>, qui concernent notamment, conformément à l'article 11 de la loi précitée, la mise en place d'un dispositif de signalement accessible et visible, 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 la promotion des contenus d'information et l'éducation aux médias et à l'information. Le CSA s'est également entouré d'un comité d'experts sur la lutte contre la désinformation en ligne qui apporte à ses travaux un éclairage économique, scientifique, technologique et éthique.

La **pandémie de la Covid-19** a rendu plus aigüe encore la problématique des fausses informations à l'ère où l'accès à l'information passe par une pluralité inédite de médias, dont les réseaux sociaux<sup>5</sup>. L'Organisation mondiale de la santé a ainsi évoqué, en parallèle de l'épidémie de la Covid-19, une « *infodémie* ». Selon son directeur général, « *les informations fausses se propagent plus vite et plus facilement que ce virus, et elles sont tout aussi dangereuses* »<sup>6,7</sup>. C'est également dans ce contexte que la Commission européenne a rendu publique la consultation visant à définir les grandes lignes du Digital Services Act, un texte qui devrait être présenté à la

---

<sup>2</sup> La définition des opérateurs de plateformes retenue dans le décret d'application de la loi du 22 décembre 2018 est celle de l'article L. 111-7 du code de la consommation : « toute personne physique ou morale proposant, à titre professionnel, de manière rémunérée ou non, un service de communication au public en ligne reposant sur : 1° Le classement ou le référencement, au moyen d'algorithmes informatiques, de contenus, de biens ou de services proposés ou mis en ligne par des tiers ; 2° ou la mise en relation de plusieurs parties en vue de la vente d'un bien, de la fourniture d'un service ou de l'échange ou du partage d'un contenu, d'un bien ou d'un service ».

<sup>3</sup> Ce seuil a été fixé par le décret n° 2019-297 du 10 avril 2019 relatif aux obligations d'information des opérateurs de plateforme en ligne assurant la promotion de contenus d'information se rattachant à un débat d'intérêt général.

<sup>4</sup> Recommandation n° 2019- 03 du 15 mai 2019 du Conseil supérieur de l'audiovisuel aux opérateurs de plateforme en ligne dans le cadre du devoir de coopération en matière de lutte contre la diffusion de fausses informations.

<sup>5</sup> Selon le Digital News Report 2019 du Reuters Institute for the Study of Journalism, l'usage des réseaux sociaux pour s'informer est en croissance en France et atteint 42 %. Les Français s'informent désormais quasiment autant en ligne (69 %), en incluant les réseaux sociaux, que via la télévision (71 %).

<sup>6</sup> OMS. *Conférence sur la sécurité Munich*. 15/02/2020.

<sup>7</sup> Parmi les fausses informations les plus partagées sur le coronavirus, on peut citer l'affirmation selon laquelle les ondes 5G seraient responsables de la pandémie, celle selon laquelle le coronavirus contiendrait des insertions de VIH créées par l'homme, celle selon laquelle la pulvérisation de chlore tuerait le virus, celle selon laquelle le « vaccin de Bill Gates » contiendrait une puce de suivi ou encore celle selon laquelle la vitamine C serait un traitement efficace contre le virus. Digimind, juillet 2020, « *Covid : les Fake News les plus répandues sur les médias sociaux en France* ».



fin de l'année 2020 et qui entend notamment faire évoluer le régime de responsabilité des opérateurs de plateformes<sup>8</sup>.

C'est dans ce contexte qu'intervient cette étude sur les phénomènes de propagation des fausses informations. Elle ne vise cependant ni à vérifier le respect par Twitter de ses obligations ni à dresser un bilan de son action pendant la pandémie. Elle n'a pas plus pour objectif de réaliser un état des lieux transversal de l'ensemble des problématiques soulevées par la diffusion des fausses informations. En s'appuyant sur des **analyses quantitatives et qualitatives**, elle propose un éclairage sur **les phénomènes de propagation des fausses informations** sur un réseau social dont l'usage est significatif en France.

Ce travail vient enrichir la connaissance du Conseil dans l'exercice de ses missions et témoigne de sa montée en compétence technique en matière d'outils et de méthodes de traitements des données. Il vise également à contribuer à une meilleure compréhension des phénomènes de diffusion des fausses informations par le grand public et par la sphère publique.

### 3.1. Éléments de définition et champ d'étude

L'attention croissante portée à la propagation de fausses informations a fait entrer dans le vocabulaire commun plusieurs concepts, parfois utilisés de manière équivalente, mais qui décrivent pourtant des réalités parfois différentes.

La **désinformation** recouvre les « *informations dont on peut vérifier qu'elles sont fausses ou trompeuses, qui sont créées, présentées et diffusées dans un but lucratif ou dans l'intention délibérée de tromper le public et qui sont susceptibles de causer un préjudice public* »<sup>9</sup>. Il s'agit d'une technique utilisée par exemple dans le cadre de campagnes visant à déstabiliser un groupe de personnes, une entreprise voire une élection. Ce terme, ainsi que l'expression « **fake news** » et sa traduction française – « fausses informations » - a été popularisé par les enquêtes sur la possible ingérence russe dans la campagne des élections présidentielles américaines de 2016. La notion de « fausses informations » ne permet cependant pas d'embrasser toutes les facettes du phénomène de la manipulation de l'information car « *la plupart du temps, le manipulateur ne se positionne pas par rapport à la vérité : il cherche simplement à produire un effet* »<sup>10</sup>.

La **mésinformation** quant à elle est une information dont l'inexactitude « *n'est pas intentionnelle* »<sup>11</sup>, mais est due par exemple à une mauvaise interprétation de données. Des internautes peuvent également partager des contenus issus de leurs contacts de bonne foi, sans pour autant en avoir vérifié l'origine, contribuant à propager de fausses informations.

Avec ces différents concepts est également apparue l'idée d'une **vérification de l'information**, rendue nécessaire par « *une vague toujours plus forte de fausses informations* »<sup>12</sup>. Également

<sup>8</sup> Commission européenne. *The Digital Services Act package*. 11/06/2020.

<sup>9</sup> Commission européenne. *A Multi-Dimensional Approach to Disinformation, Report of the Independent High Level Group on Fake News and Online Disinformation*. 03/2018.

<sup>10</sup> CAPS et IRSEM. *Les Manipulations de l'information : un défi pour nos démocraties*. 08/2018.

<sup>11</sup> *ibid.*

<sup>12</sup> Le Monde. *Le Décodex, un outil de vérification de l'information*. 23/01/2017.



appelée **fact-checking**, cette démarche consiste à « repérer et traiter une à une les informations qui [...] semblent mériter un éclairage »<sup>13</sup> notamment dans la perspective de les remettre en contexte et d'apporter un éclairage aux internautes. Des cellules composées de journalistes **fact-checkers** se sont développées au cours des dernières années dans les rédactions françaises, par exemple *FakeOff* à 20 Minutes, *AFP Factuel* pour l'Agence France-Presse, *Les Observateurs* de France 24, *Vrai ou fake* de France Info, *CheckNews* pour Libération ou encore *Les Décodeurs* du Monde.

Il ne relève pas de la compétence ni des missions du Conseil d'apprécier la véracité d'une information. La présente étude portera donc de façon générale sur les fausses informations sans chercher à identifier si elles visent un objectif de manipulation ou constituent des erreurs non intentionnelles. **Une fausse information sera considérée ici comme une information ayant fait l'objet d'une analyse par des journalistes, notamment ceux spécialisés dans le fact-checking, et qui ont constaté son caractère erroné.** Les sites, médias et comptes de réseaux sociaux identifiés par ces journalistes comme ayant déjà propagé de fausses informations seront également considérés ici comme des sources propices à la diffusion de celles-ci.

### 3.2. L'enjeu fort posé par les réseaux sociaux en matière d'accès à l'information et aux fausses informations

Selon le Digital News Report 2019 du Reuters Institute for the Study of Journalism, **l'usage des réseaux sociaux pour s'informer est en croissance en France et atteint 42 % de la population.** Les Français s'informent désormais quasiment autant en ligne (69 %), en incluant les réseaux sociaux, que via la télévision (71 %) <sup>14</sup>. Selon ce même rapport, à l'usage important des réseaux sociaux dans l'accès à l'information s'ajoute une défiance croissante envers les médias. Ainsi, **la confiance accordée par les Français dans les informations entendues au sens large est désormais la plus basse d'Europe**, à 24 % (en baisse de 11 points par rapport au précédent rapport). Les réseaux sociaux, encore très largement plébiscités dans leur usage, n'échappent pas à cette crise de la confiance : 14 % des sondés jugent qu'ils « *diffuse[nt] des théories conspirationnistes, partage[nt] des informations biaisées et filtre[nt] les informations grâce à des algorithmes* ».

Les fausses informations et leur propagation sur les réseaux sociaux portent donc des enjeux démocratiques forts, d'autant qu'ils ne sont pas le fait d'une seule catégorie de population mais peuvent concerner des catégories d'individus variées <sup>15</sup>.

Le choix a été fait de réaliser la présente étude à partir de données collectées sur le réseau social Twitter. Celui-ci autorise en effet l'accès aux données via une API <sup>16</sup>, condition préalable à la mise en œuvre de la méthodologie d'analyse envisagée par le Conseil. À l'inverse, Facebook, qui

<sup>13</sup> Ibid.

<sup>14</sup> Ibid.

<sup>15</sup> Andrew Guess, Jonathan Nagler, Joshua Tucker. *Science Advances Less than you think: Prevalence and predictors of fake news dissemination on Facebook*. 09/01/2019.

<sup>16</sup> Une API est une interface de programmation mise à disposition par un service afin d'établir des connexions entre plusieurs logiciels pour échanger des données. La mise à disposition d'une API permet aux administrateurs de définir les modalités d'utilisation (payante ou gratuite, seuil de données extractibles, etc.) et/ou les informations auxquelles les utilisateurs auront accès (inaccessibilité de données personnelles, impossibilité d'accéder à des résultats à partir d'une certaine date...).



compte un nombre considérable d'utilisateurs dans le monde (2,5 milliards<sup>17</sup>) comme en France (37 millions), a pris la décision, depuis 2018 et l'affaire « Cambridge Analytica », de restreindre l'accès à ses données, notamment en modifiant les conditions d'utilisation de ses API. La discussion engagée par les services du Conseil avec les responsables de la société afin de disposer de modalités d'accès aux contenus publics de la plateforme pour les besoins de l'étude n'a pas été conclusive.

Le Conseil est conscient des limites de cette approche, notamment dans la mesure où le profil sociologique des utilisateurs de Twitter n'est pas représentatif de l'ensemble de la population des internautes. Selon plusieurs chercheurs du MédiaLab de SciencesPo, « *Twitter constitue une importante caisse de résonance aux articles produits par les médias. Toutes les études montrent que le public dominant sur Twitter (à la différence de Facebook) est très spécifique : il rassemble principalement des individus au niveau social et culturel élevé, et, qui, pour ceux qui partagent des informations d'actualité, ont des convictions politiques fermes* »<sup>18</sup>. Par ailleurs, selon ces auteurs, qui ont conduit une analyse des liens partagés par les comptes de près de 400 médias, Twitter se caractérise par « *une porosité plus forte, notamment entre les médias de droite et d'extrême-droite ce qui peut aider certaines informations marginalisées à bénéficier de relais plus large que leur bassin naturel* ».

### 3.3. Description du service Twitter

Twitter est un **réseau social consistant en la publication de brefs messages** – les **Tweets** – ne dépassant pas 280 caractères. Le service se décrit comme « *la vitrine de ce qu'il se passe dans le monde et des sujets de conversation du moment* »<sup>19</sup>. Il est accessible gratuitement, sur inscription pour publier et sans inscription pour consulter les Tweets des comptes publics.

Ces Tweets peuvent contenir différents types de médias et des mots-dièses (« # » suivi d'un mot ou d'une phrase écrite sans espaces, comme #BlackLivesMatter). L'usage de ces **mots-dièses – hashtags** – permet de rassembler des Tweets sur un même sujet et fait l'objet d'une mise en avant sur le réseau social – les *trending topics* – lorsqu'ils concentrent par exemple de nombreux messages sur un temps réduit. Il est ainsi possible de suivre les « conversations » qui sont populaires sur le réseau social à un instant t, leur affichage pouvant être personnalisé pour chaque utilisateur (selon sa localisation par exemple).

L'ensemble des Tweets est rassemblé d'une part sur un **profil** lié à l'utilisateur qui les publie, qui peut être public ou privé, et d'autre part au sein d'un **fil d'actualités** séparé – affiché dès la connexion – qui se compose d'un agrégat de toutes les publications des profils auxquels un utilisateur donné est abonné (*timeline*). On dit dans ce cas que l'utilisateur suit – **follow** – un ou plusieurs utilisateurs et qu'il est un **abonné** (*follower*) de ces derniers. Chaque utilisateur inscrit sur Twitter dispose d'un profil rassemblant ses Tweets et d'un fil d'actualités. Ce fil d'actualités est proposé dans une organisation ante-chronologique des contenus depuis les débuts du service.

<sup>17</sup> Lefigaro.fr. Facebook a franchi la barre des «37 millions d'utilisateurs» en France en 2019. 10/02/2020.

<sup>18</sup> Dominique Cardon,, Benjamin Ooghe-Tabanou, Guillaume Plique, Jean-Philippe Cointet. SciencesPo Médialab. *Les nouveaux circuits de l'information numérique*. 2019.

<sup>19</sup> Twitter, A propos.



Depuis 2016, l'utilisateur peut également choisir une version alternative de ce fil d'actualités avec une **composante algorithmique**<sup>20</sup> : l'affichage des Tweets y sera alors personnalisé en fonction de plusieurs critères comme le nombre d'interactions de l'utilisateur sur les contenus. Twitter offre à tout moment la possibilité de passer d'un fil d'actualité avec apport algorithmique à un fil uniquement ante-chronologique, qui existe toujours et reste disponible pour chaque utilisateur.

Les utilisateurs ont en effet plusieurs possibilités d'interaction avec chaque Tweet, qu'il soit issu d'un compte auxquels ils sont abonnés ou bien d'un compte public qu'ils ne suivent pas. Ils peuvent ainsi le republier afin qu'il apparaisse sur leur profil et dans le fil d'actualité de leurs *abonnés* (**ReTweet**). Ils peuvent également le republier en le **commentant**, toujours dans la limite de 280 caractères (**QuoteTweet**), mettre ce Tweet en **favoris** (« aimer ») et le conserver sur son profil dans une section dédiée nommée « J'aime ») ou bien le **partager**, soit en copiant le lien du Tweet soit par message privé. Cette dernière fonctionnalité permet d'envoyer des messages à un autre utilisateur sans qu'ils ne soient accessibles aux autres ni publiés sur son profil ou dans les fils d'actualité de ses *abonnés*.

### 3.4. Les objectifs de l'étude

Cette étude a pour objectif de **mieux comprendre les mécanismes de propagation des fausses informations sur le réseau social Twitter**. Elle prend la forme de deux analyses reposant sur des méthodologies en grande partie distinctes. La première est basée sur une analyse de comptes Twitter labellisés selon leur fiabilité. La deuxième est, elle, basée sur l'analyse d'un certain nombre de sujets d'actualité ayant donné lieu à la diffusion de fausses informations.

Dans sa première partie, l'étude consiste donc à aborder la question de la désinformation par le biais d'une labellisation de comptes d'informations actifs sur Twitter. La labellisation adoptée par le Conseil pour cette analyse reprend une classification réalisée par le *Décodex* du Monde en fonction de la fiabilité des informations partagées par ces comptes. L'analyse a alors consisté à comparer les caractéristiques des différentes catégories de comptes (notamment les comptes d'informations fiables et ceux connus pour partager de fausses informations). Cette comparaison porte notamment sur le nombre d'abonnés des différents types de comptes et sur la propension de ces différents types de comptes à voir leurs Tweets repris sous la forme de ReTweets. Cette section présente également une analyse des thématiques abordées par les comptes connus pour propager de fausses informations. Elle comprend par ailleurs une analyse de la diversité des comptes d'information suivis par les utilisateurs de Twitter de façon à identifier la présence de phénomènes dits de « chambres d'échos ». Enfin, cette section présentera également une analyse de l'activité de certains comptes de *fact-checkers* présents sur le réseau social.

Dans une deuxième partie, l'étude présente les résultats d'une analyse fondée sur la labellisation de Tweets liés à certaines fausses informations identifiées auprès de *fact-checkers*. Cette section propose notamment une analyse quantitative de la chronologie de la diffusion des fausses informations et de leur réfutation et propose également une analyse qualitative du contenu des Tweets ayant présenté la viralité la plus importante.

---

<sup>20</sup> Slate. *Twitter's New Order. Inside the Changes That Could Save Twitter's Business – and Reshape Civil Discourse*. 05/03/2017.



L'étude se concentre sur les comptes d'information francophones et sur les Tweets publiés en français. Elle part donc du postulat que les utilisateurs francophones utiliseront principalement des sources d'information francophones<sup>21</sup>. Ses conclusions ne sont donc pas nécessairement transposables à l'univers non francophone.

### **3.5. La collecte de données sur Twitter**

La présente étude est basée sur une collecte de données effectuée sur Twitter par l'intermédiaire des API proposées gratuitement à l'ensemble du public par la plateforme. À la connaissance du Conseil, aucune plateforme similaire à Twitter n'offre d'API ouverte et gratuite permettant un accès aussi complet aux données de sa plateforme.

L'API proposée par Twitter permet à des développeurs externes de s'appuyer sur les données de la plateforme pour développer des programmes ou applications ou encore de mener des projets à des fins statistiques. À ce titre, les données mises à disposition par l'API concernent notamment les Tweets émis par un compte ainsi que l'ensemble des statistiques d'interactions propres à chaque Tweet (nombre de ReTweet, nombre de « j'aime »...) ou l'identité des comptes qui suivent ou sont suivis par un compte donné. L'API permet également la recherche de Tweets contenant un mot-clé donné. Les versions de l'API mises à disposition gratuitement par Twitter et utilisées par le Conseil présentent de nombreuses restrictions en matière de nombre de requêtes pouvant être adressées à l'API sur une durée donnée mais également en matière de profondeur de l'historique dans le cas des recherches de Tweets par mots clés.

Afin d'interroger l'API de Twitter, les services du Conseil ont eu recours aux « bibliothèques » Tweepy et Twython (pour les codes en langage Python) et à la « bibliothèque » rtweet (pour les codes en langage R). La mise en œuvre de requêtes établies à partir de ces bibliothèques a permis la collecte de données et la constitution de bases de données dont le contenu sera présenté de façon détaillée dans les sections suivantes.

### **3.6. Les précautions prises par le CSA afin de s'assurer du respect du règlement général sur la protection des données**

L'extraction et le traitement des données mentionnées *supra* ont été effectués dans le respect du règlement général sur la protection des données personnelles (RGPD) et de la loi du 6 janvier 1978 (loi « informatique et libertés »). Sont protégés par ces textes tout traitement, informatique ou non, qui, en dehors du cadre domestique, porte sur des « informations se rapportant à des personnes identifiées ou identifiables ». Si l'étude ne vise aucun examen ciblant des individus donnés, elle a cependant impliqué le traitement de certaines informations personnelles. D'une part, le traitement de données personnelles a été rendu nécessaire par le fonctionnement de l'API de Twitter. Par exemple, l'étude s'est intéressée à la diversité des comptes suivis par certains utilisateurs de Twitter. Cette analyse a supposé dans un premier temps de collecter les identifiants de ces individus, qui constituent des données personnelles. Ces identifiants ont

---

<sup>21</sup> Le Digital News Report 2019 du Reuters Institute for the Study of Journalism confirme par exemple que les seize sites d'information en ligne les plus consultés par les français sont tous francophones.



ensuite été utilisés pour adresser des requêtes sur l'API de Twitter permettant de collecter la liste des comptes suivis par ces utilisateurs. La collecte de la donnée d'identification est purement incidente et ne vise pas un traitement individuel. Ces données d'identification ont donc été supprimées à la fin de la collecte. D'autre part, le texte de certains Tweets a également été collecté. Les Tweets concernés par cette collecte sont ceux émis par des comptes d'informations et ceux contenant certains mots clés renvoyant à une fausse information donnée. La nature même des Tweets collectés diminue la probabilité qu'ils contiennent des informations personnelles, sans cependant que cette possibilité puisse être exclue.

Si l'étude visait donc une minimisation de la collecte et du traitement de données personnelles, de telles données pouvaient donc être collectées incidemment. Une communication a donc été publiée le 13 septembre 2019 sur le site du CSA de façon à informer le public de l'existence d'une étude sur le phénomène de propagation de l'information sur les réseaux sociaux<sup>22</sup>. En cohérence avec les principes du RGPD, cette information indiquait les finalités du traitement des données, la base légale du traitement, les catégories de données personnelles traitées, la source des données, les destinataires des données et leur durée de conservation. La présente étude se conforme aux finalités de traitement annoncées par le CSA lors de cette information qui indiquait que le Conseil cherchait à mieux comprendre les ressorts de la diffusion de l'information sur la plateforme Twitter.

Les données collectées par le CSA dans le cadre de cette étude ne seront par ailleurs pas réutilisées pour d'autres finalités. Le RGPD prévoit six bases légales distinctes pour la collecte et le traitement de données personnelles. Dans le cas présent, le CSA a mobilisé comme base légale l'exercice de l'autorité publique dont il est investi en application de la loi n° 86-1067 du 30 septembre 1986, notamment son article 17-2. Depuis l'entrée en vigueur de la loi n° 2018-1202 du 22 décembre 2018, le CSA est en effet tenu de contribuer à la lutte contre la diffusion de fausses informations et a notamment pour mission, en cas de nécessité, d'adresser aux opérateurs de plateformes en ligne des recommandations visant à améliorer la lutte contre la diffusion de telles informations. Afin d'assurer ces nouvelles missions, le CSA doit disposer d'une compréhension adéquate du phénomène de diffusion des fausses informations sur les plateformes en ligne, et notamment les réseaux sociaux comme Twitter. La présente étude contribue à l'expertise du Conseil sur ce sujet.

---

<sup>22</sup> Conseil supérieur de l'audiovisuel. *Données personnelles : le CSA mène une étude sur le phénomène de propagation des fausses informations sur les réseaux sociaux*. 13/09/2019.



## 4. Analyse de l'activité des comptes d'information fiables, non fiables et des comptes de *fact-checkers*

Cette première section présente une analyse des caractéristiques de différentes catégories de comptes Twitter : les comptes Twitter catégorisés comme fiables ou non fiables par l'initiative *Décodex* du *Monde* et les comptes de certains services journalistiques de vérification de l'information (*fact-checkers*) actifs sur la plateforme Twitter.

### 4.1. Présentation du *Décodex*

Afin de mener son analyse, le Conseil n'a procédé par lui-même à aucune catégorisation de comptes, mais a repris telle quelle celle du *Décodex* du *Monde*. Cette catégorisation de sources d'information en fonction de leur fiabilité est mise à disposition à l'adresse suivante : <https://www.lemonde.fr/web-service/decodex/updates>.

Cet annuaire des sources du *Décodex* a pour vocation, selon *Le Monde*<sup>23</sup>, d'aider les internautes à se repérer dans un contexte où de nombreuses fausses d'information circulent sur Internet et où les sources de ces fausses informations sont régulièrement les mêmes. La classification est basée sur l'analyse des sources du site, des auteurs, de leur identification, de la présence de correctifs en cas d'erreur, des renvois vers des sites douteux... L'outil se décline notamment en un moteur de recherche sur le site du *Monde*, où l'internaute peut entrer l'adresse d'un site à propos duquel il souhaite des informations, et d'une extension pour navigateurs permettant, pendant la navigation, de consulter des informations sur les sites consultés. La classification met en avant des sites « bleus » (catégorisés « 1 » dans la base mise en libre accès par *Le Monde*) pour les sites parodiques, des sites « rouges » (catégorisés « 2 » dans la base) qui diffusent régulièrement de fausses informations, des sites « oranges » (catégorisés « 3 » dans la base) pour les sites dont la fiabilité ou la démarche sont douteuses. Les sites réputés fiables n'ont pas de couleur spécifique et sont placés en catégorie « 4 » dans la base.

Comme il sera détaillé *infra*, la classification opérée par les journalistes du *Monde* a notamment été reprise par plusieurs chercheurs étudiant les phénomènes de désinformation. Par ailleurs, dans une étude de 2019, la cartographie de 14 000 liens hypertexte échangés entre 391 médias français entre avril et novembre 2018 a permis de faire apparaître de façon automatique une classification de sites « très proche de celle effectuée à la main par les Décodeurs du journal *Le Monde* pour désigner la fiabilité des sources d'information selon les critères de la profession »<sup>24</sup>.

<sup>23</sup> Le Monde. *L'annuaire des sources du Décodex : mode d'emploi*. 23/1/2017.

<sup>24</sup> Dominique Cardon,, Benjamin Ooghe-Tabanou, Guillaume Plique, Jean-Philippe Cointet. SciencesPo Médialab. *Les nouveaux circuits de l'information numérique*. 2019.



Aucune donnée concernant les caractéristiques sociodémographiques des abonnés aux comptes catégorisés par le *Décodex* n'a été collectée pour la présente étude. Des informations publiées par Le Monde<sup>25</sup> indiquent cependant que les sites de catégorie « rouge » et « orange » ne touchent pas seulement les plus jeunes et les personnes peu diplômées. La tranche d'âge qui se distinguerait par son appétence pour les sources douteuses est la tranche des 25 à 49 ans. Les catégories socioprofessionnelles supérieures sont par ailleurs plus nombreuses à consulter les sources moins fiables.

## 4.2. Analyse statistique des différentes catégories de comptes

L'analyse statistique des différentes catégories de comptes du *Décodex* est basée sur la collecte des 1000 derniers Tweets du fil d'actualité (*timeline*) de chaque compte (collecte effectuée le 13 septembre 2019). Les données relatives à ces Tweets (nombre de ReTweets<sup>26</sup>, nombre de favoris, etc.) ont également été collectées. Pour chacun de ces comptes ont également été collectées les données relatives au nombre d'abonnés, au nombre de comptes suivis et au nombre total de Tweets publiés depuis la création du compte. Le tableau suivant présente les caractéristiques des comptes analysés à la date de la collecte. Ces statistiques excluent les Tweets qui n'ont pas été émis par le compte d'information étudié mais qui sont des ReTweets de Tweets provenant d'un autre compte<sup>27</sup>.

CATÉGORIE DÉCODEX	NOMBRE MOYEN D'ABONNÉS	NOMBRE MOYEN DE RETWEETS PAR TWEET	NOMBRE MOYEN DE FAVORIS	NOMBRE DE COMPTES
1 (parodique)	152 004	47	173	12
2 (diffusant des fausses informations)	14 239	18	17	29
3 (douteux)	85 335	20	22	21
4 (a priori fiables)	516 962	14	29	157

<sup>25</sup> Le Monde. „La désinformation ne touche pas seulement les jeunes et les personnes peu diplômées. 05/08/2019.

<sup>26</sup> Les retweets prenant la forme de citations (« quotes ») sont également pris en compte dans cette analyse. Ces tweets prenant la forme de citations sont cependant parfois utilisés, non pas pour participer à la propagation du tweet originel, mais plutôt pour le critiquer. L'analyse effectuée dans cette section pourrait donc conduire à surestimer la viralité des différents types de comptes (et notamment des comptes propageant de fausses informations).

<sup>27</sup> Les données diffusées par l'API de Twitter sur le nombre de retweets ne permettent en effet pas de distinguer si les retweets ont été effectués à partir du compte d'information étudié, intégrer les retweets de Tweets qui ne sont pas émis par le compte d'information conduirait donc à fausser l'analyse de la viralité de ces comptes (si un compte d'information B retweete un Tweet émis par un compte A, les statistiques attachés au Tweet du compte B comprennent l'ensemble des retweets effectués à partir du compte A ou d'autres comptes tiers).



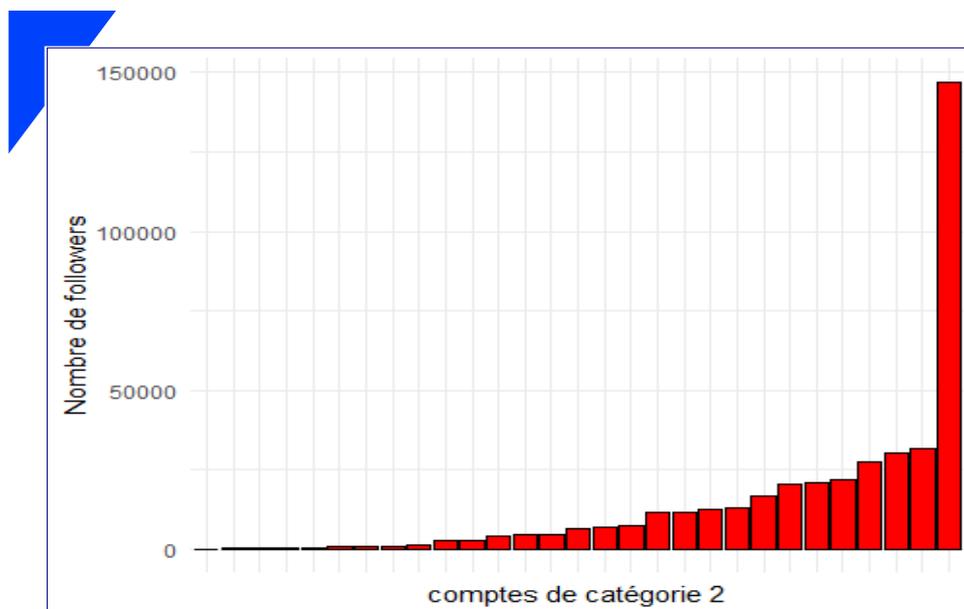
En raison de la forte dispersion des indicateurs entre les différents comptes, cette analyse des moyennes peut utilement être complétée par une analyse des médianes<sup>28</sup>:

CATÉGORIE DÉCODEX	NOMBRE MEDIAN D'ABONNÉS	NOMBRE MEDIAN DE RETWEETS PAR TWEET	NOMBRE MÉDIAN DE FAVORIS	NOMBRE DE COMPTES
1	18 968	10	17	12
2	6 409	2	2	29
3	21 596	8	5	21
4	73 945	4	8	157

Une analyse plus détaillée du nombre d'abonnés et du nombre de ReTweets par Tweet pour les différentes catégories de comptes est présentée dans les sections suivantes.

#### 4.2.1. Les comptes fiables possèdent une base d'abonnés significativement plus élevée que les autres types de comptes

Que l'analyse porte sur les médianes ou les moyennes, il apparaît que les comptes les moins fiables (de catégorie 2 ou 3) possèdent un nombre d'abonnés et donc une audience directe très significativement inférieurs à ceux des comptes de catégorie 4 considérés comme fiables. Cette analyse des médianes et des moyennes cache cependant une dispersion importante du nombre d'abonnés. Ainsi, les comptes de catégorie 2 possèdent de 285 à près de 150 000 abonnés. On observe par ailleurs un écart très élevé entre le compte de catégorie 2 le plus suivi et les comptes suivants :



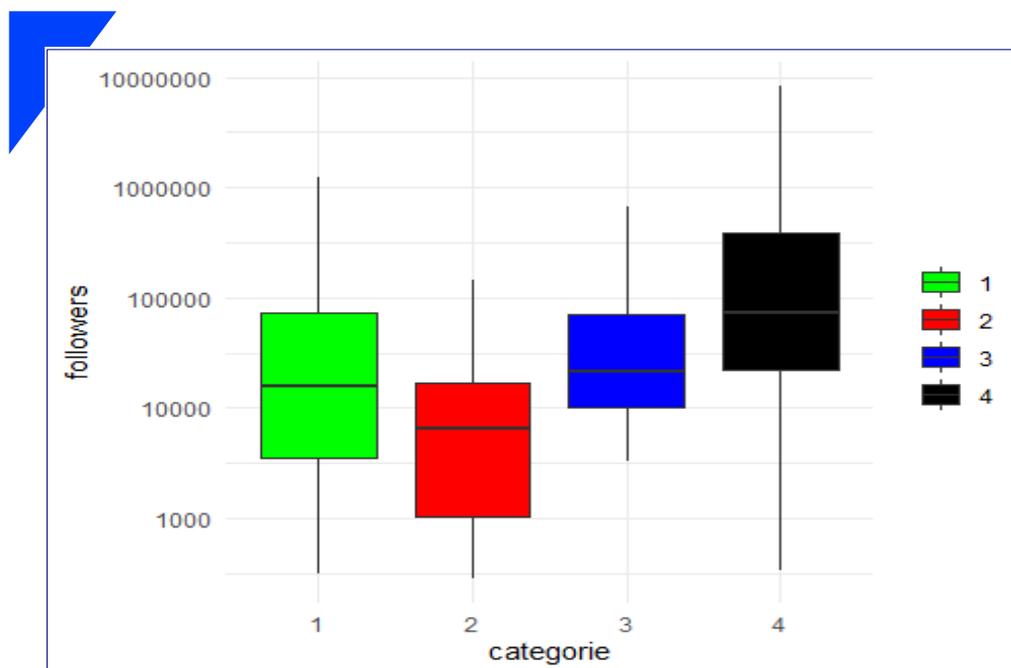
<sup>28</sup> Pour les statistiques relatives au nombre médian de retweets et au nombre médian de favoris, il s'agit plus précisément de calculer la moyenne de retweets par Tweets ou de favoris pour chaque compte et de calculer ensuite la médiane de ces moyennes.



Les comptes classés en catégorie 4 présentent une dispersion encore plus marquée puisqu'ils recueillent de 334 à plus de 8 millions d'abonnés. Un nombre significatif de comptes de catégories 4 présentent un nombre d'abonnés très significativement plus élevé que le compte de catégorie 2 le plus important. Il existe ainsi 59 comptes qualifiés de fiables qui disposent de plus de 150 000 abonnés et 26 comptes disposant de plus de 1 000 000 d'abonnés.

Les comptes de catégorie 3 présentent une situation intermédiaire. Ils obtiennent pour certains un nombre d'abonnés substantiellement plus élevé que les comptes de catégorie 2 mais sans rivaliser avec l'audience des plus importants comptes de catégorie 4.

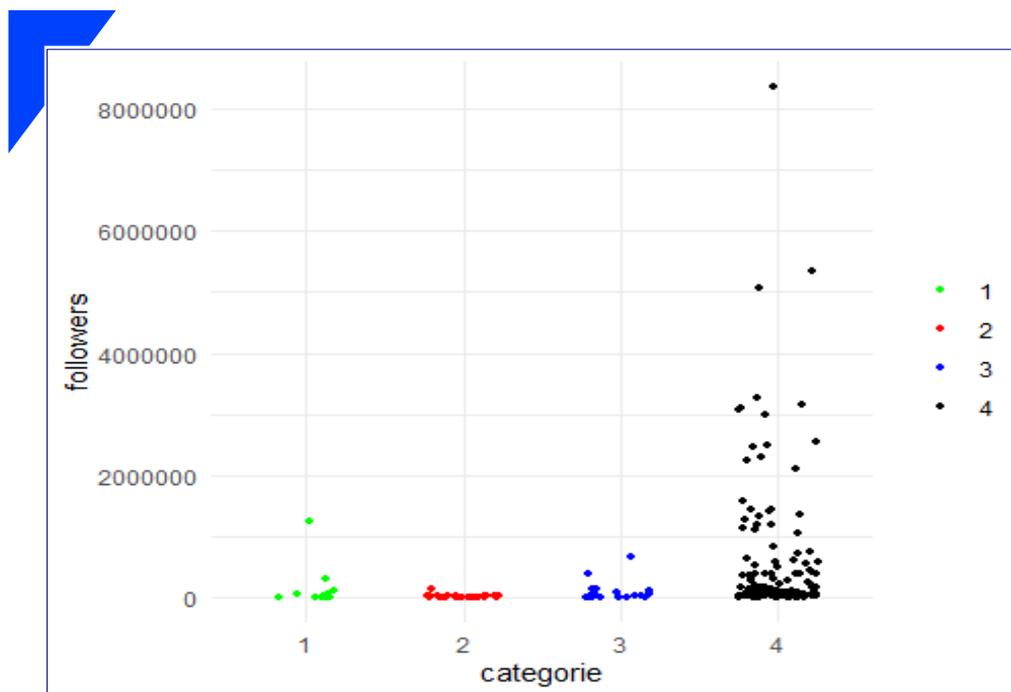
Le graphique suivant permet de représenter la dispersion du nombre d'abonnés pour chacune des catégories de compte. Notons que le graphique utilise une échelle logarithmique qui écrase les différences.



*Lecture : pour chaque catégorie, 25 % des comptes se situent sous le bord inférieur du rectangle et 25 % au-dessus du bord supérieur. La ligne horizontale située dans le rectangle représente la médiane. Les lignes s'étendent par ailleurs jusqu'aux valeurs minimales et maximales mais excluant les points hors normes<sup>29</sup>.*

Le graphique suivant présente la même dispersion sans recourir à une échelle logarithmique et en représentant l'ensemble des comptes. Cette représentation permet également d'illustrer la disproportion entre l'audience des comptes de catégorie 2 et celle des comptes de catégorie 4.

<sup>29</sup> Les points sont considérés comme étant hors normes s'ils se situent à une distance du bord du rectangle qui excède 1,5 fois la hauteur du rectangle.



Les résultats de cette section présentent des similarités avec les résultats mis en avant dans de précédentes études. Ainsi l'analyse de Fletcher et al. (2018)<sup>30</sup> des données d'audience des sites web recensés par le *Décodex* a montré que l'audience des sites internet catégorisés comme partageant de fausses informations est faible et que ces sites ont été consultés par moins de 1 % des internautes français. Ils ont notamment été très significativement moins consultés que ceux des principaux journaux d'information français (*Le Monde* et *Le Figaro*) ou que celui de France Info.

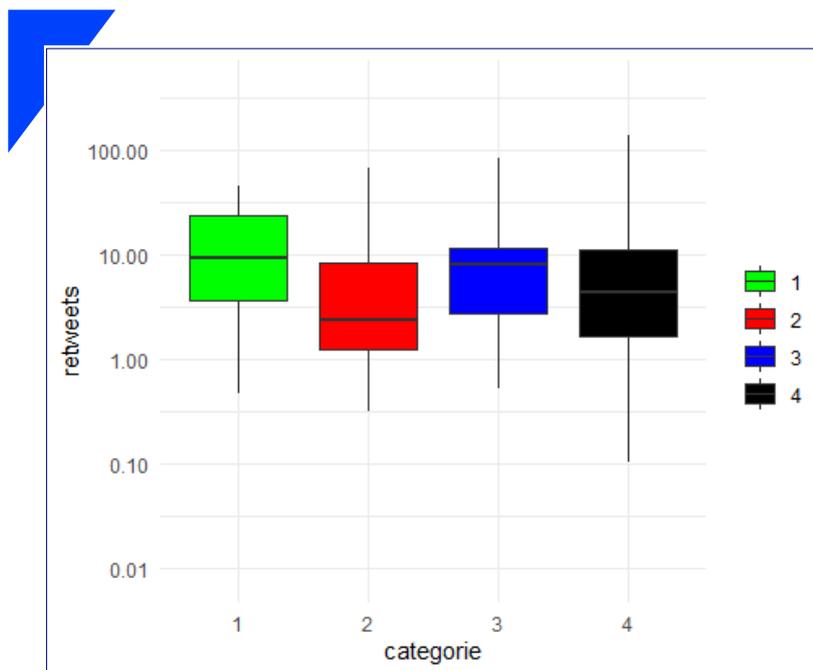
#### 4.2.2. Les comptes peu fiables présentent des nombres de ReTweets significatifs, en particulier par rapport à leur nombre d'abonnés

L'analyse du nombre d'abonnés d'un compte ne fournit qu'une vision partielle de son audience et de sa visibilité sur Twitter. Lorsqu'un utilisateur suit un compte, les Tweets publiés par ce compte sont certes affichés dans son fil d'actualités. Toutefois, en raison des modifications récentes de l'algorithme de Twitter (voir description *supra*), la visibilité du Tweet dans la timeline n'est pas connue. En outre, l'attention prêtée par l'utilisateur à un Tweet apparaissant dans sa timeline est également incertaine. Enfin et surtout, le nombre d'abonnés d'un compte ne donne qu'une vision partielle de la visibilité des Tweets publiés par ce compte. Les Tweets peuvent en effet être visibles au-delà des abonnés du compte ayant publié le Tweet s'ils font l'objet de ReTweets. Dans ce cas, les abonnés des comptes ayant procédé au ReTweet verront également apparaître le Tweet dans leur timeline. Autrement dit, l'audience d'un compte n'est pas seulement impactée par son nombre d'abonnés mais également par sa capacité à être repris sous forme de ReTweets par des comptes tiers.

<sup>30</sup> Fletcher Richard, Cornia Alessio, Graves Lucas, Nielsen Rasmus Kleis. *Measuring the reach of "fake news" and online disinformation in Europe*. 2018.



À partir des tableaux présentés *supra*, il apparaît que les Tweets émis par les comptes de catégorie 4 font l'objet de moins de ReTweets que les comptes de catégorie 1, 2 et 3. Si l'analyse est en revanche effectuée en termes de médiane, les comptes de catégorie 4 font toujours l'objet de moins de ReTweets que les comptes de catégorie 1 et 3 mais dépassent cette fois les comptes de catégorie 2. L'analyse de la dispersion est encore une fois intéressante et il apparaît que certains comptes de catégories 2 font jeu égal avec les comptes de catégorie 4 :



Ce résultat est également cohérent avec les résultats mis en avant par des études antérieures. Ainsi, Fletcher et al. (2018) ont montré que certains comptes connus pour partager de fausses informations sur Facebook avaient pu générer des engagements<sup>31</sup> similaires, voire supérieurs, à ceux de certains des plus grands médias catégorisés comme fiables. Le Monde (2018)<sup>32</sup> a également analysé les engagements sur les comptes Facebook indexés par *Décodex* : il en ressort que les comptes fiables totalisent 72,3 % des engagements sur Facebook. Le Monde relève alors que si les comptes de désinformation n'ont jamais supplanté les comptes fiables, on constate que certains comptes de désinformation rivalisent avec des rédactions bien installées. Les deux comptes de désinformation les plus populaires ont ainsi suscité plus d'engagements que Le Point ou Libération.

Les comptes de catégorie 2 disposent donc d'un nombre d'abonnés faible par rapport aux comptes de catégorie 4 mais font jeu égal avec ces derniers en termes de nombre de ReTweets. Ils génèrent une viralité particulièrement élevée par rapport à leur nombre d'abonnés. Afin d'étudier de façon plus détaillée cette propriété, le tableau suivant présente le nombre de ReTweets pour 1 000 abonnés pour les différentes catégories de comptes :

<sup>31</sup> Sur Facebook, l'engagement est calculé comme la somme du nombre de commentaires, de partages et de réactions sur un post.

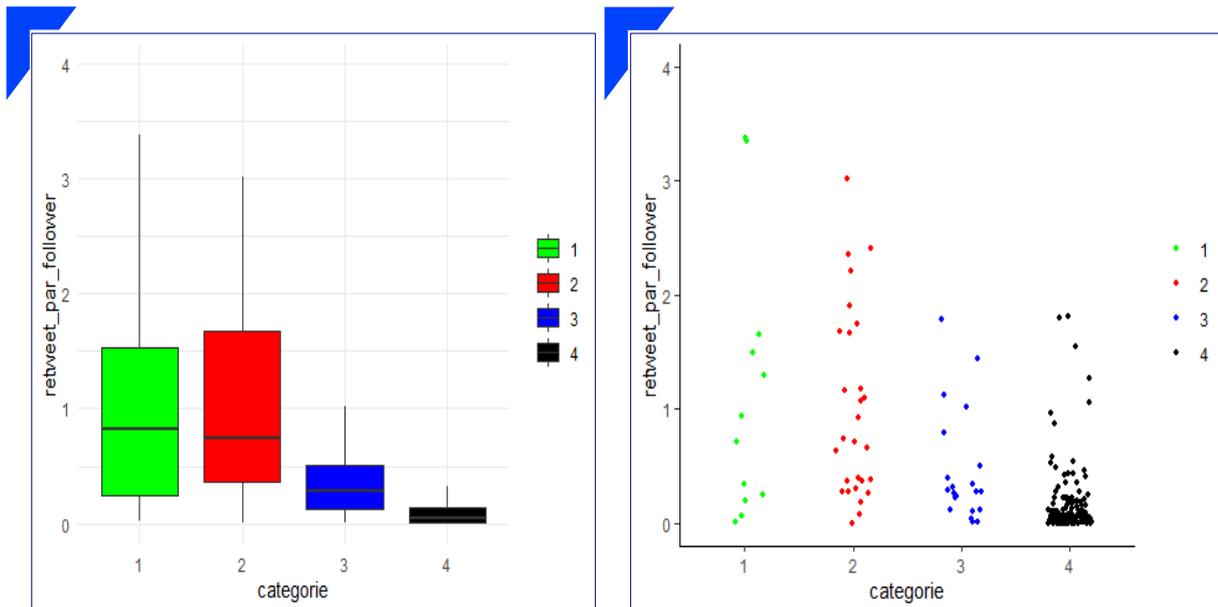
<sup>32</sup> Le Monde. *Les fausses informations circulent de moins en moins sur Facebook.*



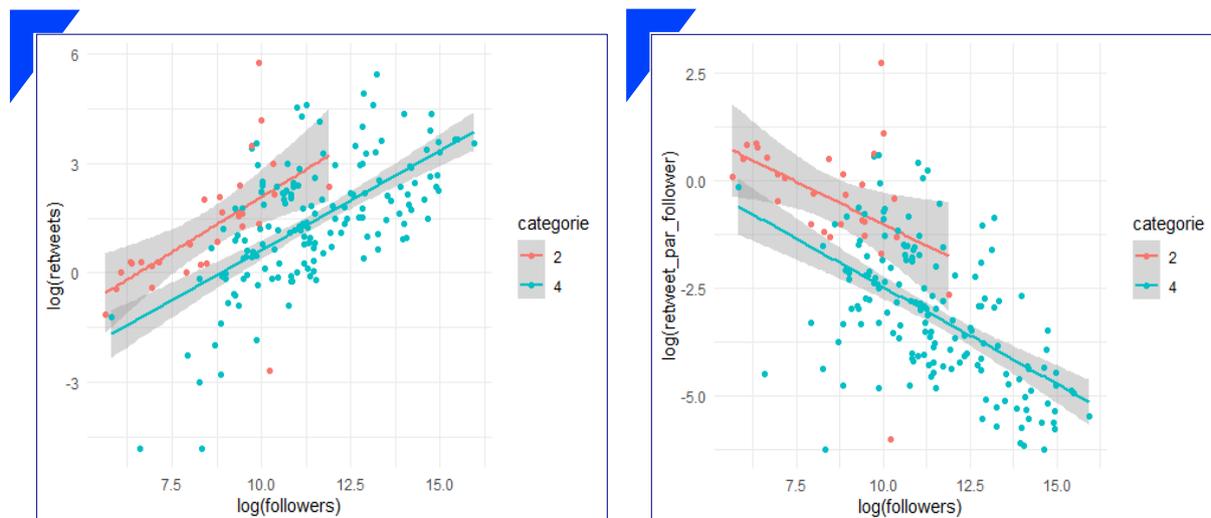
CATÉGORIE DÉCODEX	NOMBRE MOYEN DE RETWEET POUR 1000 ABONNÉS	NOMBRE MÉDIAN DE RETWEET POUR 1000 ABONNÉS
1	1,14	0,82
2	1,51	0,74
3	0,46	0,28
4	0,15	0,04

Lecture: Pour un compte de catégorie 2, chaque millier d'abonnés génère en moyenne 1,51 Tweets.

Le nombre moyen ou médian de ReTweets par Tweet et par abonné est ainsi beaucoup plus élevé pour les comptes de catégorie 1, 2 et 3 que pour les comptes de catégorie 4 considérés comme fiables. Les abonnés de comptes connus pour partager de fausses informations ont donc une propension beaucoup plus élevée à contribuer à la diffusion des Tweets de ces comptes. Les analyses de dispersion confirment ces résultats :



Le graphique ci-dessous représente par ailleurs la relation entre le nombre moyen de ReTweets pour un compte et le nombre d'abonnés du compte. Cette analyse permet de montrer que le nombre de ReTweets est une fonction croissante du nombre d'abonnés et que la croissance du nombre de ReTweets avec le nombre d'abonnés semble identique pour les comptes de catégorie 2 et de catégorie 4. Cependant, les comptes de catégorie 2 bénéficient, à nombre d'abonnés identique, d'un nombre plus important de ReTweets que les comptes de catégorie 4. Il est également intéressant d'observer que le nombre de ReTweets par abonné est une fonction décroissante du nombre d'abonnés :



Lecture : le graphique de gauche montre qu'un nombre croissant d'abonnés s'accompagne d'un nombre croissant de ReTweets, aussi bien pour les comptes de catégories 2 que de catégorie 4. À nombre d'abonnés identique, les comptes de catégorie 2 disposent cependant de plus de ReTweets que les comptes de catégorie 4. Le graphique de droite montre que le nombre de ReTweets par abonné diminue quand le nombre d'abonnés augmente. À nombre d'abonnés identiques, le nombre de ReTweets par abonné des comptes de catégorie 2 est plus élevé que celui des comptes de catégorie 4. Les axes des deux graphiques sont par ailleurs exprimés en logarithme pour améliorer leur lisibilité ce qui ne modifie cependant pas leur lecture.

Ce constat pourrait s'expliquer par le fait qu'un compte qui est suivi par un nombre élevé d'abonnés va rassembler une audience généraliste dont la proportion d'individus impliqués personnellement dans les thématiques partagées par le compte est relativement plus faible. La population totale de ce compte sera donc moins engagée et aura donc une propension moindre à retweeter les contenus diffusés.

#### Synthèse :

Les comptes d'information les moins fiables possèdent un nombre d'abonnés significativement plus faible que la majorité des comptes d'information fiables. Cependant, ils font jeu égal, en termes de ReTweets, avec les comptes fiables. Les abonnés des comptes non fiables ont en effet une propension beaucoup plus élevée à contribuer à la diffusion des informations présentes sur ces comptes que les abonnés des comptes fiables.



### 4.3. Analyse du contenu des Tweets des comptes connus comme étant propices à la diffusion de fausses informations

La section précédente a permis de mettre en évidence la particularité des comptes propageant des fausses informations en matière de viralité. Afin de mieux comprendre cette situation et de mieux appréhender la nature des informations partagées par ces comptes, une analyse du contenu des Tweets les plus retweetés de ces comptes (parmi les 1 000 derniers Tweets collectés<sup>33</sup>) a été menée en étudiant leur mise en forme, le message, la présence de vidéos et photos, la présence de fausses informations<sup>34</sup>, etc. L'analyse qualitative du contenu des Tweets sera complétée par une analyse quantitative.

#### 4.3.1. Analyse qualitative du contenu des Tweets des comptes de catégorie 2

Parmi les 1 000 derniers Tweets publiés par chacun des 29 comptes de catégorie 2 du *Décodex*, une collecte systématique des 10 Tweets les plus retweetés a été réalisée. 282 Tweets ont ainsi été récoltés en suivant cette méthodologie<sup>35</sup>.

L'étude de ces Tweets s'est attachée à analyser les ressorts suscitant l'engagement des utilisateurs, ici entendu comme leur propension à partager un contenu (*ReTweet*). Pour ce faire, dans un premier temps, l'ensemble des 282 Tweets récoltés ont été classés selon les thématiques abordées<sup>36</sup>. Dans un second temps, une analyse qualitative de leur contenu a été menée afin de dégager des traits communs en termes de traitement de l'information.

##### 4.3.1.1. Analyse des thématiques abordées

Les 282 Tweets les plus retweetés pour chacun des 29 comptes étudiés ont été classés selon la thématique principale qu'ils abordent avec une attention particulière portée à la mention de fausses informations ou d'informations trompeuses<sup>37</sup>. La collecte de données pour cette seule partie de l'étude a précédé l'apparition du nouveau coronavirus et la crise sanitaire qu'il a déclenchée<sup>38</sup>.

---

<sup>33</sup> La collecte a été effectuée le 13 septembre 2019. L'analyse porte donc sur les 1 000 derniers Tweets publiés par chaque compte à cette date.

<sup>34</sup> Afin de mener son analyse, le Conseil n'a procédé par lui-même à aucune catégorisation de comptes, mais a repris telle quelle la catégorisation du *Décodex* du *Monde* (voir section 3.2.). Cette catégorisation de sources d'information en fonction de leur fiabilité est mise à disposition par le monde à l'adresse suivante : <https://www.lemonde.fr/web-service/decodex/updates>

<sup>35</sup> Certains comptes ayant une audience limitée, ces 10 Tweets retweetés n'ont parfois pas pu être collectés pour chaque compte. Ainsi, pour l'un d'entre eux, seuls 2 Tweets ont été repartagés et les 8 autres collectés n'ont pas suscité de retweets.

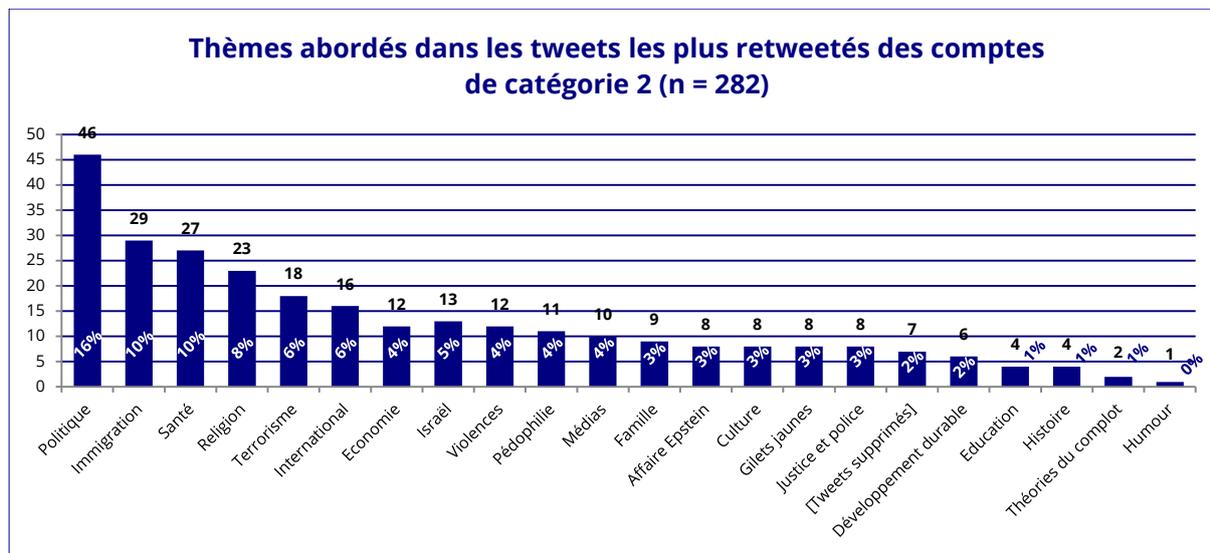
<sup>36</sup> Chaque Tweet n'a été compté qu'une seule fois, selon la thématique principale qu'il abordait. Cette méthode peut présenter des limites en ce que certains Tweets peuvent aborder plusieurs thématiques. Les tendances et proportions mises en avant restent cependant valables compte tenu des écarts importants constatés entre les thèmes.

<sup>37</sup> Cette qualification n'est pas le fait de cette étude mais d'une veille réalisée sur les thèmes couverts par les journalistes spécialistes de la vérification de l'information.

<sup>38</sup> Les analyses de la partie suivante comporteront cependant des éléments relatifs aux fausses informations concernant l'épidémie de coronavirus.



Leur contenu mêle parfois plusieurs thèmes : une publication aborde ainsi les potentielles affections de santé dont souffriraient des personnes en situation irrégulière. Elle a été classée comme relevant de la thématique « Immigration », le sujet « Santé » représentant un type de traitement de l'information. Par ailleurs, 7 Tweets collectés étaient devenus inaccessibles au moment de l'étude<sup>39</sup>.



L'analyse des thèmes se concentrera sur **ceux qui regroupent le plus grand nombre de Tweets et qui rassemblent 50 % ou plus des Tweets étudiés. Il s'agit de « Politique » (46 Tweets), « Immigration » (29), « Santé » (27), « Religion » (24) et « Terrorisme » (18), qui totalisent 51,1 % du corpus.**

- **Politique (46 Tweets)**

Parmi les 46 Tweets classés dans cette catégorie, 34 évoquent nommément des partis ou des personnalités politiques. Le Président de la République Emmanuel Macron est le plus cité, avec 9 occurrences. Il est également le sujet du Tweet le plus retweeté (2 919 ReTweets), une vidéo issue d'une émission de LCI montrant un chroniqueur qui le critique. Plusieurs personnalités proches du chef de l'État font également l'objet de publications de la part de ces comptes, comme le délégué national de La République En Marche ! ou une ancienne ministre et députée européenne, à chaque fois pour critiquer leur action ou partager des informations relatives à des polémiques les concernant.

Des personnalités politiques classées à gauche comme à droite de l'échiquier politique ressortent également de ce corpus de Tweets, de même qu'une militante écologiste et des structures liées à la mouvance identitaire. Les publications les concernant sont dans certains cas critiques, ou relèvent dans d'autres cas de démarches plus informatives, faisant la publicité d'un évènement organisé par ces partis (3) ou d'une entrée en campagne pour les élections municipales (1).

<sup>39</sup> Cette inaccessibilité peut être le résultat de plusieurs facteurs : l'auteur(e) du compte a pu le supprimer, rendant la consultation de son compte impossible, soit de son fait soit sur demande de Twitter après par exemple des signalements d'autres utilisateurs. Enfin, lorsqu'un Tweet est une réponse à un autre Tweet, si celui-ci est supprimé pour l'une des raisons évoquées, celui-ci disparaît également.



Enfin, 6 publications évoquent le monde politique dans son ensemble, en y attachant des thématiques autour de la corruption ou de l'existence d'une dictature en France.

- **Immigration (29 Tweets)**

Les 29 Tweets abordant le sujet de l'immigration évoquent des actes commis par des personnes présentées comme étrangères (11) ou la mise en place de politiques anti-migratoires hors de l'Hexagone (9). L'un de ces Tweets semble par ailleurs faire un lien entre risque sanitaire et immigration clandestine. L'utilisation du champ sémantique lié aux personnes en situation irrégulière est notable : les trois termes « clandestins », « clandos » ou « passeur » sont utilisés dans autant de publications.

Au sein de cette catégorie, les deux Tweets les plus partagés sont deux vidéos : la première montre des supporters présentés comme étant des Algériens qui « *entour[ent] la voiture de deux jeunes femmes françaises* » et qui voleraient leurs affaires tout en soulignant le silence du gouvernement à ce sujet (2 805 ReTweets). La seconde est celle d'un échange entre Éric Zemmour et Éric Dupond-Moretti durant lequel les deux s'opposent sur l'immigration (2 244 ReTweets). La troisième publication la plus retweetée ne comptabilise, elle, que 162 ReTweets.

- **Santé (27 Tweets)**

L'un de ces Tweets répond en miroir à celui mêlant immigration et risque sanitaire en faisant cette fois le lien entre « *immigration exotique* » et développement de pathologies réputées disparues en France.

Le plus populaire est toutefois une publication abordant « *l'emprise de l'empire vaccinal sur nos enfants* », qui a été retweetée 307 fois. De fait, la majeure partie de ces Tweets aborde la problématique des vaccins (14) et partage un scepticisme centré sur la commercialisation de vaccins contenant de l'aluminium.

On relève également que 2 Tweets font le lien entre alimentation et santé, accusant des produits notamment américains (isoglucose, Coca-Cola) d'être responsables de certaines affections. Le compteur Linky, qui a fait l'objet de nombreuses communications de la part d'associations de personnes électro-sensibles ces derniers mois, ne se retrouve que dans un seul Tweet.



- **Religion (23 Tweets)**

Cette thématique rassemble des publications qui abordent des faits d'actualité sous l'angle religieux ou bien utilisent la religion comme facteur d'explication de certaines problématiques. On note que c'est surtout l'Islam qui est cité (12 Tweets), suivi du judaïsme (11) et du catholicisme (7). Par ailleurs, 8 Tweets font mention conjointement de deux ou trois de ces religions.

Deux publications liées à l'Islam lient cette religion aux destructions de lieux de cultes catholiques, d'origine criminelle ou non, quand deux autres s'attachent à démontrer l'existence d'une incompatibilité supposée entre l'Islam et les sociétés européennes.

Parmi les publications faisant mention de la religion juive ou des personnes d'origine juive (11 Tweets), le Conseil représentatif des institutions juives de France (CRIF) est mentionné (et critiqué) 5 fois tandis qu'1 Tweet partage un lien sur l'Holocauste. Enfin, 3 Tweets mentionnent les mots « *synagogue* », « *kabbaliste* », « *juif-sioniste* » ou encore l'« *étoile tricolore de David* » qu'aurait porté le Président de la République au dîner du CRIF aux fins de montrer, d'après leurs auteurs, une forme de bienveillance des autorités vis-à-vis du judaïsme au détriment du catholicisme.

- **Terrorisme (18 Tweets)**

Cette thématique se concentre sur le terrorisme d'origine islamiste, notamment celui lié à Daech. Plusieurs d'entre eux critiquent une forme d'influence des terroristes sur les décisions politiques : ainsi, un ancien Président de la République aurait été vu « *pos[ant] avec un terroriste de Daesh* ». 5 Tweets vont plus loin et visent à démontrer une implication des autorités publiques dans le développement du terrorisme islamiste. Par exemple, le dirigeant de l'organisation terroriste Daech serait un « *agent de la CIA* ».

Enfin, 3 Tweets identiques publiés par le même compte abordent la thématique du terrorisme sous l'angle de la politique de solidarité française. Ils font état d'une information selon laquelle un djihadiste serait revenu de Syrie en demandant « *le versement de prestations sociales* » qu'il n'aurait pas perçues pendant son séjour hors de France.

#### 4.3.2. Analyse des traitements de l'information

Pour cette analyse, ont été extraits pour chaque compte la publication la plus partagée si elle totalise moins de 100 partages<sup>40</sup>, et les deux publications les plus partagées dès lors qu'elles comptabilisent 100 partages ou plus chacune (n = 37<sup>41</sup>). Les 37 Tweets étudiés ont fait l'objet à la fois d'une analyse sémantique et visuelle, prenant en compte aussi bien le message du Tweet que les différents contenus non-textuels (liens, images, vidéos). Il en résulte que tous les Tweets

---

<sup>40</sup> L'objectif est ici de présenter l'analyse des Tweets parvenant à obtenir les nombres de partages les plus élevés de façon à comprendre les ressorts de la viralité. Certains comptes présentent une faible popularité et sont donc peu retweetés, il reste cependant intéressant de présenter les Tweets les plus partagés de ces comptes afin de conserver une diversité des comptes étudiés. L'échantillon de Tweets retenu donne cependant une part plus importante aux comptes les plus partagés en retenant pour eux deux Tweets contre un pour les comptes les moins partagés (ceux qui n'ont jamais connu de publication partagée plus de cent fois).

<sup>41</sup> 42 Tweets répondaient à l'origine à cette qualification. L'un a été supprimé car il était, au sein d'un même compte, retweeté un nombre égal de fois à un autre, tous deux étant les plus retweetés, tout en restant sur un nombre de retweets inférieur à 100. C'est donc le plus récent des deux qui a été retenu. Quatre autres Tweets de plus n'ont pas été pris en compte dans ce corpus car devenus inactifs depuis leur collecte.



collectés comportent du texte avec un autre type de contenu (liens, images, vidéos, etc.) à l'exception d'un seul qui ne comporte que du texte. Ces Tweets collectés peuvent être analysés comme poursuivant **4 types d'objectifs : informer (15 Tweets), susciter le doute (5), la critique (11) ou encore l'adhésion (6) selon une gradation croissante, de la neutralité à une forme d'émotion**, comme la crainte ou la colère<sup>42</sup>.

Les 15 publications qui cherchent à informer l'utilisateur s'appuient sur plusieurs ressorts. En premier lieu, elles adoptent un ton neutre, très factuel. Dans certains cas, des verbatim sont repris dans des Tweets en identifiant nommément leurs auteurs. Ceux-ci convoquent également des références à des personnes dont les fonctions ou professions semblent être utilisées comme une source de légitimité. Ainsi en est-il du « *président de la république serbe de Bosnie* », du « *chancelier autrichien Sebastian Kurz* » ou encore d'un « *dentiste* » dont les actions ou propos sont relayés. On note que deux de ces publications utilisent des contenus liés à des médias ou agences de presse (France 3 et Associated Press). Par ailleurs, la qualité de certaines vidéos ou certains clichés utilisés laisse penser que, s'ils n'ont pas été pris par des professionnels, ils ne tiennent pas pour autant de l'amateurisme et font montre d'une recherche certaine de qualité ou bien d'un choix conscient des supports utilisés.

6 Tweets peuvent être considérés comme ayant pour visée de faire s'interroger l'utilisateur. Ceux-ci se rapprochent, dans leur fond comme dans leur forme, de théories du complot. Ils utilisent en effet des formes interrogatives et établissent des liens entre des sujets sans apporter une autre forme de preuve qu'un contenu dont l'origine n'est pas directement identifiable. Le ressort le plus utilisé ici est l'énumération de faits lesquels, pris ensemble, justifieraient la remise en question d'une information officielle ou d'une affirmation. Les contenus non textuels venant corroborer ces propos sont issus de sites liés au compte Twitter initial (4), d'un autre média sur lequel le Décodex du *Monde* appelle à la vigilance eu égard à la véracité des informations qu'il partage, ou encore d'une source non identifiable (photographie tronquée d'un écran).

Les 10 Tweets dont l'objectif est la critique adoptent un vocabulaire et un ton relevant de la prise de position. Ils recourent à des expressions telles que « *comme par magie* », « *arrogant* » et « *pour une fois que [...] a servi à quelque chose* ». Contrairement aux Tweets visant à informer, l'utilisation de sources clairement identifiées de médias d'information se fait pour mieux soutenir ces publications : ainsi en est-il de cette vidéo de franceinfo : dans laquelle les propos tenus par le président russe Vladimir Poutine à une conférence de presse conjointe avec le président français Emmanuel Macron sont présentés comme ayant été « *édulcoré[s]* » par une traductrice partisane (« *une stagiaire avec une carte d'adhérent LREM* »).

Enfin, les 6 Tweets cherchant à susciter l'adhésion utilisent un vocabulaire et des expressions relevant d'un registre de langage très faible (« *#fautrigoler* », « *ces ordures d'islamistes* » ou « *préparons résistance* »). Certains utilisent des caricatures ou des liens vers des pétitions.

La quasi-totalité des publications étudiées tendent ainsi vers la **construction d'une légitimité** des informations partagées. Ainsi, 15 Tweets convoquent un contenu qui leur est externe ou à

<sup>42</sup> Des signes tels que la ponctuation (« ??? », « !!! »), l'utilisation de mots en majuscules, les insultes, les verbes dépréciatifs ou encore les constructions logiques de phrase sont parmi ceux qui ont été utilisés dans cette analyse afin d'élaborer cette catégorisation. Celle-ci n'entend cependant pas faire loi en matière d'analyse des intentions des auteurs de ces contenus ni de véracité de leurs propos mais tend à apporter un éclairage sur les objectifs manifestes de publication sur leurs audiences car, comme déjà cité, « *la plupart du temps, le manipulateur ne se positionne pas par rapport à la vérité : il cherche simplement à produire un effet* » (J.-B. Jeangène Vilmer et al, août).



tout le moins difficilement identifiable comme étant propriétaire (**validation par une source autre**) tandis que 22 autres Tweets font le lien avec un contenu qui peut leur être rattaché directement ou indirectement, comme un site lié au compte Twitter. Dans ce dernier cas, ces Tweets agissent comme des **relais d'information** visant à renvoyer vers la consultation de ces contenus produits en interne qui se présentent majoritairement comme des articles. Les moyens diffèrent dans les deux cas, mais l'intention semble bien être ici d'ajouter des éléments à l'appui d'une information ou d'une prise de position dans le but de lui donner plus de consistance, quel que soit l'objectif final recherché : informer, susciter le doute, critiquer ou créer une forme de cohésion autour des mêmes idées.

Dans le cas des 15 Tweets faisant appel à du contenu externe, 5 partagent l'URL de sites sans lien, de prime abord, avec le compte Twitter d'origine, mais qui sont tous recensés par le Décodex comme étant des sites pour lesquels la vigilance est de mise. 1 autre partage par ailleurs du contenu auparavant publié sur Twitter par un compte considéré par le *Décodex* soit comme « gris », soit non classifié mais pour lequel un message d'avertissement est néanmoins proposé<sup>43</sup>.

#### 4.3.3. Analyse quantitative du contenu des Tweets des comptes de catégorie 2

L'analyse quantitative du contenu des Tweets des comptes de catégorie 2 et 4 a été menée sur les 25 % de Tweets les plus retweetés pour chaque compte (parmi les 1000 derniers Tweets publiés), encore une fois afin de dégager des enseignements sur les thèmes susceptibles de générer une viralité importante. L'analyse a tout d'abord consisté à supprimer les termes non significatifs (mots les plus courants de la langue française ou non alphabétiques). Les mots similaires (mais présentant une orthographe différente) ont par ailleurs été regroupés<sup>44</sup>. L'analyse a alors visé à comparer la fréquence d'apparition des mots dans les deux corpus et à identifier les termes surreprésentés dans les comptes de catégorie 2 par rapport aux comptes de catégorie 4<sup>45</sup>.

Il existe deux méthodes pour apprécier l'ampleur de la surreprésentation. La première consiste à calculer la différence entre l'occurrence effective du terme dans le corpus des comptes de catégorie 2 et l'occurrence attendue si la fréquence était identique à celle des comptes de catégorie 4. Elle a pour avantage de faire ressortir plus facilement des termes surreprésentés qui

<sup>43</sup> Dans son moteur de recherche des sources d'information, le *Décodex* propose plusieurs couleurs de qualification. Un site qualifié de « rouge » s'accompagne de l'avertissement « *ce site diffuse un nombre significatif d'informations ou d'articles trompeurs. Restez vigilant et croisez avec d'autres sources plus fiables. Si possible, remontez à l'origine de l'information* » (1 Tweet concerné ici) ; un site « orange » d'un « *soyez prudents et croisez avec d'autres sources. Si possible, remontez à l'origine de l'information* » (4 Tweets). Enfin, les contenus non classifiés sont « gris » et l'avertissement affiché est le suivant : « *N'hésitez pas à confirmer l'information en croisant avec d'autres sources ou en remontant à son origine.* ». Le Monde. *Décodex : vérification de sources d'informations, pages Facebook et chaînes YouTube*. URL : <https://www.lemonde.fr/verification/>

<sup>44</sup> La méthode utilisée est celle de la distance de Levenshtein qui est une distance mathématique qui vise à mesurer la différence entre deux chaînes de caractère. Beaucoup de mots expriment la même idée mais sont orthographiés de manières différentes (pluriel, faute de frappe, espace, etc.). Afin de regrouper tous les mots identique ou quasi-identique, il a été décidé d'utiliser la distance de Levenshtein sur une base 1, à savoir, de souder ensemble tous les mots présentant jusqu'à une lettre de différence. Ainsi, les nombres d'occurrence de mots partageant une très forte proximité sont additionnés permettant à certaines thématiques d'être mise en avant plus justement.

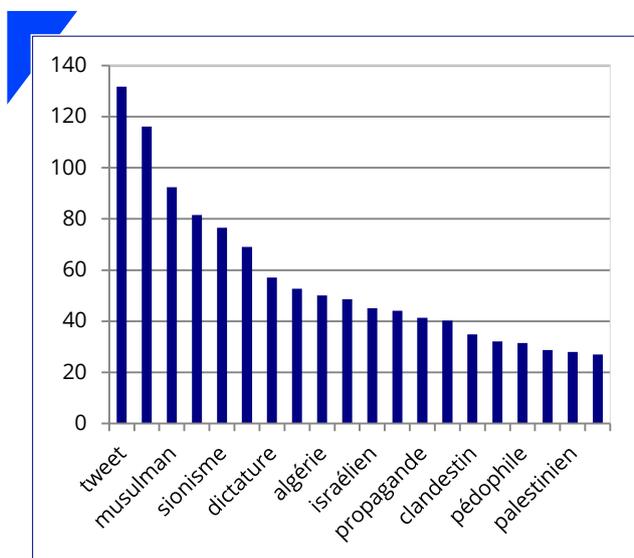
<sup>45</sup> La significativité de la surreprésentation a été estimée selon la méthode du log-likelihood  $G^2$  (Dunning, Ted. 1993. *Accurate Methods for the Statistics of Surprise and Coincidence*. Computational Linguistics, Volume 19, number 1, pp. 61-74). Seules les différences statistiquement significatives à 0,01% ont été conservées.



sont fréquemment utilisés par les comptes de catégorie 2. La deuxième méthode consiste à étudier non plus la différence absolue mais le rapport. Elle permet par exemple de faire ressortir des termes qui restent peu usités par les comptes de catégorie 2 mais qui le sont cependant significativement plus que sur les comptes de catégorie 4.

- **Méthode de la différence entre la fréquence effective et la fréquence attendue**

Le nuage de mots suivant représente les mots les plus surreprésentés au sein des comptes de la catégorie 2 par rapport à ceux de la catégorie 4 (en fonction de l'importance de la différence entre l'occurrence effective et l'occurrence attendue), tandis que le graphique présente les 20 termes les plus surreprésentés<sup>46</sup> :



<sup>46</sup> Pour la réalisation du nuage, les mots possédant une orthographe proche ont été assimilés. Le mot « français » représente alors aussi bien les occurrences de « françaï » que de « français ».





#### 4.4. Les analyses ne font pas apparaître de phénomène de chambre d'écho sur Twitter

Cette section s'intéresse à l'existence éventuelle d'un phénomène de « chambre d'écho » sur Twitter. La notion de chambre d'écho renvoie ici au risque qu'un individu exposé à des sources d'information peu fiables ne soit pas susceptible d'être exposé à des sources d'information fiables. Dans le cadre de cette étude, il s'agira d'étudier dans quelle mesure les abonnés de comptes de catégorie 2 (rouges) sont susceptibles d'être également des abonnés de comptes de catégorie 4 (fiables). Cette méthodologie est nécessairement simplificatrice pour plusieurs raisons. D'une part, elle ne permet pas de tenir compte de l'exposition des abonnés de comptes de catégorie 2 à d'autres sources d'informations fiables en dehors de Twitter. D'autre part, le fait d'être abonné d'un compte de catégorie 4 n'impliquera pas nécessairement que l'utilisateur va accorder une attention élevée aux Tweets publiés par ce compte et qu'il va par exemple lire les articles auxquels renvoient les liens pour éclairer son opinion. Enfin, il n'est pas tenu compte du fait que les abonnés de comptes de catégorie 2 sont également potentiellement exposés à des Tweets émanant de comptes de catégories 4 même si ils ne suivent pas ce type de comptes (notamment parce qu'ils suivent des personnes qui auraient retweeté ces comptes).

La définition de la chambre d'écho qui est retenue ici ne correspond pas nécessairement à celle examinée dans la littérature qui a notamment étudié l'existence de chambre d'échos idéologiques (par exemple en étudiant si les interactions entre utilisateurs se font exclusivement entre membres d'un même bord politique). Ces travaux ont pour certains confirmé la présence de chambres d'échos tandis que d'autres les ont réfutées<sup>47</sup>. Aux États-Unis, plusieurs études se sont intéressées au lien entre les phénomènes de chambre d'écho et la polarisation politique, concluant plutôt que la première est un résultat de la seconde et non l'inverse<sup>48</sup>.

Dans la présente étude, un utilisateur qui suit un compte de catégorie 2 (rouge) et un compte de catégorie 4 (fiable) ayant un positionnement politique similaire ou proche ne sera pas considéré comme étant enfermé dans une chambre d'écho.

L'analyse a été menée sur un échantillon de 30 utilisateurs choisis aléatoirement pour chaque compte de catégorie 2. Pour chacun de ces utilisateurs, la liste de l'ensemble des comptes suivis a été recueillie et, parmi cette liste, ont été identifiés les comptes labellisés par le *Décodex*. Si l'analyse ne porte que sur un échantillon restreint de près de 800 comptes Twitter, les résultats obtenus sont suffisamment marqués pour considérer que la taille de l'échantillon n'est pas susceptible d'impacter la conclusion.

<sup>47</sup> Council of Europe. *Information disorder : Toward an interdisciplinary framework for research and policy making*. 09/2017.« »

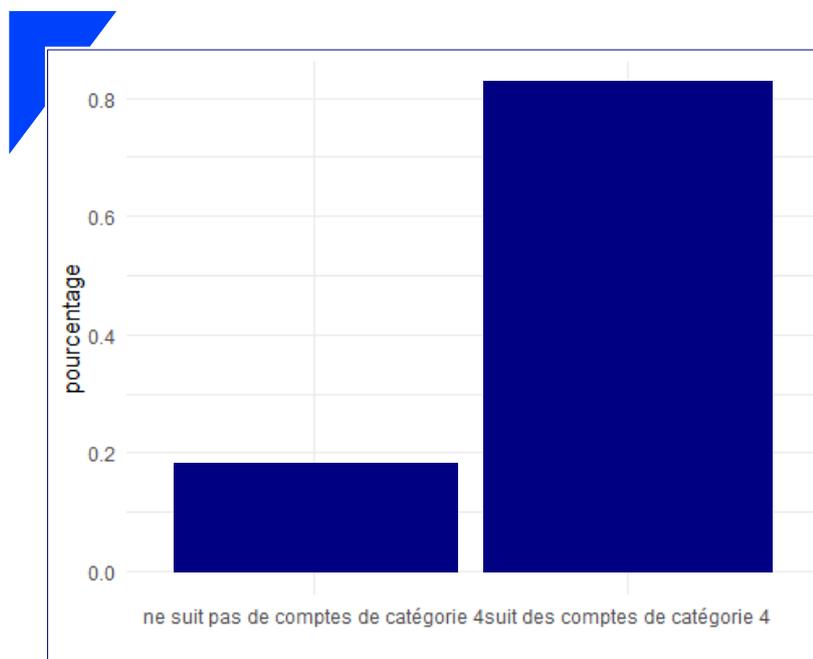
<sup>48</sup> Andrew Guess, Brendan Nyhan, Jason Reifler. *Selective Exposure to Misinformation: Evidence from the consumption of fake news during the 2016 U.S. presidential campaign*. 09/01/2018.



Le nombre médian de comptes suivis par les abonnés de comptes de catégorie 2 est résumé dans le tableau suivant :

CATÉGORIE 1	CATÉGORIE 2	CATÉGORIE 3	CATÉGORIE 4
0	2	1	6

Ces abonnés suivent un nombre plus élevé de comptes de catégorie 4 que de comptes de catégorie 2. On peut également examiner la proportion d'abonnés de comptes de catégorie 2 qui ne suivent aucun compte de catégorie 4 :



Plus de 80 % des abonnés de comptes de catégorie 2 suivent donc au moins un compte de catégorie 4. Et ils suivent plus de compte de catégorie 4 que de comptes de catégories 1, 2 et 3 cumulés. Ce résultat corrobore le constat le nombre d'abonnés de comptes de catégories 4 est beaucoup plus important que celui du nombre d'abonnés aux autres types de comptes.

Ces résultats confirment par ailleurs encore une fois les résultats mis en avant par Fletcher et al. (2018) qui montrent que les internautes qui consultent des sites recensés dans le *Décodex* comme partageant de fausses informations consultent également des sites Internet classés comme étant crédibles. Cette observation effectuée sur les sites internet est donc également vérifiée sur Twitter.



Le tableau ci-dessous représente par ailleurs les 10 comptes qui sont les plus suivis<sup>49</sup> par l'échantillon d'abonnés de catégorie 2 étudié :

COMPTE	CATÉGORIE	NOMBRE D'UTILISATEURS DE L'ÉCHANTILLON SUIVANT CE COMPTE
Le Monde	4	324
Mediapart	4	285
AFP	4	281
Le Figaro	4	258
BFMTV	4	217
Russia Today (version française)	3	211
Valeurs actuelles	3	210
TV Libertés	3	210
Franceinfo	4	208
Le Parisien-Aujourd'hui en France	4	204

Comme il a été vu précédemment, la médiane du nombre de sites de catégories 2 suivis par l'échantillon est de 2 tandis que la médiane du nombre de sites de catégorie 4 suivis est de 6. Le tableau ci-dessous présente les comptes du *Décodex* suivis par les mêmes abonnés identifiés qui suivent 2 comptes classés en catégorie 2 et entre 5 et 7 comptes classés en catégorie 4. Ces abonnés ne sont pas nécessairement représentatifs de l'ensemble des abonnés étudiés mais ils permettent d'illustrer quelques exemples de comportements d'abonnés proche de la médiane des abonnés des comptes de catégorie 2. Chaque ligne du tableau présente l'ensemble des comptes catégorisés par le *Décodex* suivis par un follower donné.

<sup>49</sup> Parmi les comptes labellisés par le *Décodex*.



---

## COMPTES SUIVIS PAR QUELQUES ABONNÉS « REPRÉSENTATIFS » DE COMPTES DE CATÉGORIE 2

---

["Al Kanz"], ["Mouv"], ["Tariq Ramadan"], ["Dieudosphere (Dieudonné)"], ["Jovanovic"], ["20 Minutes"], ["Le Monde"]]

["The Onion"], ["The Economist"], ["BuzzFeed"], ["The New York Times"], ["BBC"], ["Tprincedelamour"], ["AE911truth"], ["France 24"], ["Counterpunch"]]

["Panamza"], ["Egalité et Réconciliation"], ["L'Equipe"], ["AFP"], ["L'Obs"], ["L'Express"], ["Libération"], ["Le Figaro"], ["Le Parisien-Aujourd'hui en France"]]

["DailyMail"], ["Le Soir"], ["Boulevard Voltaire"], ["20 Minutes"], ["Hollande Démission"], ["Paris Match"], ["Valeurs actuelles"], ["Jeune Nation"], ["Canal+"], ["TF1"]]

["Panamza"], ["Le Monde diplomatique"], ["Réseau International"], ["Maître Eolas"], ["Acrimed"], ["Mediapart"], ["L'humour de droite"], ["Oumma"], ["Tariq Ramadan"], ["Al Kanz"]]

["Panamza"], ["Valeurs actuelles"], ["L'Obs"], ["Dieudosphere (Dieudonné)"], ["Mediapart"], ["Rue89"], ["Tariq Ramadan"], ["Libération"]]

["Réinformation RC"], ["TV Libertés"], ["LesObservateurs.ch"], ["Russia Today (version française)"], ["Melly"], ["Jean-Marc Morandini"], ["Valeurs actuelles"], ["Libération"], ["Fdesouche"], ["Huffington Post"], ["BFMTV"], ["AFP"]]

["Russia Today (version française)"], ["Jovanovic"], ["Agence Info Libre"], ["Etienne Chouard"], ["L'Equipe"], ["Dieudosphere (Dieudonné)"], ["Libération"], ["Le Monde"], ["Le Figaro"], ["Rue89"], ["Maître Eolas"]]

["Konbini"], ["Mediapart"], ["Korben"], ["Dieudosphere (Dieudonné)"], ["Egalité et Réconciliation"], ["Al Kanz"], ["Paris-Normandie"], ["France 2"]]

["BFMTV"], ["i24 News"], ["France 2"], ["JSS News"], ["Dreuz Info"], ["Europe Israël"], ["Nice-Matin"], ["Libération"], ["Ouest-France"], ["Valeurs actuelles"], ["Le Monde"]]

---

Les analyses précédentes conduisent donc à relativiser l'importance d'un phénomène d'enfermement qui devrait se traduire par une imperméabilité aux médias réputés fiables. Cependant, si ces phénomènes d'enfermement ne sont pas majoritaires, ils sont malgré tout susceptibles d'exister sur la plateforme. Il a ainsi été vu qu'un peu moins de 20 % des abonnés de comptes de catégorie 2 ne suivent aucun compte de catégorie 4.



Le tableau ci-dessous présente le comportement d'abonnements de certains utilisateurs suivants plus de 3 comptes propices à la désinformation et ne suivant aucun compte d'actualité catégorisé comme fiable :

---

**COMPTES SUIVIS PAR QUELQUES ABONNÉS DE COMPTES DE CATÉGORIE 2  
PRÉSENTANT UN PHÉNOMÈNE D'ENFERMEMENT**

---

["Réseau International"], ["Panamza"], ["Médias Presse Info"], ["Jovanovic"], ["Valeurs actuelles"], ["LesObservateurs.ch"], ["Boris Le Lay"], ["Dieudosphere (Dieudonné)"], ["TV Libertés"]

["Valeurs actuelles"], ["Boulevard Voltaire"], ["Boris Le Lay"], ["Breizh Info"], ["TV Libertés"], ["Le Gorafi"], ["Jeune Nation"], ["Panamza"], ["Médias Presse Info"], ["Fdesouche"], ["Le site du Professeur Henri Joyeux"], ["Sputniknews"], ["Dieudosphere (Dieudonné)"], ["Egalité et Réconciliation"]

["Antipresse"], ["Réinformation RC"], ["OJIM"], ["Fdesouche"], ["Egalité et Réconciliation"], ["TV Libertés"], ["Boris Le Lay"], ["Agence Info Libre"], ["Diktacratie"]

["Médias Presse Info"], ["Réinformation RC"], ["Paris Vox"], ["Jeune Nation"], ["Novopress"], ["Breizh Info"], ["Russia Today (version française)"], ["TV Libertés"], ["Jovanovic"], ["Etienne Chouard"], ["Agence Info Libre"], ["Sputniknews"], ["Egalité et Réconciliation"], ["Russia Today"]

["Panamza"], ["Diktacratie"], ["Etienne Chouard"], ["Jovanovic"], ["Agence Info Libre"], ["Egalité et Réconciliation"], ["Dieudosphere (Dieudonné)"]

["Hollande Démission"], ["Russia Today (version française)"], ["Europe Israël"], ["Riposte laïque"], ["Dreuz Info"], ["Russia Today"], ["Breitbart"], ["Sputniknews"], ["TV Libertés"], ["Réseau International"]

["Diktacratie"], ["Egalité et Réconciliation"], ["Dieudosphere (Dieudonné)"], ["Panamza"], ["Etienne Chouard"], ["Agence Info Libre"], ["Jovanovic"]

["Tprincedelamour"], ["Panamza"], ["Fawkes News"], ["Novopress"], ["Boulevard Voltaire"], ["Valeurs actuelles"], ["Riposte laïque"], ["Boris Le Lay"], ["TV Libertés"], ["Sputniknews"], ["Russia Today (version française)"], ["Dieudosphere (Dieudonné)"], ["Egalité et Réconciliation"], ["Etienne Chouard"], ["Agence Info Libre"], ["Jovanovic"], ["Fdesouche"]

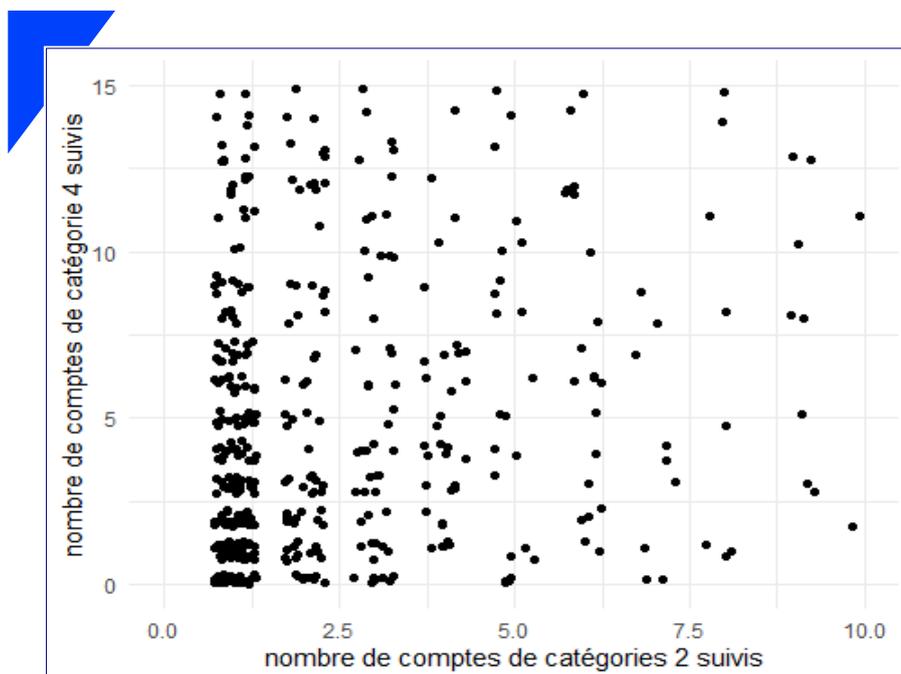
["Boris Le Lay"], ["Hollande Démission"], ["Tprincedelamour"], ["La gauche m'a tuer"]

["Breizh Info"], ["Le Gorafi"], ["Réinformation RC"], ["Russia Today (version française)"], ["Boris Le Lay"], ["Les News"], ["TV Libertés"], ["Riposte laïque"], ["OJIM"]

---



Le graphique suivant représente enfin le nombre de comptes de catégorie 4 suivis en fonction du nombre de comptes de catégorie 2 suivis par l'utilisateur. Cette analyse fait apparaître la diversité des comportements<sup>50</sup> mais montre également que le cas le plus fréquent est celui d'un utilisateur qui ne suit qu'un compte de catégorie 2 et un nombre significativement plus important de comptes de catégories 4 :



#### Synthèse :

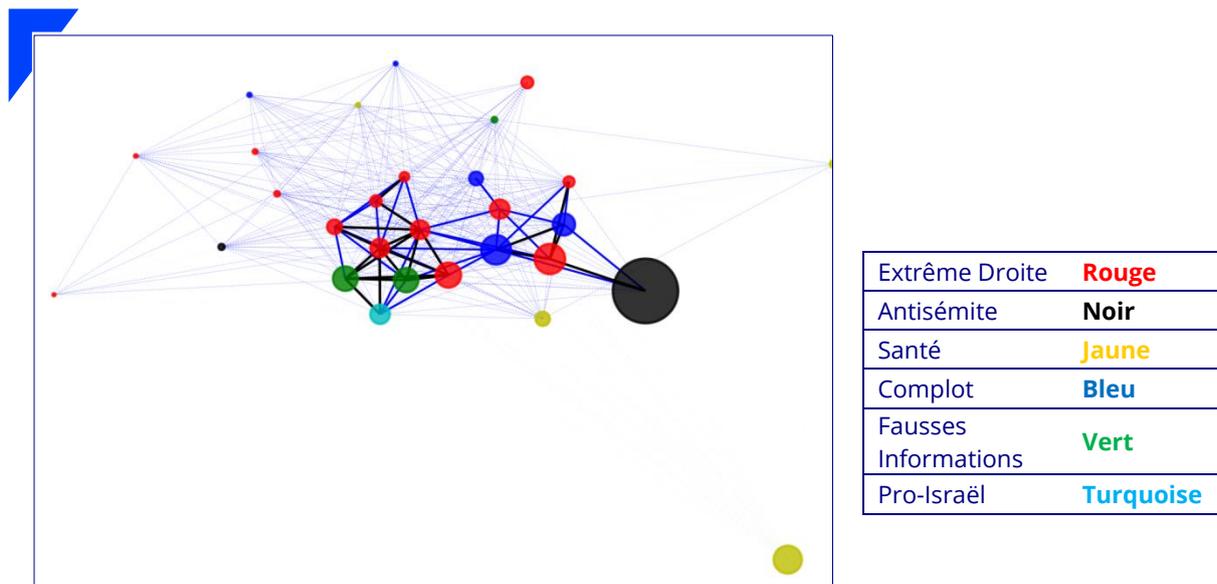
Les individus exposés à des sources connues pour partager des fausses informations, sans suivre dans le même temps des comptes fiables sont très minoritaires. Ils suivent même en moyenne une majorité de comptes fiables. Moins de 20 % des abonnés de ces comptes ne suivent en effet aucun compte fiable. Ces individus présentent ainsi des comportements d'abonnements qui ne leur permettent pas d'être confrontés à des sources d'information susceptibles de corriger les fausses informations reçues par ailleurs.

<sup>50</sup> Les comptes suivants plus de 15 comptes de catégorie 4 ou plus de 10 comptes de catégorie 2 sont très minoritaires et ne sont pas représentés pour faciliter la lecture du graphique.



## 4.5. Analyses du graphe des abonnés

L'analyse précédente a permis d'examiner sur un échantillon d'utilisateurs la diversité des comportements d'abonnements, et plus particulièrement la propension des abonnés de comptes non fiables à suivre des comptes fiables. La section présente s'intéresse également au comportement d'abonnement des utilisateurs en cherchant à identifier l'existence de proximités entre ces différents comptes. Cette analyse est basée sur un examen de la totalité des abonnés de chaque compte de catégorie 2. Elle conduit à proposer une analyse sous forme de graphe ou réseau faisant apparaître la proximité entre chaque compte de catégorie 2 en fonction de leur proportion d'abonnés communs. Les points (nœuds) au sein du graphe représentent les comptes de catégorie 2 et leur taille est proportionnelle au nombre d'abonnés. L'épaisseur des liens reliant les comptes est, elle, proportionnelle au nombre d'abonnés communs entre les comptes. En utilisant les descriptions utilisées par les Décodeurs pour justifier leurs choix de classification<sup>51</sup>, des sous-catégories de comptes de catégorie 2 ont pu être créées et sont représentés dans le graphe suivant par des couleurs distinctes<sup>52</sup> :



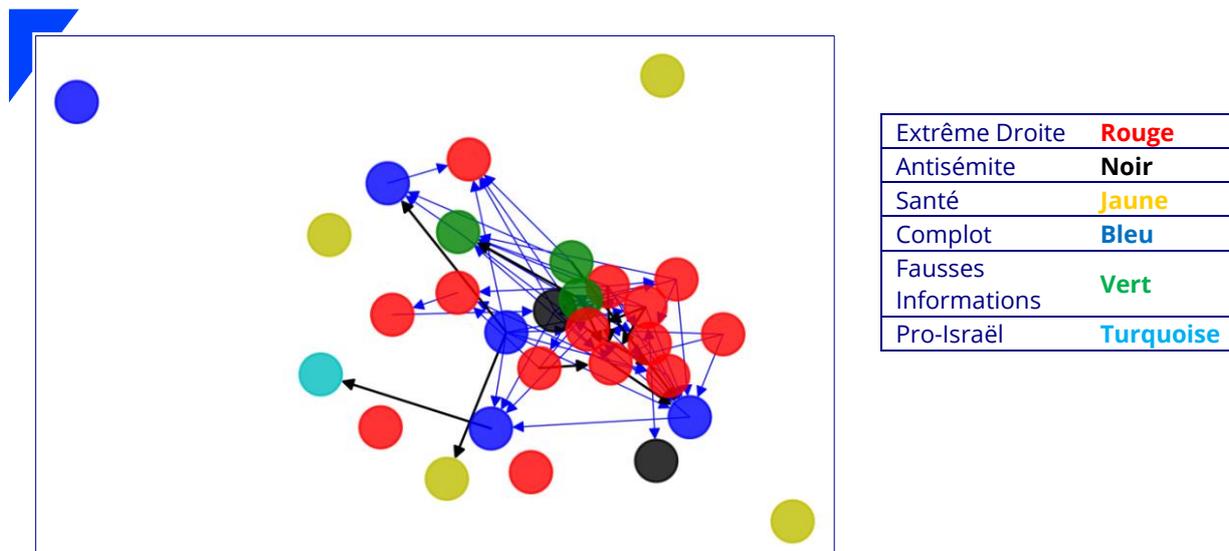
Le graphe permet d'observer que certains comptes partagent un nombre important d'abonnés communs (comptes reliés par des liens plus épais). C'est particulièrement le cas des comptes à thématiques d'Extrême Droite entre eux mais également de ces comptes avec les comptes relayant des fausses informations de manière générale.

<sup>51</sup> Le Monde. *Décodex : vérification de sources d'informations, pages Facebook et chaînes YouTube.*

<sup>52</sup> Pour être totalement exhaustif, six sous-catégories ont été constituées : la catégorie des comptes décrits comme antisémites ou antisionistes ; celle des comptes d'extrême droite ; la catégorie des comptes qui promeuvent une santé « alternative » (certains de ces comptes ont une position anti-vaccin) ; une catégorie de compte pro-israélien ; une autre catégorie regroupe les comptes catégorisés comme « complotistes » ; enfin, une dernière catégorie désigne les comptes « relais » de fausses informations sans plus de précision.



Un second graphe (reprenant le même code couleur) représente le comportement d'abonnement des comptes de catégorie 2 eux-mêmes. Des comptes de catégorie 2 sont liés si les comptes se suivent entre eux (la flèche représentant le sens de l'abonnement). Ce graphe permet de montrer une forte proximité entre certains comptes d'extrême droite, des comptes partageant des fausses informations et des comptes complotistes. D'autres comptes, notamment ceux appartenant au thème de la santé apparaissent plus autonomes :



**Synthèse :**

De nombreux comptes de catégorie 2 présentent une proximité importante : ils possèdent une proportion significative d'abonnés communs ou ils se suivent mutuellement sur Twitter.



#### 4.5.1. Examen de l'activité des comptes des vérificateurs d'information

La présente section porte sur une catégorie de comptes ne relevant pas de la typologie du *Décodex* : les comptes de journalistes vérificateurs d'information (comptes de *fact-checking*). Cette analyse ne prétend pas à l'exhaustivité. Elle est fondée sur l'étude des comptes Twitter de 20minFakeOff (*20 Minutes*), AFPFactuel (Agence France-Presse), Les Décodeurs (*Le Monde*), InfoIntoxF24 (Les Observateurs de France 24) et CheckNewsfr (*Libération*).

Les statistiques présentées dans le tableau suivant proviennent d'une collecte d'informations effectuée le 26 novembre 2019 :

NOM DU COMPTE	NOMBRE DE TWEETS ANALYSÉS	NOMBRE MOYEN DE RETWEETS PAR TWEET	NOMBRE D'ABONNÉS	NOMBRE TOTAL DE TWEETS PUBLIÉS PAR LE COMPTE	DATE DE CRÉATION DU COMPTE	NOMBRE MOYEN DE RETWEETS PAR TWEET POUR 1 000 ABONNÉS
20minFakeOff	475	6,93	2 902	480	2018-09-18 11:58:15	2,39
AFPFactuel	2 451	78,91	91 012	2 775	2017-11-20 14:24:50	0,87
CheckNewsfr	3 185	7,01	40 680	7 145	2017-09-26 16:25:10	0,17
Décodeurs	3 083	17,16	141 819	17 070	2009-11-03 10:32:57	0,12
InfoIntoxF24	891	5,04	4 752	944	2019-02-25 11:33:55	1,06

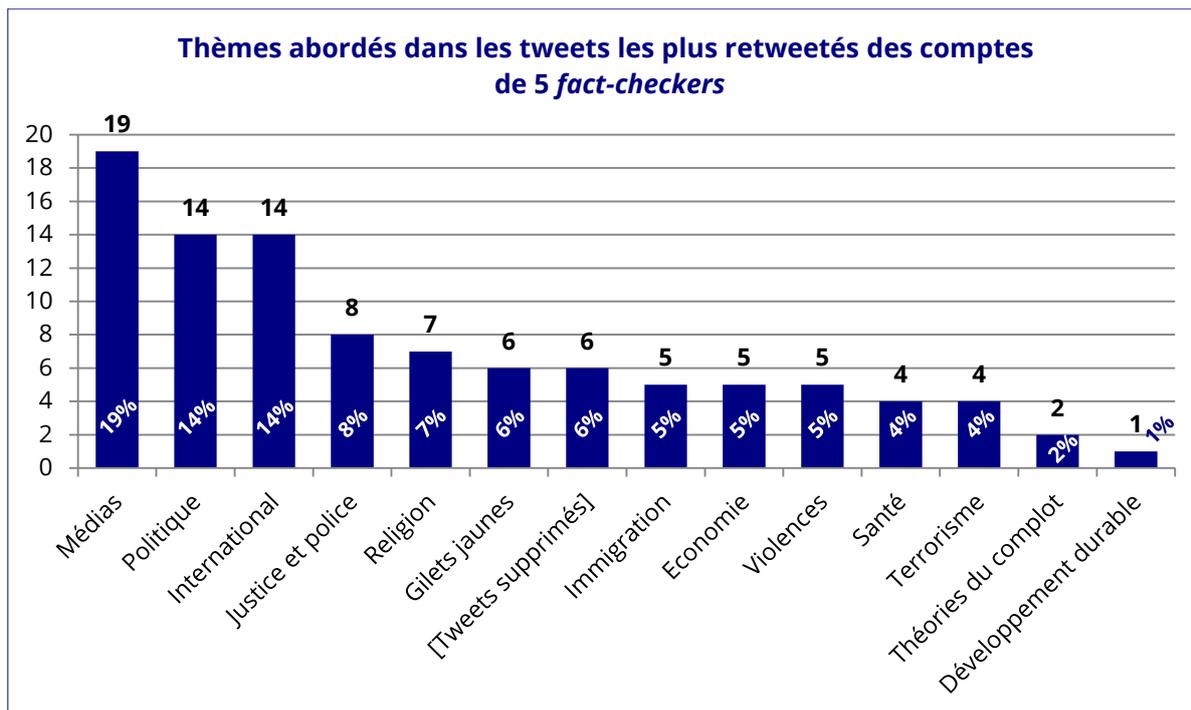
Certains comptes comme celui de AFPFactuel ou des Décodeurs et dans une moindre mesure de Checknews sont suivis par un nombre d'abonnés très important. Rappelons que le nombre moyen d'abonnés des comptes de catégorie 2 est de près de 14 000. Le nombre moyen de ReTweets par Tweet est relativement faible sauf pour les comptes de AFPFactuel et des Décodeurs. En comparaison, le nombre moyen de ReTweets par Tweet pour les comptes de catégorie 2 est de 18. Il est par ailleurs de 14 pour les comptes de catégorie 4. Les comptes de *fact-checking* parviennent enfin pour certains à obtenir un partage substantiel de la part de leurs abonnés. Les comptes des catégories 2 et 4 présentent en moyenne respectivement 1,5 ReTweet et 0,15 ReTweet pour 1 000 abonnés.

Afin de présenter de façon plus concrète l'activité des *fact-checkers* sur Twitter et afin de mieux comprendre la nature des Tweets qui sont les plus propices au partage, le contenu des vingt Tweets qui ont fait l'objet du nombre le plus élevé de ReTweets parmi les *fact-checkers* sélectionnés est analysé à la suite.



#### 4.5.2. Analyse des thématiques abordées par les Tweets les plus retweetés des *fact-checkers*

Cette analyse a été réalisée sur les 20 Tweets les plus retweetés de chaque compte<sup>53</sup>. Utilisant la même grille d'analyse que celle utilisée pour les Tweets des comptes de catégorie 2, les Tweets des vérificateurs d'information (*fact-checkers*) ont été classés selon 22 thèmes différents<sup>54</sup>.



L'analyse des thèmes se concentrera sur les plus représentés dans ce corpus : « Médias » (19 Tweets), « Politique » (14) et « International » (14) représentent ainsi 47 % des Tweets les plus retweetés. À titre de comparaison, la thématique « Immigration », la deuxième la plus représentée chez les comptes de catégorie 2, n'apparaît qu'à la 9<sup>ème</sup> place chez les comptes de *fact-checkers*. Dans une perspective comparative, elle sera néanmoins ajoutée à l'analyse thématique.

<sup>53</sup> Parmi les 100 Tweets initialement collectées, certains étaient des réponses à des publications d'internautes, pour lesquelles les *fact-checkers* ont poussé leurs articles de vérification. Dans ce type de partage, les auteurs du Tweet initial ont parfois supprimé leur publication, entraînant de fait l'inaccessibilité du contenu du *fact-checker* lié. 6 Tweets n'ont ainsi pas pu être analysés (n = 94).

<sup>54</sup> Chaque Tweet n'a été compté qu'une seule fois, selon la thématique principale qu'il abordait. Cette méthode peut présenter des limites en ce que certains Tweets peuvent aborder plusieurs thématiques. Les tendances et proportions mises en avant restent cependant valides compte tenu des écarts importants constatés entre les thèmes.



- **Médias (19 Tweets)**

Parmi les 19 Tweets de cette catégorie, 12 se concentrent sur la diffusion d'outils pour aider le public à exercer son sens critique (référence verticale), 5 communiquent autour d'initiatives de *fact-checking*, notamment européennes (référence circulaire), et 2 Tweets s'attachent à vérifier ou relever les erreurs faites par d'autres médias (référence horizontale). Contrairement aux comptes de catégorie 2, il ne s'agit aucunement de critiquer les médias dans leur ensemble, mais plutôt d'aider les internautes à naviguer dans un environnement caractérisé par l'instantanéité et la densité. Le volume de Tweets récoltés appartenant à cette thématique témoigne ainsi, au-delà du thème traité, d'une promotion de la méthode de vérification des faits utilisée par les journalistes.

Une part importante des Tweets de cette catégorie (12) propose donc aux utilisateurs des outils afin de les accompagner dans leur consommation d'information sur internet (référence verticale). Ces publications se répartissent entre diffusion d'outils de vérification de photos ou de vidéos et exemples de cas illustrant les ressorts d'une mauvaise interprétation d'une information.



Exemples de Tweets sur la référence verticale publiés par les fact-checkers.



Au sein du corpus des Tweets les plus retweetés, Checknewfr est le seul à avoir publié un message relatif à des vérifications d'informations publiées par des confrères journalistes (2 Tweets).

Exemple de Tweet sur la référence horizontale publié par un fact-checkeur.

Enfin, 5 Tweets font la promotion d'initiatives de *fact-checking* ou de l'indépendance des médias dont ils émanent (vérification circulaire). C'est notamment le cas des Décodeurs du *Monde*, média qui a été au cœur d'une problématique liée à la préservation de l'indépendance de la rédaction lors d'un changement d'actionnaires. Les exemples de communication les plus emblématiques restent ici ceux ayant trait à des initiatives entre plusieurs rédactions pour faire de la vérification de l'information. Il s'agit dans les deux cas plutôt de ReTweets de ces organismes (3 cas sur 5) que de Tweets publiés par le compte étudié.



"Nous, journalistes du Monde..." : dans une démarche inédite, plus de 460 journalistes de la rédaction du Monde signent un texte pour demander aux actionnaires du groupe de préserver leur indépendance éditoriale. [lemonde.fr/actualite-medi...](http://lemonde.fr/actualite-medi...)



Exemples de Tweets sur la référence circulaire publiés par les fact-checkers.

• Politique (14 Tweets)

Parmi les Tweets de ce corpus, une majorité traite du chef de l'État ou du gouvernement et de ses membres (10 Tweets). Le reste des Tweets concerne des personnalités spécifiques (3) tandis qu'un seul aborde le monde politique dans son ensemble.

Les publications traitant du Président de la République ou du gouvernement sont également celles qui concentrent les Tweets ayant débouché sur le plus de réactions ; il s'agit souvent de Tweets visant à évaluer les propos tenus par des personnalités (7 sur 10). Les Décodeurs analysent par exemple l'interdiction des néonicotinoïdes, dont le gouvernement actuel se serait attribué les mérites alors que celle-ci avait été prise, selon eux, avant son arrivée au pouvoir. Les 3 autres Tweets abordent par exemple des événements auxquels ont participé des membres de l'exécutif.



Exemples de Tweets sur le chef de l'État et sur le gouvernement publiés par les fact-checkers.

Les autres Tweets de cette thématique portent sur d'autres personnalités politiques et sur le monde politique dans son ensemble, partent plutôt de questionnements venant par exemple de citoyens mettant en doute la véracité des propos tenus par des personnalités. Les *fact-checkers* y font référence en utilisant la forme interrogative ou des expressions comme « *selon cette rumeur* ». Il s'agit ici de partir d'un soupçon qui aurait été lu, vu ou entendu par les journalistes et de lancer une vérification à partir de cela. Le cas de la photo du président de la commission des Finances de l'Assemblée nationale et ancien ministre Éric Woerth, dont la véracité de son escalade de l'aiguille d'Argentière a été fortement questionnée, est emblématique de cette vérification qui part de doutes exprimés par des internautes.

Comparativement aux Tweets les plus populaires des comptes de catégorie 2, les personnalités d'extrême-droite sont sous-représentées dans le corpus ici étudié.



- **International (14 Tweets)**

Les 14 Tweets de cette catégorie s'intéressent majoritairement à deux zones géographiques : l'Union européenne (5 Tweets) et la Russie (3). Les autres Tweets abordent les feux forestiers en Amazonie ainsi que d'autres thématiques mais de manière isolée (opération militaire au Mali, fait divers relatif à un footballeur français lors d'un séjour aux Émirats Arabes Unis, canular en Afrique).

Les élections européennes de mai 2019 d'une part et une prétendue interdiction des banques Rothschild en Russie, d'autre part, suscitent le plus de ReTweets. Dans le premier cas, il s'agit de faire de la pédagogie. Dans le second, c'est la vérification d'une information qui est à l'œuvre. Il s'agit donc de s'intéresser à une information tangente à l'actualité nationale, et non de vérifier une rumeur sans lien avec la France comme certains comptes de catégorie 2 ont pu le faire.



*Exemples de Tweets sur l'Union européenne et la Russie publiés par les fact-checkers.*

- **Immigration (5 Tweets)**

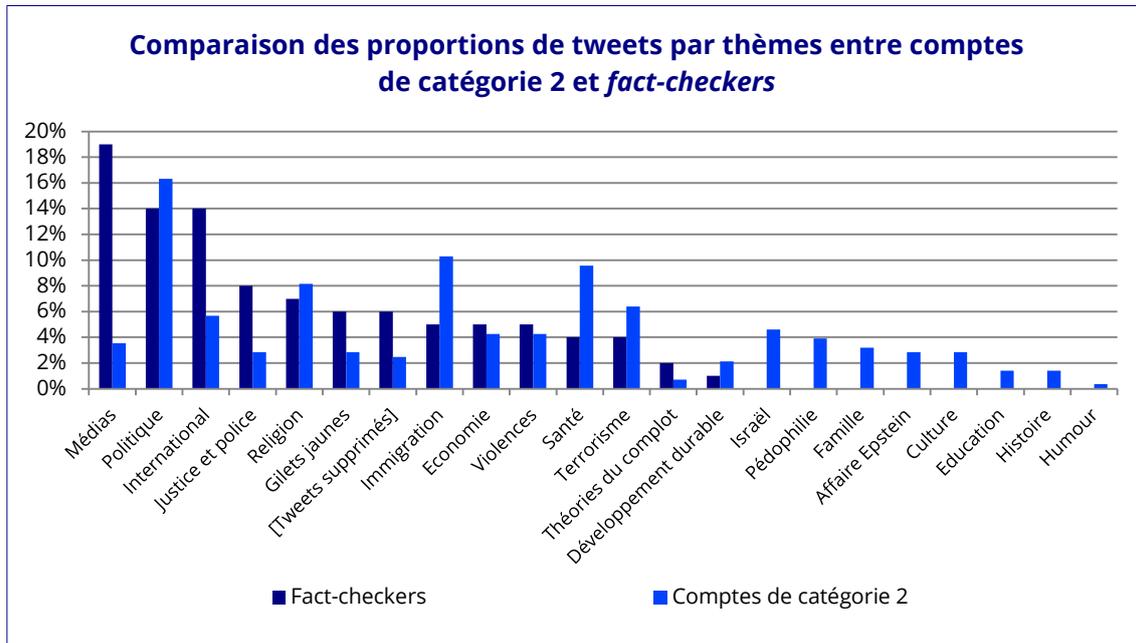
Thématique très investie par les comptes de catégorie 2, dont les Tweets les plus retweetés abordent à 29 reprises ce sujet (10,3 % du corpus), l'immigration est également traitée par les *fact-checkers*. Il s'agit ici de vérifier des propos tenus notamment par des personnalités. 4 des 5 Tweets de cette catégorie concernent ainsi des propos d'élus Rassemblement national (RN) qui s'avèrent faux selon les journalistes parce qu'ils se basent soit sur des images sorties de leur contexte, soit sur des chiffres qui s'avèrent erronés.



Exemples de Tweets sur l'immigration publiés par les fact-checkers.

#### 4.5.3. Comparaison des résultats obtenus avec ceux des comptes de catégorie 2

L'analyse comparée des Tweets les plus partagés sur les comptes de catégorie 2 et sur ceux des journalistes spécialisés dans la vérification de l'information a montré que **la politique** (1<sup>er</sup> chez les comptes de catégorie 2, 2<sup>e</sup> chez les *fact-checkers*) **et la religion** (4<sup>e</sup> contre 5<sup>e</sup>) sont également très partagées dans les deux catégories. Les thématiques portant sur l'immigration (2<sup>e</sup> contre 9<sup>e</sup>), la santé (3<sup>e</sup> contre 11<sup>e</sup>) et le terrorisme (5<sup>e</sup> contre 12<sup>e</sup>) ont suscité des engagements très différents. Elles ont fait l'objet d'une rediffusion importante chez les comptes de catégorie 2, à la différence des comptes des *fact-checkers*.



Ces résultats sont cohérents avec l'analyse menée par une doctorante et des étudiants de SciencesPo<sup>55</sup> : en janvier 2020, ils ont analysé les 50 derniers articles de trois rubriques de *fact-checking* (AFP Factuel, Les Décodeurs et Checknews) et ont trouvé que les sujets traités étaient majoritairement d'ordre politique (à 30, 33 et 28 % respectivement) tandis que l'environnement (2, 10 et 6 %), la religion (4, 6 et 6 %) et la santé (6, 8 et 2 %) y occupaient une place marginale<sup>56</sup>.

#### 4.5.4. Analyse des traitements de l'information

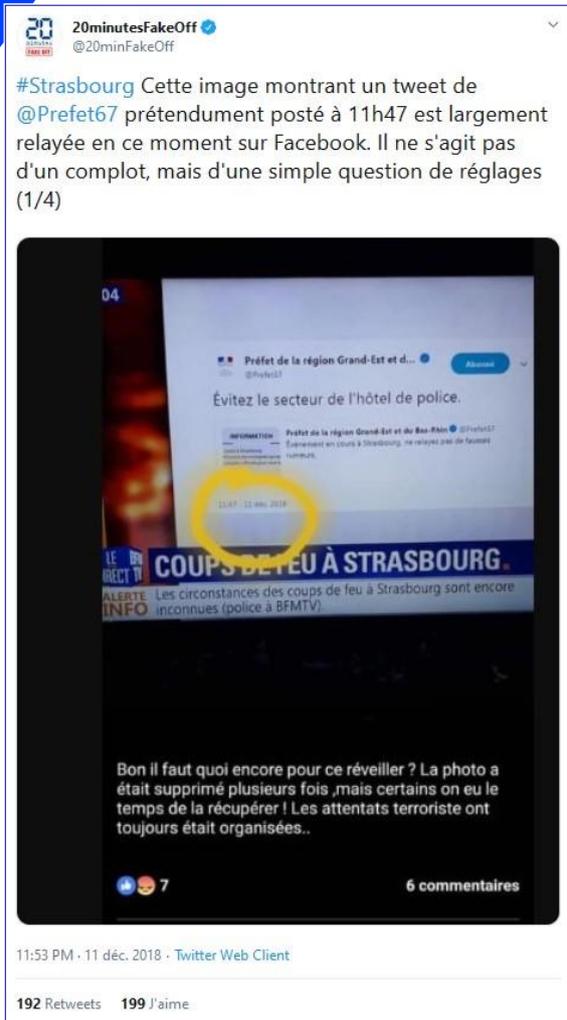
Basé sur les 5 Tweets les plus retweetés de chaque compte de *fact-checkers* étudié, un corpus de 25 Tweets a été constitué qui a permis de mettre en évidence, **parmi les ressorts qui semblent susciter le plus d'intérêt chez les internautes, la vérification d'un contenu déjà abondamment relayé et le recours à l'adresse directe** notamment en commentant le contenu directement en commentaire du *Tweet* publié par l'auteur, plutôt que la vérification de propos de personnalités politiques par exemple. Il s'agit de « *formats médiatiques facilitant le dialogue avec le grand public* » qui seraient, selon des recherches menées à SciencesPo, les plus efficaces et attractifs pour les internautes<sup>57</sup>.

7 Tweets de notre corpus font ainsi référence à des contenus largement diffusés (« *cette image [...] est largement relayée* » ou « *122.000 partages sur FB pour ces 7 photos* »). On retrouve ici le phénomène de référence verticale précédemment décrit.

<sup>55</sup> Manon Berriche. Sciences Po. *Le fact-checking est-il vraiment efficace ?*. 24/01/2020.

<sup>56</sup> Ce résultat est à remettre dans le contexte de sa date de publication, en janvier 2020, antérieure au large traitement médiatique de la Covid-19.

<sup>57</sup> Ibid.



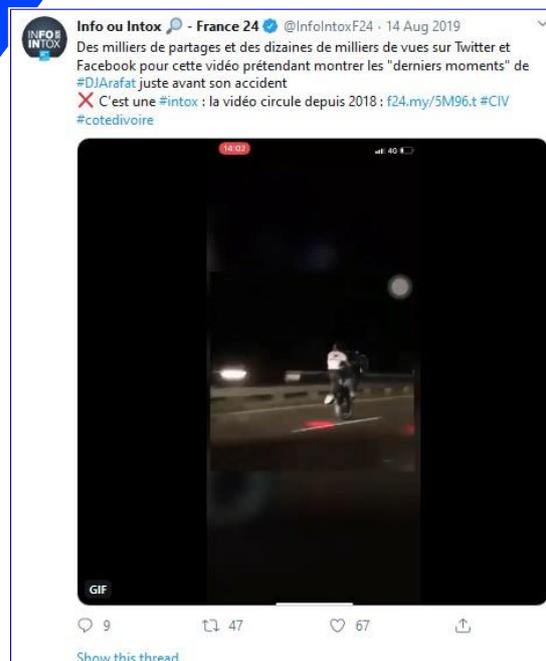
Exemples de Tweets sur la vérification d'un contenu très relayé par les fact-checkers.

Le recours à l'adresse directe (3 Tweets) est quant à lui utilisé principalement par AFP Factuel qui insère sa vérification dans un dialogue avec les internautes relayant des informations qui s'avèrent fausses, ainsi que par CheckNewsfr. Ce type de publication consiste ainsi non pas à publier un Tweet mais à répondre à un premier message, qui relaie par exemple une information ensuite qualifiée d'erronée (*thread* ou conversation). Elle permet d'une part d'atteindre la personne à l'origine de cette fausse information plus facilement, notamment parce qu'elle reçoit une notification, mais aussi d'être visible de son audience et de donner au travail de vérification une diffusion plus ciblée vers les personnes sensibles aux fausses informations.



*Exemples de Tweets utilisant l'adresse directe publiés par les fact-checkers.*

Enfin, en matière de type de contenu utilisé, le recours au non-textuel comme les photos (9), la citation de Tweets vérifiés (6) ainsi que l'animation ou *gif* (1) pour appuyer le propos de vérification est notable. C'est en particulier le cas des Tweets s'attachant à des vérifications d'images, afin de montrer la décontextualisation voire la manipulation qui peuvent se cacher derrière.



Exemples de Tweets utilisant des contenus vidéos publiés par les fact-checkers.

### Synthèse :

Les comptes de cellules journalistiques spécialisées dans la vérification de l'information sont en moyenne davantage suivis que les comptes de catégorie 2, mais leurs publications génèrent moins d'interactions que celles de ces derniers. En termes de thématiques majoritairement traitées, seules la politique et la religion sont communes à ces deux types de comptes. Les *fact-checkers* s'intéressent davantage à des thématiques en lien avec leur métier, comme l'actualité des médias ou des contenus qui relèvent de l'éducation aux médias et à l'information.



## 5. Analyse des Tweets liés à de fausses informations : chronologie de la propagation et de la correction

La présente partie aborde la problématique des fausses informations sous un angle complémentaire de la partie précédente en identifiant les Tweets se rattachant à une fausse information donnée et en se proposant d'identifier ceux propageant cette fausse information et ceux qui vont au contraire la corriger. Les comptes catégorisés par le *Décodex* étudiés dans la précédente section ne représentent en effet pas l'exhaustivité des comptes susceptibles de partager ou de corriger des fausses informations sur Twitter. À partir de quelques cas d'étude, la méthodologie utilisée dans cette section permettra donc d'acquérir une meilleure compréhension des phénomènes de propagation des fausses informations.

### 5.1. Description des fausses informations sélectionnées et de la méthodologie de collecte de données

Les fausses informations étudiées dans cette partie ont été identifiées sur les sites de vérificateurs d'informations. Pour chaque fausse information sélectionnée, des mots clés ont été choisis pour permettre d'identifier les Tweets s'y rattachant (ils ont été testés afin de vérifier qu'ils étaient les plus appropriés). L'API de Twitter a alors été utilisée afin de collecter l'ensemble des Tweets contenant ces mots clés. L'API de Twitter dans sa version gratuite (qui est celle utilisée dans la présente étude) ne permet de collecter que les Tweets émis sur les 7 derniers jours (approximativement). Cette limite temporelle constitue une contrainte forte puisqu'elle implique que les fausses informations soient identifiées moins de 7 jours après leur apparition sur le réseau social afin d'être en mesure de collecter l'ensemble de l'historique de diffusion de cette fausse information. Pour certaines des fausses informations sélectionnées, les Tweets relatifs aux mots clés sélectionnés ont par ailleurs fait l'objet de plusieurs collectes séparées dans le temps de façon à permettre l'étude de la propagation de la fausse information sur une durée supérieure à 7 jours.

Les fausses informations sélectionnées ne sont pas exhaustives et ne sont par ailleurs pas non plus nécessairement représentatives de l'ensemble des fausses informations. Ainsi, le fait de se concentrer sur les seules fausses informations ayant fait l'objet de réfutations par des vérificateurs d'informations pourrait conduire à sélectionner des fausses informations plus particulièrement virales. Les analyses constitueront donc des études de cas permettant d'apprécier des exemples de propagation de fausses informations sur Twitter sans qu'il soit nécessairement possible de généraliser l'ensemble de résultats obtenus.

Le tableau suivant synthétise les fausses informations étudiées dans cette partie. Elles ne correspondent pas toutes à des manipulations volontaires pouvant se rapprocher de la notion de désinformation mais répondent toutes au seul critère utilisé par le Conseil, à savoir qu'elles ont été vérifiées par un site de fact-checking.



THÈME	MOTS-CLÉS SÉLECTIONNÉS	DESCRIPTION DE LA FAUSSE INFORMATION ET LIEN VERS UN ARTICLE D'UN VÉRIFICATEUR
<b>Activité de Greta Thunberg sur Facebook</b>	Greta et Facebook	Les messages de Greta Thunberg sur Facebook seraient écrits par son père et un militant écologiste suédois/activiste indien <sup>58</sup> .
<b>Brevet sur le coronavirus</b>	Brevet et Coronavirus ou coronavirus et breveté	Le nouveau coronavirus aurait été breveté en 2003 aux États-Unis, ce qui prouverait que le virus a été créé volontairement <sup>59</sup> .
<b>Colis /coronavirus</b>	Colis et Coronavirus	Le coronavirus pourrait se transmettre via les colis en provenance de Chine <sup>60</sup> .
<b>Coronavirus à Argenteuil</b>	Argenteuil et Coronavirus	En janvier, des personnes atteintes du Coronavirus auraient été hospitalisées à Argenteuil <sup>61</sup> .
<b>Coronavirus à Perpignan</b>	Perpignan et Coronavirus	En janvier, des cas de coronavirus auraient été détectés à Perpignan <sup>62</sup> .
<b>Gale à l'hôpital de Nevers</b>	Epidémie et Nevers ou Gale et Nevers	L'hôpital de Nevers serait placé en quarantaine suite à une épidémie de gale due à l'hospitalisation de migrants
<b>Liens entre Bill Gates et le Coronavirus</b>	Gates et coronavirus	La fondation Bill Gates a été liée à plusieurs fausses informations : la fondation aurait prévu l'épidémie de coronavirus et elle aurait financé le groupe détenant le brevet sur le nouveau coronavirus <sup>63</sup>
<b>Panneau d'affichage à Limoges</b>	Limoges et panneau ou Limoges et panneaux ou Limoges et femmes	Un panneau d'affichage de la ville de Limoges demanderait aux femmes de ne pas être seules dans la rue et d'éviter les ruelles sombres <sup>64</sup> .
<b>Netflix</b>	Netflix et partage	Netflix mettrait fin au partage de compte <sup>65</sup> .
<b>Pollution de l'eau à Rouen</b>	Eau et Rouen	Suite à l'incendie de l'usine Lubrizol l'eau ne serait plus potable à Rouen <sup>66</sup>
<b>Propos de M. Blanquer</b>	Blanquer et 99.9%	Le ministre Jean-Michel Blanquer aurait soutenu que 99,9% des enseignants soutiennent la réforme du bac <sup>67</sup> .
<b>Rémunération des infirmières</b>	infirmière et happy meal	Une infirmière à domicile serait rémunérée l'équivalent d'un 'Happy meal' par patient <sup>68</sup> .
<b>TPMP</b>	TPMP et déprogrammé ou TPMP et CSA	Une source interne au CSA aurait annoncé la déprogrammation de l'émission « Touche pas à mon poste » (TPMP) <sup>69</sup> .

<sup>58</sup> Libération. *Un bug sur Facebook a-t-il révélé que les messages de Greta Thunberg étaient écrits par son père ?*. 16/01/2020.

<sup>59</sup> Le Monde. *Le coronavirus qui sévit en Chine n'a pas été « créé en 2003 aux USA »*. 27/01/2020.

<sup>60</sup> franceinfo. *Coronavirus : un colis en provenance de Chine peut-il transmettre la maladie ?*. 28/01/2020.

<sup>61</sup> AFP. *Nouveau coronavirus : attention aux fausses captures d'écran sur les réseaux sociaux*. 28/01/2020.

<sup>62</sup> 20 Minutes. *Coronavirus : Non, aucun cas n'a été détecté à Perpignan, indique la préfecture*. 26/01/2020.

<sup>63</sup> Le Monde. *Coronavirus : Bill Gates ciblé par des rumeurs et infox complotistes*. 05/02/2020.

<sup>64</sup> Le Monde. *« Mesdames, évitez de rentrer seule le soir... » : le message factice de la ville de Limoges*. 12/11/2019.

<sup>65</sup> franceinfo. *Netflix : la rumeur sur la suppression du partage de compte fait son grand retour*. 23/10/2019.

<sup>66</sup> Le Monde. *Après l'incendie de l'usine Lubrizol à Rouen, des intox en série*. 30/09/2019.

<sup>67</sup> Libération. *Blanquer a-t-il vraiment dit que «99,9% des enseignants soutiennent la réforme du bac» ?*. 23/01/2020.

<sup>68</sup> franceinfo. *Une infirmière à domicile est-elle payée moins par patient que le prix d'un "Happy Meal" ?*. 18/10/2019.

<sup>69</sup> 20 Minutes. *« Touche pas à mon poste » déprogrammée ? Cyril Hanouna dement*. 26/11/2019.

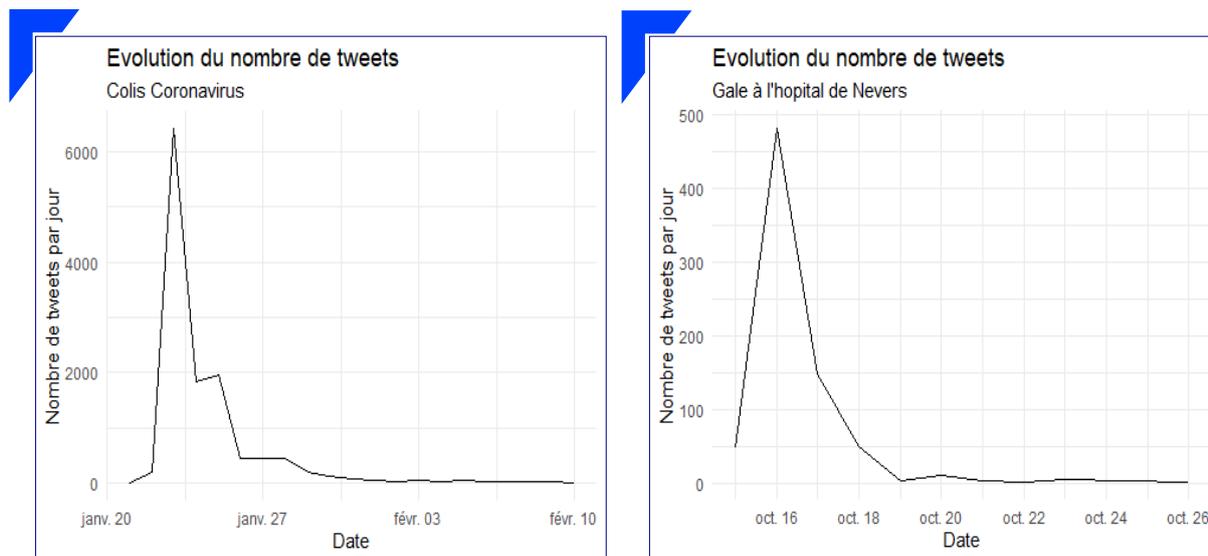


Le tableau suivant présente pour chaque fausse information sélectionnée le nombre de Tweets collectés et la période couverte par la collecte. Par exemple, la recherche de Tweets concernant la transmission du coronavirus par livraison de colis (mots clés « Colis et coronavirus ») a retourné 12 506 Tweets. Une proportion importante de ces Tweets sont des ReTweets, le nombre de Tweets différents est donc lui plus faible et s'élève à 1 438.

<b>THÈME</b>	<b>NOMBRE DE TWEETS COLLECTÉS</b>	<b>NOMBRE DE TWEETS DIFFÉRENTS COLLECTÉS</b>	<b>DURÉE DE CAPTATION (EN JOURS)<sup>70</sup></b>
<b>Activité de Greta Thunberg sur Facebook</b>	6 093	894	25
<b>Brevet sur le Coronavirus</b>	669	300	18
<b>Colis Coronavirus</b>	12 506	1 438	20
<b>Coronavirus à Argenteuil</b>	45	31	10
<b>Coronavirus à Perpignan</b>	50	23	4
<b>Gale à l'hôpital de Nevers</b>	745	60	11
<b>Gates et Coronavirus</b>	819	406	18
<b>Netflix</b>	835	339	8
<b>Panneau d'affichage - Limoges</b>	23 983	291	9
<b>Pollution de l'eau à Rouen</b>	101 734	4 470	51
<b>Propos de M. Blanquer</b>	3 100	630	15
<b>Rémunération des infirmières</b>	5 720	26	30
<b>TPMP</b>	156	29	24

La totalité des fausses informations étudiées ont la particularité de présenter une forte concentration des Tweets sur une période de temps très resserrée. Les graphiques suivants présentent l'exemple de la chronologie des Tweets pour deux des fausses informations étudiées : la transmission du coronavirus par colis et la quarantaine de l'hôpital de Nevers pour cause d'épidémie de gale (les graphiques des autres fausses informations étudiées sont reportés en annexe) :

<sup>70</sup> La durée de collecte n'a pas été équivalente pour toutes les thématiques. Cette colonne indique le nombre de jours pour lesquels des Tweets ont été collectés pour un thème donné.



Ce profil de propagation se retrouve de façon plus ou moins similaire pour la totalité des fausses informations étudiées. Il est cependant à relever que la collecte a ici porté sur une durée limitée et que, dans certains cas, la circulation des fausses informations a pu reprendre de l'ampleur au-delà de la période de collecte. La fausse information concernant l'existence d'un brevet sur le nouveau coronavirus a par exemple pris différentes formes à des périodes distinctes<sup>71</sup>.

Une hypothèse optimiste pour expliquer la forme prise par la courbe chronologique des Tweets serait que la fausse information commencerait à se propager puis qu'elle serait réfutée par des fact-checkers, dont les Tweets seraient par la suite largement retweetés, faisant disparaître la fausse information. Les sections suivantes permettront de vérifier si cette hypothèse est vérifiée ou non.

## 5.2. Méthodologie de qualification des Tweets collectés

La phase de collecte a été suivie par une phase de qualification du contenu des Tweets collectés. L'ampleur du nombre de ces derniers n'a pas permis d'envisager leur complète qualification. Seuls les 50 messages les plus retweetés pour chaque thématique ont été étudiés et qualifiés.

Dans la totalité des cas, cette approche restreinte aux messages les plus retweetés a permis de traiter une majorité de Tweets collectés (chaque Tweet collecté étant comptabilisé comme un Tweet, y compris s'il s'agit d'un ReTweet). Cette observation permet d'emblée de mettre en avant un phénomène de concentration de la conversation sur Twitter autour de certains Tweets très largement retweetés. Il a été relevé par exemple que la recherche de Tweets relative à la transmission du coronavirus par colis avait conduit à la collecte de 12 506 Tweets (originaux ou retweetés). La labellisation des seuls 50 Tweets les plus retweetés sur ce sujet a permis de labelliser 86 % des Tweets collectés.

<sup>71</sup> 20 Minutes. *Coronavirus : Non, le Covid-19 n'a pas été breveté par l'Institut Pasteur en 2004*. 18/03/2020.



Le tableau suivant propose une synthèse du pourcentage de Tweets qualifiés pour chacune des fausses informations étudiées :

<b>THÈMES</b>	<b>POURCENTAGE DE TWEETS QUALIFIÉS</b>
<b>Activité de Greta Thunberg sur Facebook</b>	82 %
<b>Brevet sur le Coronavirus</b>	62 %
<b>Colis Coronavirus</b>	86 %
<b>Coronavirus à Argenteuil</b>	100 %
<b>Coronavirus à Perpignan</b>	100 %
<b>Gale à l'hôpital de Nevers</b>	100 %
<b>Gates et Coronavirus</b>	55 %
<b>Netflix</b>	64 %
<b>Panneau d'affichage - Limoges</b>	99 %
<b>Pollution de l'eau à Rouen</b>	89 %
<b>Propos de M. Blanquer</b>	77 %
<b>Rémunération des infirmières</b>	100 %
<b>TPMP</b>	100 %

De façon plus précise, la catégorisation a consisté à classer les Tweets sélectionnés en catégorie « 1 » pour ceux contribuant à la propagation d'une fausse information telle qu'identifiée par des fact-checkers, en catégorie « 2 » pour les Tweets humoristiques, en catégorie « 3 » pour les Tweets qui contiennent une information vérifiée par les fact-checkers et en catégorie « 4 » lorsque le message n'est pas lié à la fausse information analysée et a été collecté par erreur car il contenait les mots clés recherchés (ces derniers Tweets ont alors été supprimés de la base qualifiée).

Cet exercice a recherché autant que possible à se baser sur les articles de fact-checkers et à éviter une classification subjective. Si dans la plupart des cas, la nature des Tweets a permis sans difficulté de les classer, cette opération a pu être parfois plus délicate, notamment dans le cas de Tweets adoptant un ton humoristique mais partageant une fausse information<sup>72</sup>.

<sup>72</sup> Plusieurs Tweets ayant rencontré un certain succès ont été catégorisés comme contribuant à la propagation de fausses informations en ce qu'ils les abordaient avec humour et sans les remettre en question. C'est le cas d'une reprise de la séquence de la série télévisée d'animation *Les Simpson* qui aurait « prédit » que les colis en provenance de Chine pourraient être vecteurs du virus, reprenant une fausse information qui a circulé un temps sur les réseaux sociaux mais en n'y apportant pas de correction.



### 5.3. Analyse de la base de Tweets qualifiée

Le tableau ci-dessous résume le nombre de Tweets qualifiés pour chaque thématique (pour rappel chaque ReTweet compte comme un Tweet) ainsi que la répartition de ces Tweets selon leur catégorisation :

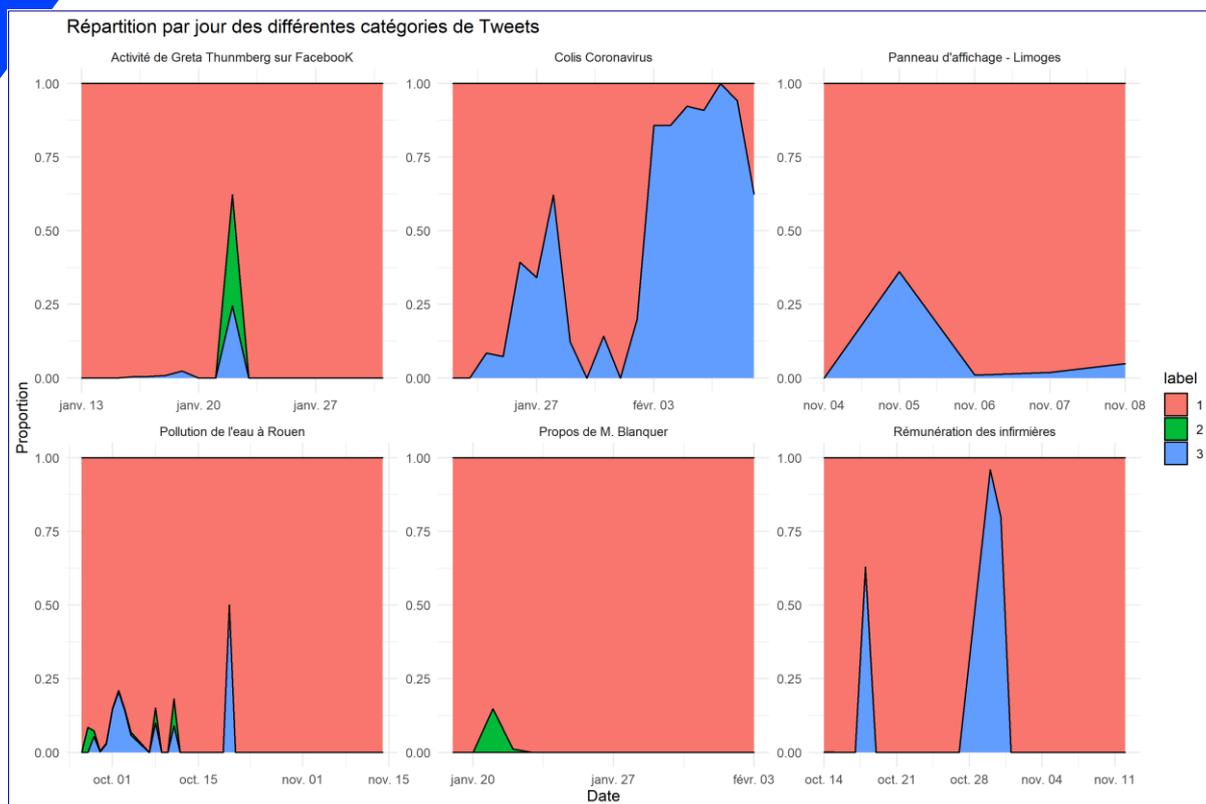
THÈME	NOMBRE DE TWEETS QUALIFIÉS	POURCENTAGE DE TWEETS QUALIFIÉS « 1 »	POURCENTAGE DE TWEETS QUALIFIÉS « 2 »	POURCENTAGE DE TWEETS QUALIFIÉS « 3 »
<b>Activité de Greta Thunberg sur Facebook</b>	4 967	99,19 %	0,34 %	0,46 %
<b>Brevet sur le Coronavirus</b>	416	31,73 %	0,00 %	68,27 %
<b>Colis Coronavirus</b>	10 783	93,46 %	0,00 %	6,54 %
<b>Coronavirus à Argenteuil</b>	45	42,22 %	6,67 %	51,11 %
<b>Coronavirus à Perpignan</b>	50	48,00 %	0,00 %	52,00 %
<b>Gale à l'hôpital de Nevers</b>	745	86,04 %	2,28 %	11,68 %
<b>Gates et Coronavirus</b>	453	24,72 %	0,00 %	75,28 %
<b>Netflix</b>	538	15,06 %	0,00 %	84,94 %
<b>Panneau d'affichage - Limoges</b>	23 708	98,62 %	0,00 %	1,38 %
<b>Pollution de l'eau à Rouen</b>	90 048	95,24 %	1,16 %	3,60 %
<b>Propos de M. Blanquer</b>	2 393	99,16 %	0,84 %	0,00 %
<b>Rémunération des infirmières</b>	5 720	98,72 %	0,02 %	1,26 %
<b>TPMP</b>	156	4,49 %	0,00 %	95,51 %

Indépendamment de l'importance de l'information elle-même, il apparaît que les thèmes se répartissent à peu près équitablement entre ceux pour lesquels la fausse information apparaît de façon majoritaire et ceux pour lesquels l'information vérifiée est au contraire prépondérante. Cependant, pour les thèmes qui présentent une volumétrie plus élevée (plus de 1 000 Tweets), la fausse information est dans tous les cas majoritaire et elle est même le plus souvent ultra-majoritaire. Les Tweets partageant une information vérifiée sont peu susceptibles d'atteindre des niveaux de viralité très élevés, alors que les sujets qui connaissent une viralité plus significative sont ceux qui ont atteint cette viralité grâce à un partage massif de fausses informations.

Si l'analyse de la volumétrie des Tweets selon leur nature (partage d'information fausse ou avérée) fournit des enseignements intéressants, il peut également être utile d'étudier la chronologie du partage des différentes catégories de Tweets. Comme souligné supra, une hypothèse - optimiste - consisterait à considérer que cette chronologie pourrait être la suivante.

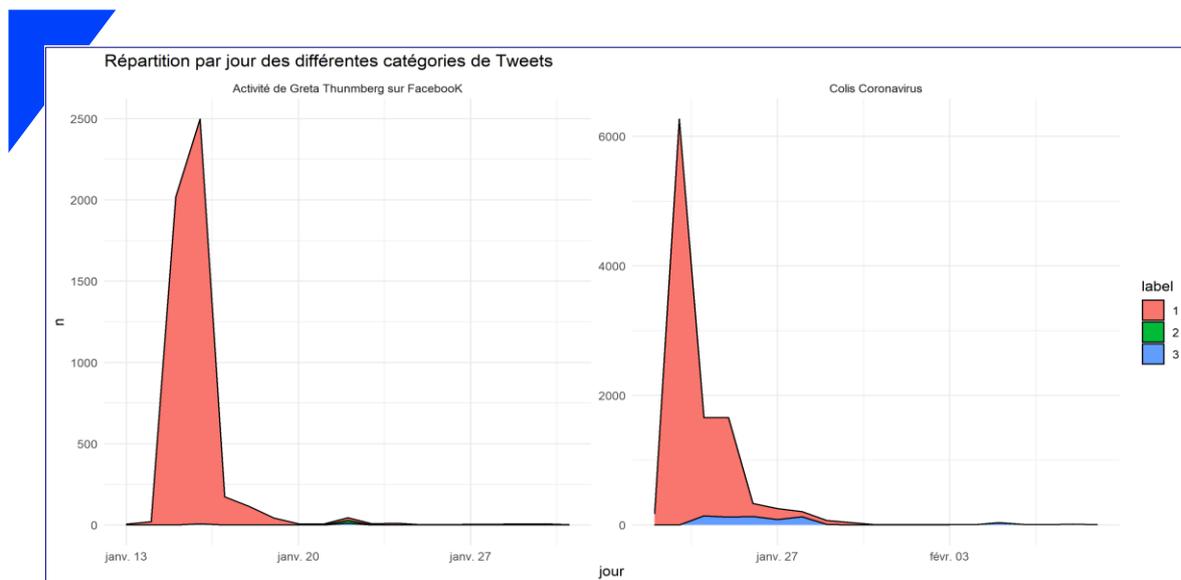


Lorsqu'un sujet apparaît sur Twitter sous la forme d'une fausse information, des Tweets viennent corriger cette fausse information et la fausse information est alors délaissée tandis que les Tweets reprenant l'information vérifiée deviennent à leur tour majoritaires. Les chronologies observées, variables d'un sujet à l'autre, ne correspondent cependant pas à cette hypothèse. Les graphiques ci-dessous présentent ainsi l'évolution de la répartition des labels dans le temps pour les thématiques ayant fait l'objet de la diffusion la plus importante, soit plus de 1000 Tweets (les graphiques correspondant aux autres thématiques sont renvoyés en annexe) :



Il apparaît ainsi que c'est uniquement pour le thème relatif à la transmission du coronavirus par colis que l'information vérifiée devient finalement dominante, dans les autres cas, l'information vérifiée connaît un bref pic de croissance avant de céder à nouveau la place aux fausses informations.

Ces graphiques dépeignent par ailleurs une vision sans doute trop optimiste encore de la chronologie de la diffusion des fausses informations. En effet, ils mettent en évidence l'évolution de la répartition des catégories de Tweets en faisant abstraction de l'évolution de leur volumétrie en valeur absolue. Les graphiques suivants présentent l'évolution du nombre de Tweets de chaque catégorie (et non plus l'évolution de la répartition) pour les exemples de la transmission du coronavirus par colis et du compte Facebook de Greta Thunberg. Les autres graphiques sont reportés en annexe.



Pour ce qui concerne par exemple la transmission du coronavirus par colis, si les vraies informations ont fini par devenir majoritaires sur ce sujet c'est uniquement à partir du moment où le sujet a quasiment cessé d'être discuté sur Twitter. De la même façon, pour le cas du compte Facebook de Greta Thunberg, le pic dans la proportion d'information vérifiée correspond à une période pour laquelle ce sujet n'était plus que marginalement diffusé sur Twitter.

#### Synthèse :

Toutes les fausses informations étudiées présentent une forte concentration des Tweets sur une période de temps très réduite. Elles se partagent par ailleurs équitablement entre les thèmes pour lesquels la fausse information est partagée de façon majoritaire et les thèmes pour lesquels c'est au contraire l'information vérifiée qui est majoritairement partagée. Cependant, pour les thèmes qui présentent les volumes de Tweets les plus élevés, la fausse information est dans tous les cas majoritaire. L'étude de la chronologie de la diffusion des fausses informations montre que, contrairement à ce qui aurait pu être attendu ou espéré, les « vraies » informations ne chassent pas les fausses. Pour les sujets sur lesquels les fausses informations sont majoritaires, la discussion sur ces sujets s'éteint rapidement sur Twitter sans que l'information vérifiée n'ait jamais atteint une position majoritaire.

### 5.4. Analyse de l'audience potentielle de chaque catégorie de Tweet

Les analyses précédentes ont permis d'apprécier l'importance volumétrique du nombre de Tweets portant sur des informations fausses ou avérées. Ces analyses sont cependant insuffisantes pour apprécier la visibilité des différentes catégories de Tweets puisqu'elles ne tiennent pas compte du nombre d'abonnés de chaque compte ayant partagé une information vraie ou fausse. Le tableau suivant présente pour chaque thématique analysée la somme du nombre d'abonnés de chaque compte ayant émis un Tweet sur cette thématique dans la base labellisée de l'étude. Ce nombre d'abonnés est ensuite réparti entre les comptes ayant émis un Tweet catégorisé en 1 (fausse information), en 2 (humour) et en 3 (information vérifiée). Il est à noter que cette analyse est soumise à une forte marge d'erreur : si un compte disposant d'une



audience importante publie une information vérifiée sur la thématique étudiée mais que ce Tweet ne fait pas l'objet de nombreux ReTweets, il ne sera pas retenu pour notre labellisation et ne sera donc pas comptabilisé.

THÈME	SOMME DU NOMBRE D'ABONNÉS DES COMPTES AYANT PARTAGÉ UN TWEET SUR LE SUJET <sup>73</sup>	POURCENTAGE DU NOMBRE D'ABONNÉS EXPOSÉS A UN TWEET QUALIFIÉ 1	POURCENTAGE DU NOMBRE D'ABONNÉS EXPOSÉS A UN TWEET QUALIFIÉ 2	POURCENTAGE DU NOMBRE D'ABONNÉS EXPOSÉS A UN TWEET QUALIFIÉ 3
Activité de Greta Thunberg sur Facebook	6 541 160	98,78 %	0,86 %	0,35 %
Brevet sur le Coronavirus	6 431 347	4,41 %	0,00 %	95,59 %
Colis Coronavirus	6 396 890	82,99 %	0,00 %	17,01 %
Coronavirus à Argenteuil	27 914	29,03 %	4,95 %	66,02 %
Coronavirus à Perpignan	3 959 948	0,65 %	0,00 %	99,35 %
Gale à l'hôpital de Nevers	5 672 059	11,95 %	0,55 %	87,50 %
Gates et Coronavirus	761 422	38,83 %	0,00 %	61,17 %
Netflix	1 028 193	46,37 %	0,00 %	53,63 %
Panneau d'affichage - Limoges	11 244 122	97,34 %	0,00 %	2,66 %
Pollution de l'eau à Rouen	56 654 501	80,88 %	1,26 %	17,86 %
Propos de M. Blanquer	2 618 692	99,82 %	0,18 %	0,00 %
Rémunération des infirmières	4 821 826	67,90 %	0,00 %	32,10 %
TPMP	264 733	0,35 %	0,00 %	99,65 %

On s'attendrait à ce que la part d'audience des informations fiables soit supérieure à celle des fausses informations, puisque les comptes d'information fiables qui sont susceptibles de publier des Tweets partageant l'information avérée disposent d'audiences fréquemment élevées. Or si, pour certaines thématiques, la part d'audience des sites partageant des informations vérifiées est effectivement plus élevée que la part volumétrique de ces Tweets, certaines fausses informations restent majoritaires en termes de visibilité (compte Facebook de Greta Thunberg, transmission du coronavirus par colis, etc.).

<sup>73</sup> Relevons que pour chaque thématique un même abonné peut être comptabilisé plusieurs fois si il est abonné à différents comptes ayant partagé un Tweet sur cette thématique.



### Synthèse :

Une analyse en termes d'audience potentielle des informations fausses et vérifiées fait apparaître la place plus importante de l'information vérifiée par rapport à une analyse portant uniquement sur le volume de Tweets. Cependant, pour certains thèmes, la fausse information reste malgré tout plus visible que l'information vérifiée.

## 5.5. Analyse des délais de correction

Le tableau suivant représente les délais entre la première fausse information enregistrée dans la base labellisée et la première information vérifiée sur le même thème. Encore une fois, cette analyse qui porte sur la base labellisée ne considère que les Tweets les plus retweetés et des informations avérées peu repartagées ne sont pas prises en compte, ce qui peut conduire à surestimer le délai de correction.

THÈME	DATE DE LA PREMIÈRE INFORMATION FAUSSE	DATE DE LA PREMIÈRE INFORMATION AVÉRÉE	DÉLAIS ENTRE FAUSSE INFORMATION ET INFORMATION AVÉRÉE
Activité de Greta Thunberg sur Facebook	2020-01-13 22:03:31	2020-01-16 14:50:45	64.79 heures
Brevet sur le Coronavirus	2020-01-22 05:48:34	2020-01-24 16:42:18	58.90 heures
Colis Coronavirus	2020-01-22 10:19:28	2020-01-24 14:35:00	52.26 heures
Coronavirus à Argenteuil	2020-01-26 13:27:35	2020-01-26 23:09:44	9.70 heures
Coronavirus à Perpignan	2020-01-26 17:50:44	2020-01-26 13:46:18	-4.07 heures
Gale à l'hôpital de Nevers	2019-10-15 21:53:39	2019-10-16 15:44:13	17.84 heures
Gates et Coronavirus	2020-01-24 19:59:38	2020-01-29 19:20:54	119.35 heures
Netflix	2019-10-21 08:32:33	2019-10-21 10:07:26	1.58 heures
Panneau d'affichage - Limoges	2019-11-04 22:53:48	2019-11-05 09:17:21	10.39 heures
Pollution de l'eau à Rouen	2019-09-26 06:34:20	2019-09-28 05:26:15	46.87 heures
Propos de M. Blanquer	2020-01-19 15:00:26	<sup>74</sup>	-
Rémunération des infirmières	2019-10-14 09:05:10	2019-10-15 00:35:06	15.50 heures
TPMP	2019-11-25 14:49:35	2019-11-25 15:54:33	1.08 heures

<sup>74</sup> Absence de Tweet corrigeant la fausse information dans la base.



Les fausses informations apparaissent sur le réseau Twitter de manière assez naturelle le plus souvent avant leur réfutation (à l'exception du cas de Coronavirus à Perpignan<sup>75</sup>). Dans certains cas, la réfutation apparaît très rapidement comme c'est le cas pour l'annulation de l'émission TPMP. Dans d'autres cas, les délais apparaissent très importants, notamment si l'on tient compte de la durée de vie parfois très limitée des discussions sur une thématique donnée. Ces éléments ainsi que les éléments présentés *supra* illustrent la difficulté du travail des fact-checkers. S'ils veulent influencer la diffusion des fausses informations sur Twitter, ils doivent réagir extrêmement vite puisqu'une fausse information sera le plus souvent partagée sur une durée très réduite alors que dans le même temps, ils sont soumis aux contraintes classiques d'un travail journalistique de vérification des sources qui ne peut être instantané et leur travail peut reposer sur l'identification de fausses informations ayant déjà atteint une certaine popularité en ligne.

De ce point de vue, les partenariats mis en œuvre par certains réseaux sociaux avec des fact-checkers<sup>76</sup> afin d'afficher des corrections à proximité des fausses informations détectées pour en limiter la viralité seront d'autant plus susceptibles de produire un effet important que ces corrections seront ajoutées très rapidement après la publication du Tweet.

## 5.6. Analyse des messages les plus retweetés

Afin de mieux comprendre la nature des Tweets propageant ou réfutant de fausses informations, une analyse qualitative a été menée, pour chaque thématique, sur les deux Tweets les plus partagés contenant des informations vérifiées et les deux Tweets les plus partagés reprenant des informations non vérifiées<sup>77</sup>, qu'ils soient issus ou non de comptes de médias. Il en résulte un corpus de 48 Tweets, à parts égales entre messages erronés et correctifs (24 Tweets à chaque fois).

Parmi ces 24 Tweets, les **principaux messages erronés** sont basés sur de fausses informations de différents types : il peut soit s'agir de faits non avérés voire tronqués, soit d'une absence de contextualisation qui en modifie le sens. Ils **expriment une critique directe (16), une forme de panique (3) ou relèvent du registre de l'humour (2)**<sup>78</sup>. Dans un seul cas, décrit à la fin de cette section, la fausse information a été publiée par des médias ayant le statut « vérifié » sur Twitter (2).

---

<sup>75</sup> Il semble que pour ce cas, la fausse information soit tout d'abord apparue sur le réseau Facebook, elle a ensuite été démentie par la presse locale et ce démenti a été partagé par le titre de presse sur Twitter : <https://www.lindependant.fr/2020/01/26/coronavirus-non-aucun-cas-na-ete-detec-te-a-perpignan,8688548.php>. La fausse information a ensuite fait son apparition sur Twitter après l'information vérifiée.

<sup>76</sup> Par exemple pour le cas de Facebook : <https://www.facebook.com/business/help/182222309230722>

<sup>77</sup> Le corpus devrait se composer de 52 Tweets (4 par thème avec 13 thèmes à l'étude) labellisés pour moitié comme étant faux (1) et pour l'autre moitié comme constituant des vérifications ou partage de vérification journalistique (3) parmi les plus retweetés, mais pour deux thèmes étudiés, aucun Tweet labellisé 1 et 3 respectivement n'a été suffisamment retweeté pour être inclus dans notre analyse (n = 48).

<sup>78</sup> Un Tweet ne fait que reprendre un article ayant ensuite été qualifié de faux par les cellules de vérification de l'information de Libération et de franceinfo ; sans autre texte que le titre et le lien. Il n'est donc pas commenté dans cette analyse. On retrouve ainsi bien n= 24.



Les accusations directes sont majoritaires dans le corpus étudié. Sont ainsi pris pour cibles des membres du gouvernement (« *#Blanquer qui trouve que tout va bien puisque 99,9 % des profs le soutiennent* »), des élus (« *L'hôpital de #Nevers en quarantaine suite à une épidémie de gale due à l'hospitalisation de #migrants Omerta du maire #LREM de [...]* ») et des personnalités (« *Un bug sur Facebook révèle que les messages postés par Greta Thunberg sont écrits par son père... Seuls les idiots pouvaient penser que cette décrochée scolaire pouvait écrire ses messages, qu'elle passe son bac avant de donner des leçons au monde entier* »). Des mentions plus générales (« *on* », « *continuez* ») semblent accuser les autorités de mentir, sans les nommer : c'est par exemple le cas des Tweets liés à l'incendie de l'usine Lubrizol, à Rouen (« *Heureusement qu'on nous a dit que l'eau est potable à Rouen* »). D'autres le font de manière détournée, en utilisant par exemple l'expression « *hasard ?* ». Ces accusations se basent majoritairement sur des témoignages non-recoupés, sous forme d'images ou d'écrits, ou bien sur des faits qui ont été qualifiés d'inexactes ou faux par des journalistes.

La crise sanitaire mondiale liée à l'épidémie de la Covid-19 semble quant à elle particulièrement favoriser le partage de Tweets exprimant un sentiment de panique. Ceux-ci prennent deux formes principales : soit ils sont très émotionnels avec l'utilisation de majuscules alliée à des formes interrogatives et des émoticônes négatifs (« *La fondation Gates a-t-elle vraiment financé « le groupe qui détient le brevet » du coronavirus ? Je crains que ça soit \* LA VÉRITÉ \* 😞😞😞* »), soit ils font une compilation de chiffres et de faits à l'appui de leur propos (« *Brevet américain pour le brevet sur le coronavirus (brevet n ° 10 130 701 délivré le 20 novembre 2018)* ») bien que ceux-ci aient été qualifiés de faux par des vérificateurs journalistiques.

Par ailleurs, deux Tweets ont été labellisés comme contribuant à la propagation de fausses informations en ce qu'ils les abordaient avec humour et sans les remettre en question. Des quatre catégories de cette analyse, ce sont ceux qui ont la plus forte moyenne de repartage (29 511 ReTweets en tout, dont 21 600 pour le plus partagé des deux). C'est le cas d'une reprise de la séquence de la série télévisée d'animation *Les Simpson* qui aurait « *prédit* » que les colis en provenance de Chine pourraient être vecteurs du virus, reprenant une fausse information qui a circulé un temps sur les réseaux sociaux mais en n'y apportant pas de correction.

Enfin, les deux seuls Tweets publiés par des médias et propageant une fausse information abordent la même thématique : Netflix aurait annoncé le renforcement de la lutte contre le partage de compte. Les propos du directeur des produits du service ont été repris dans l'un d'entre eux à l'appui de cette affirmation, bien qu'il soit fait mention que cela serait « *sous-entend[re]* ». Des liens vers les sites de ces médias accompagnent ces Tweets. Il semble ici que ces deux médias ont repris une information sans la recouper, entraînant par la suite des démentis publiés par des confrères (« *personne n'a dit ça* ») tweetés par des consœurs de *Numerama*. Ils totalisent seulement 45 Tweets à eux deux.



Concernant les **Tweets correctifs**, la moitié sont issus de comptes « certifiés »<sup>79</sup> (12 se partageant en 11 médias dont des cellules spécialisées dans la vérification des informations et 1 institutionnel) et l'autre moitié de comptes non certifiés. Pour certains, il s'agit de comptes se qualifiant de médias (4), tandis que d'autres semblent davantage relever de comptes personnels (8).

Les premiers alternent entre reprises des termes de la fausse information sous forme interrogative avant d'inclure un lien vers leur article de vérification (« *L'hôpital de Nevers est-il confiné à cause d'une épidémie de gale amenée par des migrants ?* », Franceinfo, 17 octobre 2019). D'autres se rapprochent davantage des formats spécifiques à Twitter, par exemple en étant plus diserts et en donnant directement des éléments détaillés dans leurs Tweets (« *Le message affiché sur le panneau numérique rue Jean-Jaurès est un message factice qui a été affiché hier de 13h à 16h et qui sera renouvelé aujourd'hui dans les mêmes horaires pour les besoins d'un film long métrage actuellement en tournage dans le centre-ville de Limoges* »), Ville de Limoges, 5 novembre 2019) ou en créant un fil composé de plusieurs Tweets – thread (« *[A dérouler] Explosions, oiseaux morts, qualité de l'air et eau potable...Les inquiétudes légitimes sur les conséquences de l'incendie de l'usine Lubrizol à Rouen ont servi de terreau à de nombreuses manipulations. Voici nos conseils pour vérifier certaines affirmations* ↓ », AFP Factuel 📍, 1<sup>er</sup> octobre 2019). Ces Tweets concentrent par ailleurs un grand nombre de ReTweets (1 653 en tout, 138 par Tweet en moyenne), pour moitié (810 reprises) issus de la seule publication de l'AFP Factuel 📍 qui vient d'être citée.

Du côté des correctifs apportés par des comptes non certifiés, l'analyse qualitative des Tweets correctifs qu'ils ont publiés met en exergue certaines caractéristiques communes. Ainsi, plusieurs citent des médias dits historiques ou des sources faisant autorité (« *Le CSA vient de démentir en parlant d'une intox ainsi que @Cyrilhanouna à l'instant en direct à l'antenne.* ») ou déconstruisent les origines de la fausse information (« *Vous avez copié le journal parodique @Nordpresse (cousin BE du #Gorafi)* »). Les Tweets qui apparaissent issus de comptes personnels s'attachent de leur côté soit à faire des récapitulatifs concernant les fausses informations qui circulent sur un sujet donné (« *Quelques théories du net sur le Coronavirus: [...]*») soit, dans un cas, à critiquer le correctif apporté par un média vérifié (« *France Info qui reprend mon Tweet en affichant pseudo+photo sans m'avertir [...]*»). Ces Tweets concentrent moins de ReTweets (1 179 en tout, 98 par Tweet en moyenne), le plus repartagé étant issu d'un compte apparaissant comme un média (Brèves de presse, 610 reprises).

Au global, l'analyse des Tweets les plus partagés en matière de désinformation permet de mettre en évidence le rôle très important joué par les médias traditionnels en matière de lutte contre la désinformation. Dans la majorité des cas, ce sont plus précisément les cellules spécialisées dans la vérification des informations qui parviennent à obtenir la visibilité la plus importante. Ce succès ne reste cependant que partiel puisque les Tweets diffusés par ces vérificateurs ne parviennent que rarement à atteindre le degré de viralité des Tweets partageant de fausses informations.

---

<sup>79</sup> Twitter appose à côté du nom de certains comptes un badge constitué d'une encoche blanche sur fond bleu devant permettre « d'assurer aux utilisateurs l'authenticité d'un compte d'intérêt public » par exemple dans le domaine de la politique, des médias et des pouvoirs publics. Twitter indique cependant que ce badge « n'implique en aucun cas une recommandation » de la part du réseau social. Twitter, A propos des comptes certifiés, Centre d'assistance (page consultée le 30 avril 2020). URL : <https://help.twitter.com/fr/managing-your-account/about-twitter-verified-accounts>



### **Synthèse :**

L'analyse de quelques fausses informations permet de dégager plusieurs caractéristiques intéressantes. D'une part, ces contenus, lorsqu'ils n'émanent pas de fact-checkers, sont souvent ardu à qualifier : le statut de l'humour pose ainsi question dans l'appréhension des Tweets. Ces fausses informations sont par ailleurs majoritairement le support d'une critique adressée aux autorités ou expriment une forme de panique, par exemple sur des sujets sanitaires sensibles. Elles sont appuyées par des informations non recoupées ou qualifiées de fausses par les journalistes, qui apportent de leur côté des correctifs empruntant les codes spécifiques aux réseaux sociaux. Le faible degré d'engagement (ReTweet, commentaires) de ces publications par rapport aux comptes susceptibles de diffuser de fausses informations semble faire état d'un succès relatif de cette méthode. Il faut néanmoins relever que dans plusieurs cas, le correctif a pu être directement apporté par des internautes, en s'appuyant sur les éléments de ces articles de médias, montrant que ce travail rencontre son public.



## 6. Conclusion

Cette étude a permis de montrer que les comptes d'information les moins fiables sur Twitter possèdent un nombre d'abonnés significativement plus faible que la majorité des comptes d'information fiables. Cependant, les comptes connus pour partager de fausses informations font jeu égal, en termes de ReTweets, avec les comptes fiables. Les abonnés des comptes non fiables ont en effet une propension beaucoup plus élevée à contribuer à la diffusion des informations diffusées sur ces comptes que les abonnés des comptes fiables, avec un nombre de ReTweets par abonnés 10 à 20 fois plus élevé selon l'indicateur retenu.

L'analyse a également montré que les comptes non fiables privilégient largement des thématiques d'actualité et liées à des sujets clivants. Ainsi les Tweets traitant de politique, d'immigration, de santé, de religion et de terrorisme représentent plus de la moitié du corpus étudié. L'analyse quantitative montre la surreprésentation de termes liés la délinquance, l'immigration, Israël et la Palestine, la pédophilie, l'Islam ou encore la franc-maçonnerie. Ces comptes utilisent massivement des contenus non textuels (images, vidéos, liens, etc.) et renvoient fréquemment aux sites internet liés aux comptes mais également à des sources issues de médias traditionnels.

Les individus suivant des comptes non fiables et aucun fiable sont minoritaires sur Twitter. En effet, moins de 20 % des abonnés à ces comptes n'en suivent aucun fiable. Une analyse en termes de graphe fait par ailleurs ressortir que de nombreux comptes peu fiables possèdent une proportion significative d'abonnés communs ou se suivent mutuellement sur Twitter.

Les comptes de cellules journalistiques spécialisées dans la vérification de l'information sont en moyenne davantage suivis que les comptes catégorisés comme non fiables, mais génèrent moins d'interactions sur leurs publications que ces derniers. En termes de thématiques suscitant le plus d'engagement, seules la politique et la religion sont communes entre ces deux types de comptes.

L'étude présente enfin une analyse de Tweets se rapportant à des fausses informations identifiées comme telles par des journalistes. Toutes les fausses informations étudiées présentent une forte concentration de Tweets sur une période de temps très réduite. Les fausses informations étudiées se partagent par ailleurs équitablement entre les thèmes pour lesquels la fausse information est partagée de façon majoritaire et ceux pour lesquels c'est au contraire l'information vérifiée qui est majoritairement partagée. Cependant, pour les thèmes qui présentent le volume de Tweets le plus élevé, la fausse information est dans tous les cas majoritaire. L'étude de la chronologie de la diffusion des fausses informations montre que les « vraies » informations ne chassent pas les fausses. Pour les sujets sur lesquels les fausses informations sont majoritaires, la discussion sur ces sujets s'éteint rapidement sur Twitter sans que l'information vérifiée n'ait jamais occupé une place majoritaire.

L'analyse des Tweets ayant témoigné de la viralité la plus importante dans le corpus étudié permet enfin de faire ressortir plusieurs traits intéressants. Les fausses informations sont majoritairement le support d'une critique adressée aux autorités ou expriment une forme de panique, par exemple sur des sujets sanitaires sensibles. Elles s'appuient sur des informations non recoupées ou qualifiées de fausses par les journalistes, qui apportent de leur côté des correctifs empruntant les codes spécifiques aux réseaux sociaux. Ces correctifs récoltent



## La propagation des fausses informations sur les réseaux sociaux Étude du service Twitter

---

néanmoins un faible degré d'engagement (ReTweet, commentaires) par rapport aux comptes susceptibles de diffuser de fausses informations.

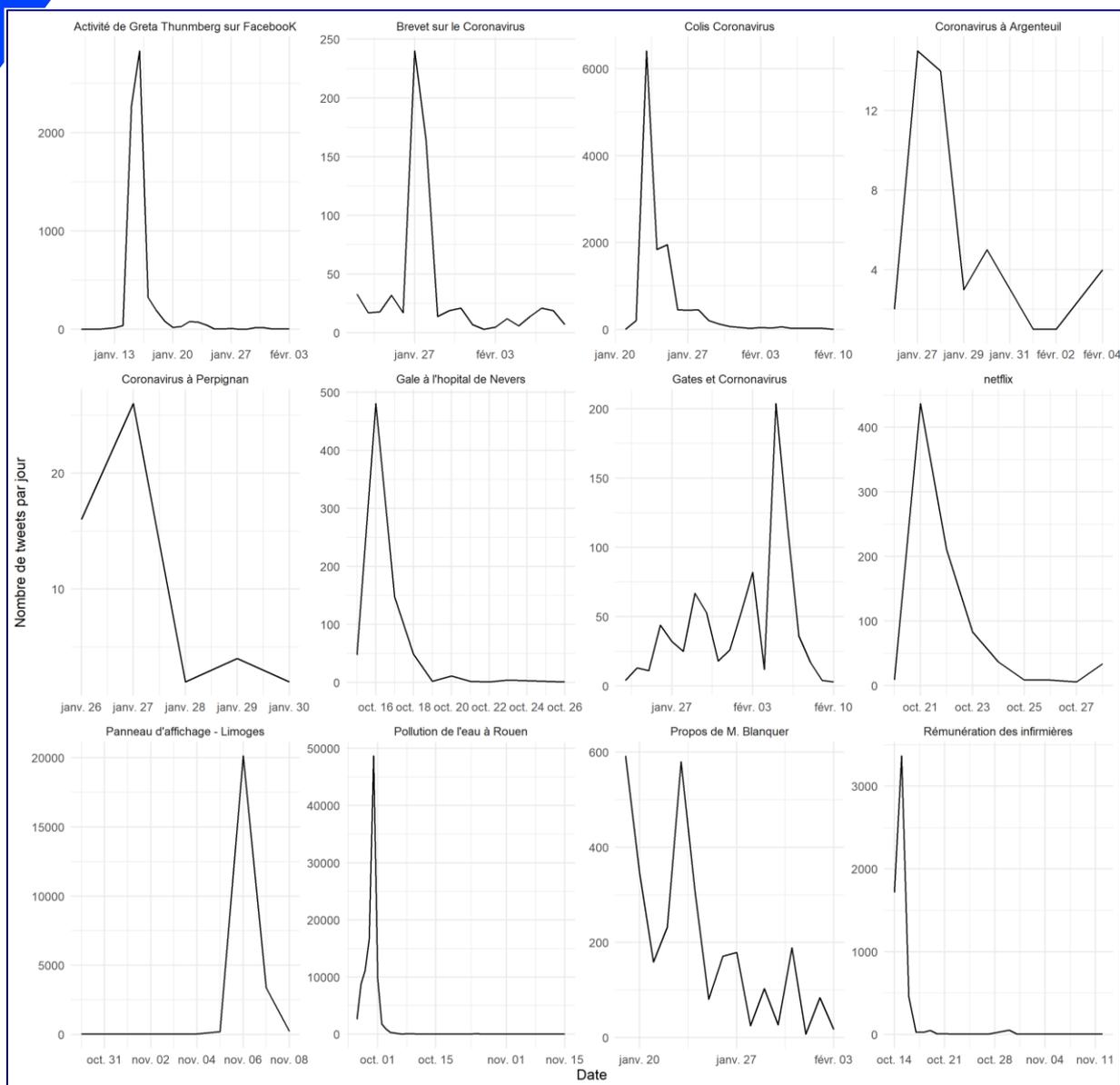
Si cette étude s'est attachée à étudier la viralité des fausses informations sur un réseau social, le contexte de sa publication rappelle avec force que les réseaux sociaux sont susceptibles de faciliter les partages d'autres formes de contenus préjudiciables, tels ceux appelant à la haine et à la violence. La compréhension des modalités de propagation des différentes formes de contenus sur les plateformes restera donc un enjeu central pour le Conseil.



## 7. Annexes

### 7.1. Chronologie de diffusion des Tweets

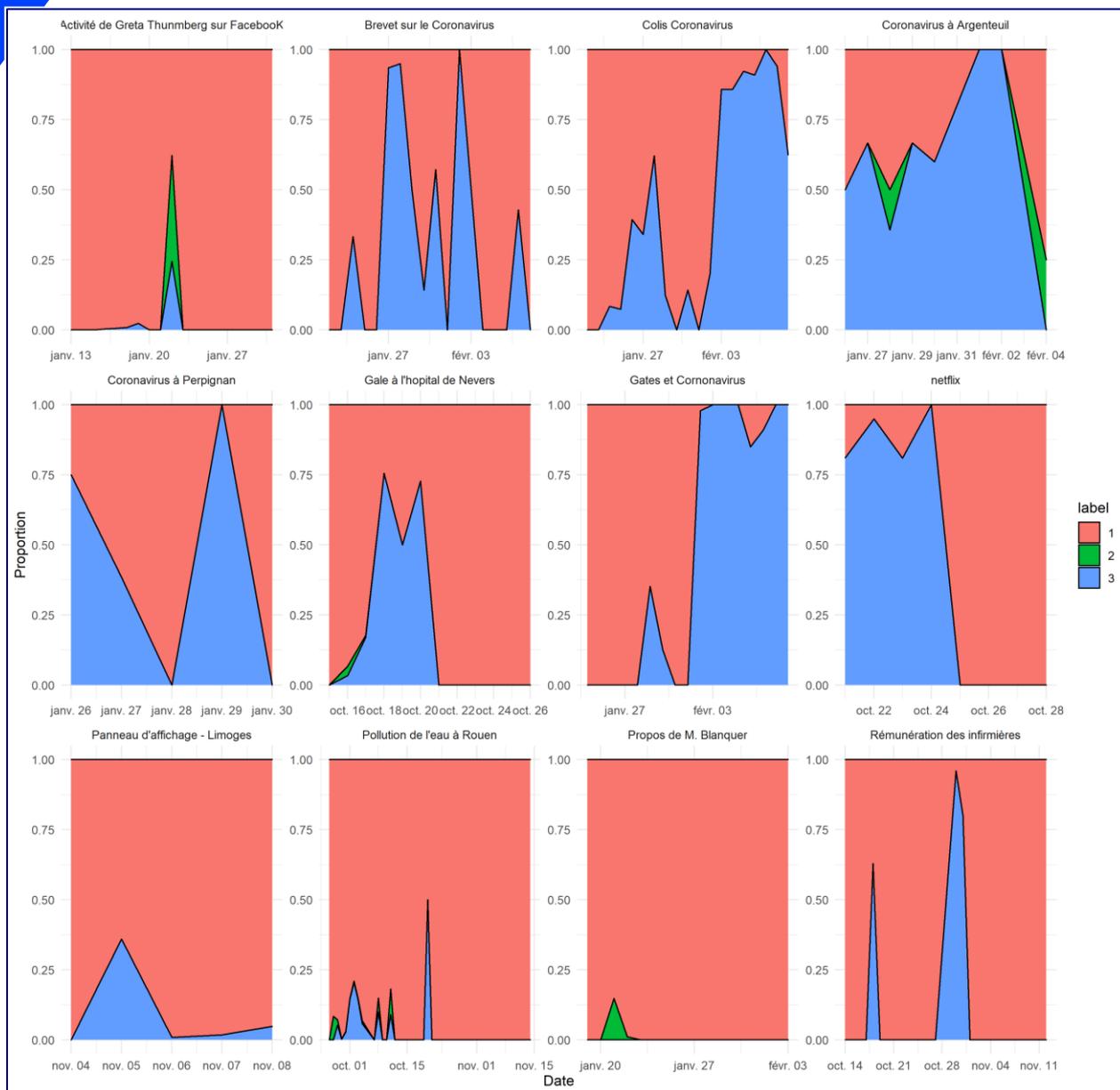
#### Évolution du nombre de Tweets





## 7.2. Évolution de la répartition des Tweets par catégorie

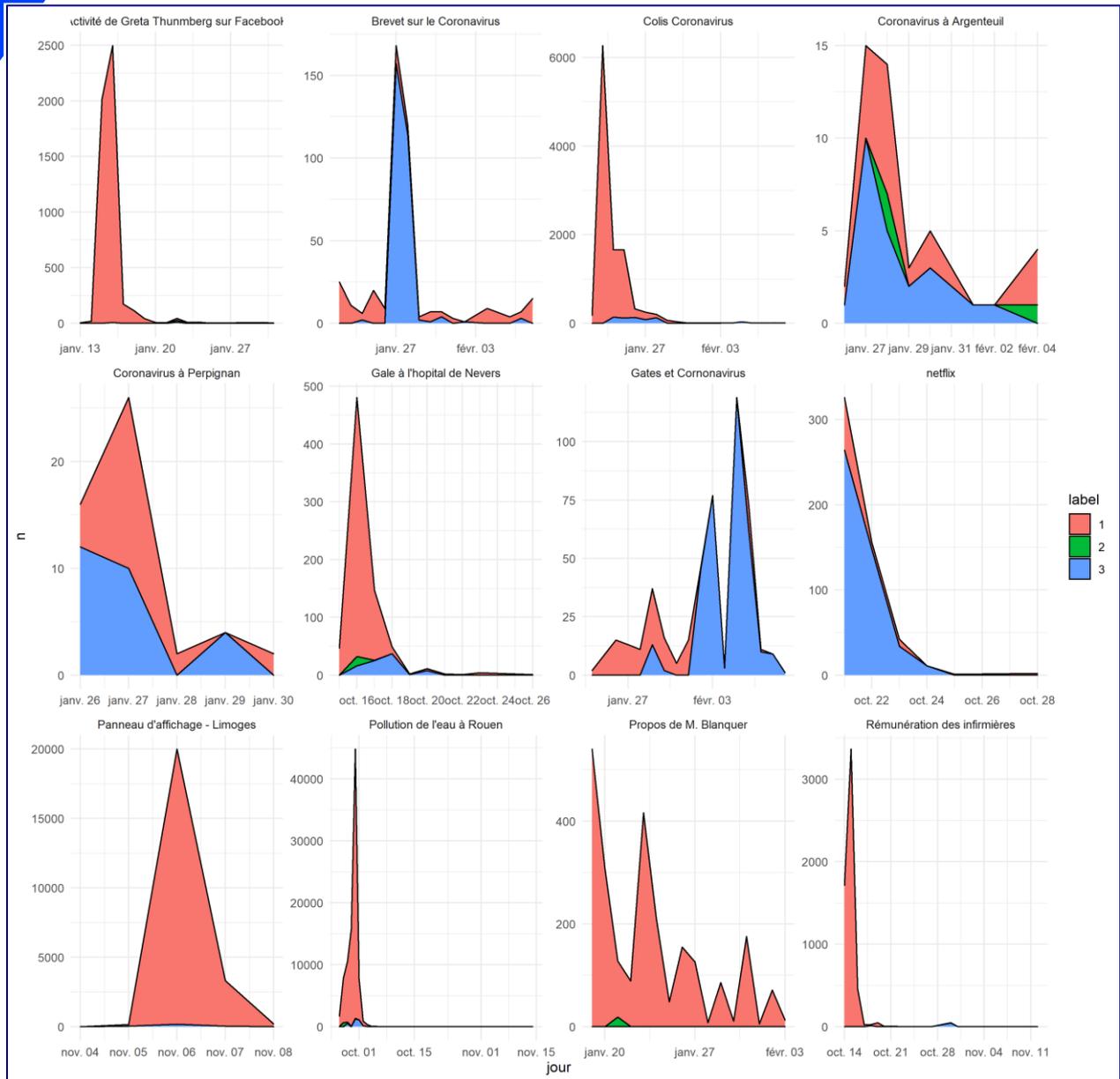
Répartition par jour des différentes catégories de Tweets





### 7.3. Évolution de la volumétrie des Tweets par catégorie

Répartition par jour des différentes catégories de Tweets





## 8. Bibliographie

### Sources juridiques

*Loi n° 2018-1202 du 22 décembre 2018 relative à la lutte contre la manipulation de l'information.* Journal officiel de la République française, 23/12/2018. URL : <https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000037847559&categorieLien=id> .

*Recommandation n° 2019- 03 du 15 mai 2019 du Conseil supérieur de l'audiovisuel aux opérateurs de plateforme en ligne dans le cadre du devoir de coopération en matière de lutte contre la diffusion de fausses informations.* Journal officiel de la République française, 17/5/2019. URL : <https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000038480745&categorieLien=id>.

*Décret n° 2019-297 du 10 avril 2019 relatif aux obligations d'information des opérateurs de plateforme en ligne assurant la promotion de contenus d'information se rattachant à un débat d'intérêt général.* Journal officiel de la République française, 11/4/2019. URL : <https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000038359165&categorieLien=id> ..

*The Digital Services Act package.* (Commission européenne), URL : <https://ec.europa.eu/digital-single-market/en/digital-services-act-package>. Consulté le : 11/6/2020.

### Littérature scientifique

Hunt Allcott, Matthew Gentzkow. *Social media and fake news in the 2016 election.* Journal of Economic Perspectives 2: 1–28. 2017.

Dimitrios Bountouridis et al. *Annotating credibility: Identifying and mitigating bias in credibility datasets.* 25/7/2019.

Peter Burger, Soeradj Kanhai, Alexander Pleijter, Suzan Verberne. *The Reach of Commercially Motivated Junk News on Facebook.* PLOS ONE 14(8): e0220446. 1/8/2019.

Cardon, Dominique, Ooghe-Tabanou, Benjamin, Plique, Guillaume, Cointet, Jean-Philippe. *Les nouveaux circuits de l'information numérique.* SciencesPo Médialab. 11/6/2019. URL : <https://medialab.sciencespo.fr/en/news/les-nouveaux-circuits-de-linformation-numerique>. Consulté le : 28/10/2019.

Ted Dunning. *Accurate Methods for the Statistics of Surprise and Coincidence.* Computational Linguistics 19(1): 61–74. 1993.



Emilio Ferrara. *Disinformation and social bot operations in the run up to the 2017 French presidential election*. First Monday 22(8). 31/7/2017. URL : <http://journals.uic.edu/ojs/index.php/fm/article/view/8005>. Consulté le : 24/10/2019.

Richard Fletcher, Rasmus Kleis Nielsen. *Are people incidentally exposed to news on social media? A comparative analysis*. New Media & Society 20(7). 17/8/2017. URL : <https://journals.sagepub.com/doi/abs/10.1177/1461444817724170>. Consulté le : 4/6/2020.

Lucas Graves. *A smarter conversation about how (and why) fact-checking matters*. Nieman Lab. 12/2019. URL : <https://www.niemanlab.org/2019/12/a-smarter-conversation-about-how-and-why-fact-checking-matters/>. Consulté le : 4/6/2020.

Andrew Guess, Jonathan Nagler, Joshua A. Tucker. *Less than you think: Prevalence and predictors of fake news dissemination on Facebook*. Science Advances 5(1). 9/1/2019.

Andrew Guess, Brendan Nyhan, Benjamin Lyons, Jason Reifler. *Avoiding the Echo Chamber about Echo Chambers: Why Selective Exposure to like-Minded Political News Is Less Prevalent than You Think*. Knight Foundation White Paper. 2018.

Andrew Guess, Brendan Nyhan, Jason Reifler. *Selective exposure to misinformation: Evidence from the consumption of fake news during the 2016 U.S. presidential campaign*. 9/1/2018. URL : <https://www.dartmouth.edu/~nyhan/fake-news-2016.pdf>.

Kristoffer Holt, Tine Ustad Figenschou, Lena Frischlich. *Key dimensions of alternative news media*. Digital Journalism 7(7): 860-69. 2019.

Edda Humprecht. *How Do They Debunk « Fake News »? A Cross-National Comparison of Transparency in Fact Checks*. Digital Journalism. 14/10/2019. URL : <https://www.tandfonline.com/doi/abs/10.1080/21670811.2019.1691031>. Consulté le : 4/6/2020.

David M.J. Lazer et al. *The science of fake news*. Science 359(380): 1094-96. 2018.

Briony Swire, Ullrich K.H. Ecker, Stephan Lewandowsky. *The role of familiarity in correcting inaccurate information*. Journal of Experimental Psychology : Learning, Memory and Cognition 43(12): 1948-61. 12/2017.

Soroush Vosoughi, Deb Roy, Sinan Aral. *The spread of true and false news online*. Science 359: 1146-51. 9/3/2018.

## Rapports

*A Multi-Dimensional Approach to Disinformation, Report of the Independent High Level Group on Fake News and Online Disinformation*. 3/2018.

*Digital news report 2019*. 2019.



Richard Fletcher, Alessio Cornia, Lucas Graves, Rasmus Kleis Nielsen. *Measuring the reach of "fake news" and online disinformation in Europe*. Factsheet, 2/2018. URL : <https://www.press.is/static/files/frettamyndir/reuterfake.pdf> .

Divina Frau-Meigs. *Societal costs of "fake news" in the Digital Single Market*. 12/2018. URL : <https://sanef.org.za/wp-content/uploads/2019/02/European-Parliament-Societal-costs-of-fake-news-December-2018.pdf> .

Lucas Galan, Jordan Osserman, Tim Parker, Matt Taylor. *How young people consume news and the implications for mainstream media*. 8/2019. URL : <https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2019-08/FlamingoxREUTERS-Report-Full-KG-V28.pdf>. Consulté le : 24/10/2019.

J.-B. Jeangène Vilmer, A. Escorcía, M. Guillaume, J. Herrera. *Les Manipulations de l'information : un défi pour nos démocraties*. 8/2018. URL : [https://www.diplomatie.gouv.fr/IMG/pdf/les\\_manipulations\\_de\\_l\\_information\\_2\\_\\_cle04b2b6.pdf](https://www.diplomatie.gouv.fr/IMG/pdf/les_manipulations_de_l_information_2__cle04b2b6.pdf). Consulté le : 12/6/2020.

Joshua A. Tucker et al. *Social Media, Political Polarization, and Political Disinformation: A Review of the Scientific Literature*. 3/2018. URL : <http://eprints.lse.ac.uk/87402/1/Social-Media-Political-Polarization-and-Political-Disinformation-Literature-Review.pdf> .

Claire Wardle, Hossein Derakhshan. *Information disorder: Toward an interdisciplinary framework for research and policy making*. 27/9/2017. URL : <https://rm.coe.int/information-disorder-toward-an-interdisciplinary-framework-for-research/168076277c>.

## Article d'encyclopédie

Vérification des faits (rubrique France). Wikipédia. URL : [https://fr.wikipedia.org/wiki/V%C3%A9rification\\_des\\_faits#France](https://fr.wikipedia.org/wiki/V%C3%A9rification_des_faits#France). Consulté le : 28/10/2019.

## Presse

20 Minutes. *Coronavirus : Non, aucun cas n'a été détecté à Perpignan, indique la préfecture*. 26/1/2020. URL : <https://www.20minutes.fr/sante/2703999-20200126-coronavirus-non-aucun-cas-detecte-perpignan-indique-prefecture>. Consulté le : 23/7/2020.

20 Minutes. *Coronavirus : Non, le Covid-19 n'a pas été breveté par l'Institut Pasteur en 2004*. 18/3/2020. URL : <https://player.acast.com/5ab28781627e02027cf860fe/episodes/5e68ec5b7936a86027281fb6?ref=facebook>. Consulté le : 23/7/2020.

20 Minutes. « *Touche pas à mon poste* » déprogrammée ? Cyril Hanouna dément. 26/11/2019. URL : <https://www.20minutes.fr/arts-stars/television/2660459-20191126-touche-poste-deprogramme-cyril-hanouna-dement>. Consulté le : 23/7/2020.



AFP Factuel. *Australie : les causes des incendies, mine de désinformation sur les réseaux sociaux*. 13/1/2020. URL : <https://factuel.afp.com/australie-les-causes-des-incendies-mine-de-desinformation-sur-les-reseaux-sociaux>. Consulté le : 17/1/2020.

AFP Factuel. *Nouveau coronavirus : épidémie mondiale de fausses informations*. 28/1/2020. URL : <https://factuel.afp.com/nouveau-coronavirus-epidemie-mondiale-de-fausses-informations>. Consulté le : 31/1/2020.

Axios. *Big Tech — and China — Struggle with Coronavirus Misinformation*. 28/1/2020. URL : <https://www.axios.com/coronavirus-misinformation-facebook-twitter-google-china-246a0325-b4ea-4465-92ae-5f364a7e965c.html>. Consulté le : 3/2/2020.

Bloomberg. *Sandberg Says Facebook 'Filter Bubble' Problem Is Misunderstood*. 22/10/2019. URL : <https://www.bloomberg.com/news/articles/2019-10-22/sandberg-says-facebook-filter-bubble-problem-is-misunderstood>. Consulté le : 24/10/2019.

BuzzFeed. *Facebook Won't Do Anything To Stop Climate Deniers Capitalising On Australia's Bushfire Crisis*. 22/1/2020. URL : <https://www.buzzfeed.com/hannahryan/facebook-australia-bushfires-climate-change-deniers-facebook>. Consulté le : 27/1/2020.

franceinfo. *#VraiOuFake Un colis en provenance de Chine peut-il transmettre le coronavirus ?*. 28/1/2020. URL : [https://www.francetvinfo.fr/sante/maladie/coronavirus/coronavirus-un-colis-en-provenance-de-chine-peut-il-transmettre-la-maladie\\_3801681.html](https://www.francetvinfo.fr/sante/maladie/coronavirus/coronavirus-un-colis-en-provenance-de-chine-peut-il-transmettre-la-maladie_3801681.html). Consulté le : 23/7/2020.

franceinfo. *Coronavirus : quels sont les comptes Twitter « super propagateurs » de fausses informations en France ?* 4/6/2020. URL : [https://www.francetvinfo.fr/sante/maladie/coronavirus/coronavirus-quels-sont-les-comptes-twitter-super-propagateurs-de-fausses-informations-en-france\\_3982961.html](https://www.francetvinfo.fr/sante/maladie/coronavirus/coronavirus-quels-sont-les-comptes-twitter-super-propagateurs-de-fausses-informations-en-france_3982961.html). Consulté le : 5/6/2020.

franceinfo. *Netflix : la rumeur sur la suppression du partage de compte fait son grand retour*. 23/10/2019. URL : [https://www.francetvinfo.fr/culture/series/netflix/netflix-la-rumeur-sur-la-suppression-du-partage-de-compte-fait-son-grand-retour\\_3671539.html](https://www.francetvinfo.fr/culture/series/netflix/netflix-la-rumeur-sur-la-suppression-du-partage-de-compte-fait-son-grand-retour_3671539.html). Consulté le : 23/7/2020.

franceinfo. *Une infirmière à domicile est-elle payée moins par patient que le prix d'un « Happy Meal » ?* 18/10/2019. URL : [https://www.francetvinfo.fr/sante/politique-de-sante/une-infirmiere-a-domicile-est-elle-payee-moins-par-patient-que-le-prix-dun-happy-meal\\_3664829.html](https://www.francetvinfo.fr/sante/politique-de-sante/une-infirmiere-a-domicile-est-elle-payee-moins-par-patient-que-le-prix-dun-happy-meal_3664829.html). Consulté le : 23/7/2020.

L'ADN. *Nous vivons une véritable épidémie de fake news et il n'y a pas vraiment de remède*. 29/11/2019. URL : <https://www.ladn.eu/media-mutants/reseaux-sociaux/fakes-news-continuent-progression-reseaux-sociaux/>. Consulté le : 4/6/2020.

L'ADN. *Un outil pour comprendre l'orientation politique des médias*. 21/1/2020. URL : <https://www.ladn.eu/media-mutants/reseaux-sociaux/comment-echapper-bulles-filtre-reseaux-sociaux/>. Consulté le : 22/1/2020.



L'Indépendant. *Coronavirus : non, aucun cas n'a été détecté à Perpignan*. 26/1/2020. URL : <https://www.lindependant.fr/2020/01/26/coronavirus-non-aucun-cas-na-ete-detecte-a-perpignan,8688548.php>. Consulté le : 23/7/2020.

Le Figaro. *Facebook a franchi la barre des «37 millions d'utilisateurs» en France en 2019*. 10/2/2020. URL : <https://www.lefigaro.fr/secteur/high-tech/facebook-a-franchi-la-barre-des-37-millions-d-utilisateurs-en-france-en-2019-20200210>. Consulté le : 23/7/2020.

Le Monde. *« Mesdames, évitez de rentrer seule le soir... » : le message factice de la ville de Limoges*. 12/11/2019. URL : [https://www.lemonde.fr/les-decodeurs/article/2019/11/12/mesdames-evitez-de-rentre-seule-le-soir-le-message-factice-de-la-ville-de-limoges\\_6018869\\_4355770.html](https://www.lemonde.fr/les-decodeurs/article/2019/11/12/mesdames-evitez-de-rentre-seule-le-soir-le-message-factice-de-la-ville-de-limoges_6018869_4355770.html). Consulté le : 23/7/2020.

Le Monde. *Après l'incendie de l'usine Lubrizol à Rouen, des intoxic en série*. 30/9/2019. URL : [https://www.lemonde.fr/les-decodeurs/article/2019/09/30/oiseaux-morts-eau-non-potable-communique-bidon-intox-en-serie-apres-l-incendie-de-rouen\\_6013661\\_4355770.html](https://www.lemonde.fr/les-decodeurs/article/2019/09/30/oiseaux-morts-eau-non-potable-communique-bidon-intox-en-serie-apres-l-incendie-de-rouen_6013661_4355770.html). Consulté le : 23/7/2020.

Le Monde. *Coronavirus : Bill Gates ciblé par des rumeurs et infox complotistes*. 5/2/2020. URL : [https://www.lemonde.fr/les-decodeurs/article/2020/02/05/coronavirus-bill-gates-cible-par-des-rumeurs-et-infox-complotistes\\_6028482\\_4355770.html](https://www.lemonde.fr/les-decodeurs/article/2020/02/05/coronavirus-bill-gates-cible-par-des-rumeurs-et-infox-complotistes_6028482_4355770.html). Consulté le : 23/7/2020.

Le Monde. *Décodex : vérification de sources d'informations, pages Facebook et chaînes YouTube*. URL : <https://www.lemonde.fr/verification/>. Consulté le : 23/7/2020.

Le Monde. *L'annuaire des sources du Décodex : mode d'emploi*. 23/1/2017.a. URL : [https://www.lemonde.fr/les-decodeurs/article/2017/01/23/l-annuaire-des-sources-du-decodex-mode-d-emploi\\_5067719\\_4355770.html](https://www.lemonde.fr/les-decodeurs/article/2017/01/23/l-annuaire-des-sources-du-decodex-mode-d-emploi_5067719_4355770.html). Consulté le : 23/7/2020.

Le Monde. *Le coronavirus qui sévit en Chine n'a pas été « créé en 2003 aux USA »*. 27/1/2020. URL : [https://www.lemonde.fr/les-decodeurs/article/2020/01/27/le-coronavirus-qui-sevit-en-chine-n-a-pas-ete-cree-en-2003-aux-usa\\_6027390\\_4355770.html](https://www.lemonde.fr/les-decodeurs/article/2020/01/27/le-coronavirus-qui-sevit-en-chine-n-a-pas-ete-cree-en-2003-aux-usa_6027390_4355770.html). Consulté le : 23/7/2020.

Le Monde. *Le Décodex, un outil de vérification de l'information*. 23/1/2017.b. URL : [https://www.lemonde.fr/les-decodeurs/article/2017/01/23/le-decodex-un-premier-premier-pas-vers-la-verification-de-masse-de-l-information\\_5067709\\_4355770.html](https://www.lemonde.fr/les-decodeurs/article/2017/01/23/le-decodex-un-premier-premier-pas-vers-la-verification-de-masse-de-l-information_5067709_4355770.html). Consulté le : 12/6/2020.

Le Monde. *Incendie à Rouen : intoxic en série*. 30/9/2019. URL : [https://www.lemonde.fr/les-decodeurs/article/2019/09/30/oiseaux-morts-eau-non-potable-communique-bidon-intox-en-serie-apres-l-incendie-de-rouen\\_6013661\\_4355770.html](https://www.lemonde.fr/les-decodeurs/article/2019/09/30/oiseaux-morts-eau-non-potable-communique-bidon-intox-en-serie-apres-l-incendie-de-rouen_6013661_4355770.html). Consulté le : 23/7/2020.

Le Monde. *Les fausses informations circulent de moins en moins sur Facebook*. 17/10/2018. URL : [https://www.lemonde.fr/les-decodeurs/article/2018/10/17/les-fausses-informations-perdent-du-terrain-sur-facebook\\_5370461\\_4355770.html](https://www.lemonde.fr/les-decodeurs/article/2018/10/17/les-fausses-informations-perdent-du-terrain-sur-facebook_5370461_4355770.html). Consulté le : 4/6/2020.

Les Echos. *Coronavirus : un rapport épingle la persistance des « fake news » sur Twitter*. 20/5/2020. URL : <https://www.lesechos.fr/tech-medias/hightech/coronavirus-un-rapport-epingle-la-persistance-des-fake-news-sur-twitter-1204358>. Consulté le : 20/5/2020.



Libération. *A propos - Checknews*. URL : <https://www.liberation.fr/checknews/about/>. Consulté le : 28/10/2019.

Libération. *Blanquer a-t-il vraiment dit que « 99,9% des enseignants soutiennent la réforme du bac » ?*. 23/1/2020. URL : [https://www.liberation.fr/checknews/2020/01/23/blanquer-a-t-il-vraiment-dit-que-999-des-enseignants-soutiennent-la-reforme-du-bac\\_1774760](https://www.liberation.fr/checknews/2020/01/23/blanquer-a-t-il-vraiment-dit-que-999-des-enseignants-soutiennent-la-reforme-du-bac_1774760). Consulté le : 23/7/2020.

Libération. *Brèves de presse, Alertes infos, Les News... : qui est derrière ces comptes d'info sur Twitter ?* 16/1/2020. URL : [https://www.liberation.fr/checknews/2020/01/16/breves-de-presse-alerte-infos-les-news-qui-est-derriere-ces-comptes-d-info-sur-twitter\\_1770396](https://www.liberation.fr/checknews/2020/01/16/breves-de-presse-alerte-infos-les-news-qui-est-derriere-ces-comptes-d-info-sur-twitter_1770396). Consulté le : 4/6/2020.

Libération. *Un bug sur Facebook a-t-il révélé que les messages de Greta Thunberg étaient écrits par son père ?*. 16/1/2020. URL : [https://www.liberation.fr/checknews/2020/01/16/est-il-vrai-qu-un-bug-sur-facebook-a-revele-que-les-messages-de-greta-thunberg-etaient-ecrits-par-so\\_1773407](https://www.liberation.fr/checknews/2020/01/16/est-il-vrai-qu-un-bug-sur-facebook-a-revele-que-les-messages-de-greta-thunberg-etaient-ecrits-par-so_1773407). Consulté le : 23/7/2020.

Numerama. *Conflits\_Fr : qui est derrière le compte Twitter qui partage des informations sur le coronavirus sans les vérifier ?*. 16/3/2020. URL : [https://www.numerama.com/politique/611424-conflits\\_fr-qui-est-derriere-le-compte-twitter-qui-partage-des-informations-sur-le-coronavirus-sans-les-verifier.html](https://www.numerama.com/politique/611424-conflits_fr-qui-est-derriere-le-compte-twitter-qui-partage-des-informations-sur-le-coronavirus-sans-les-verifier.html). Consulté le : 24/3/2020.

Siècle Digital. *L'histoire de la propagande chinoise sur Twitter pendant l'épidémie du Covid-19*. 31/3/2020. URL : <https://siecledigital.fr/2020/03/31/lhistoire-de-la-propagande-chinoise-sur-twitter-pendant-lepidemie-du-covid-19/>. Consulté le : 6/4/2020.

Slate. *Twitter's New Order*. 5/3/2017. URL : [https://www.slate.com/articles/technology/cover\\_story/2017/03/twitter\\_s\\_timeline\\_algorithm\\_and\\_its\\_effect\\_on\\_us\\_explained.html?via=gdpr-consent](https://www.slate.com/articles/technology/cover_story/2017/03/twitter_s_timeline_algorithm_and_its_effect_on_us_explained.html?via=gdpr-consent). Consulté le : 29/4/2020.

The Conversation. *Rating News Sources Can Help Limit the Spread of Misinformation*. 2/12/2019. URL : <http://theconversation.com/rating-news-sources-can-help-limit-the-spread-of-misinformation-126083>. Consulté le : 4/6/2020.

The Verge. *Google's ads just look like search results now*. 23/1/2020. URL : <https://www.theverge.com/tldr/2020/1/23/21078343/google-ad-desktop-design-change-favicon-icon-ftc-guidelines>. Consulté le : 27/1/2020.

Wired. *How One Particular Coronavirus Myth Went Viral*. 19/3/2020. URL : <https://www.wired.com/story/opinion-how-one-particular-coronavirus-myth-went-viral/>. Consulté le : 24/3/2020.

Wired. *YouTube Gaming's Most-Watched Videos Dominated by Scams and Cheats*. 10/2/2020. URL : <https://www.wired.com/story/youtube-gaming-scams-cheats-livestreams/>. Consulté le : 25/2/2020.



## Autres sources

À propos. Twitter. URL : <https://about.twitter.com/fr.html>. Consulté le : 23/7/2020.

À propos des comptes certifiés. Twitter. URL : <https://help.twitter.com/fr/managing-your-account/about-twitter-verified-accounts>. Consulté le : 23/7/2020.

*Ce que les éditeurs doivent savoir à propos de la vérification des informations par des tiers sur Facebook | Pages d'aide de Facebook Business.* Facebook. URL : <https://fr-fr.facebook.com/business/help/18222309230722>. Consulté le : 23/7/2020.

*Données personnelles : le CSA mène une étude sur le phénomène de propagation des fausses informations sur les réseaux sociaux.* Conseil supérieur de l'audiovisuel. 13/9/2019. URL : <https://www.csa.fr/Informer/Toutes-les-actualites/Actualites/Donnees-personnelles-le-CSA-mene-une-etude-sur-le-phenomene-de-propagation-des-fausses-informations-sur-les-reseaux-sociaux>. Consulté le : 23/7/2020.

Direction générale de l'enseignement scolaire. *Infographie sur les outils de vérification des faits (mars 2020).* Twitter. 27/3/2020. URL : [https://twitter.com/eduscol\\_emi/status/1243472130441789440](https://twitter.com/eduscol_emi/status/1243472130441789440). Consulté le : 6/4/2020.

*IFCN Covid-19 Misinformation.* Poynter. URL : <https://www.poynter.org/ifcn-covid-19-misinformation/>. Consulté le : 6/4/2020.

*La carte des théories du complot sur le coronavirus.* Conspiracy Watch. 23/3/2020. URL : <https://www.conspiracywatch.info/la-carte-des-theories-du-complot-sur-le-coronavirus.html>. Consulté le : 6/4/2020.

*Misinformation Research - Airtable Universe.* Airtable. URL : <https://airtable.com/universe/expPeddCpX0wOeNNE/misinformation-research>. Consulté le : 4/6/2020.

*NewsGuard : l'outil de confiance sur Internet.* NewsGuard. 28/10/2019. URL : <https://www.newsguardtech.com/fr/>. Consulté le : 28/10/2019.

*Signalement : campagne électorale.* CNIL. URL : <https://www.cnil.fr/fr/webform/signalement-campagne-electorale>. Consulté le : 19/2/2020.

*Superspreaders.* NewsGuard. 23/4/2020. URL : <https://www.newsguardtech.com/superspreaders/>. .

*The Truth behind Filter Bubbles: Bursting Some Myths.* Reuters Institute for the Study of Journalism. URL : <https://reutersinstitute.politics.ox.ac.uk/risj-review/truth-behind-filter-bubbles-bursting-some-myths>. Consulté le : 17/2/2020.

Dr. Tedros Adhanom Ghebreyesus. *Conférence sur la sécurité de Munich, discours du directeur général.* 15/2/2020. URL : <https://www.who.int/fr/dg/speeches/detail/munich-security-conference>. .



Manon Berriche. *Le fact-checking est-il vraiment efficace ?*. SciencesPo. 22/1/2020. URL : <https://www.sciencespo.fr/actualites/actualite/C3%A9s/le-fact-checking-est-il-vraiment-efficace/4539>. Consulté le : 24/1/2020.

Laurent Bigot. *Le fact checking à l'épreuve des fake news*. La Revue des Médias. 17/10/2019. URL : <http://larevuedesmedias.ina.fr/le-fact-checking-lepreuve-des-fake-news>. Consulté le : 4/6/2020.

Conniefan.com. *Classifying fake news*. 22/3/2017. URL : <https://www.conniefan.com/2017/03/classifying-fake-news/>. Consulté le : 24/10/2019.

Rory Smith. *How to Investigate Health Misinformation (and Anything Else) Using Twitter's API*. First Draft. 6/3/2020. URL : [https://firstdraftnews.org:443/latest/how-to-investigate-health-misinformation-and-anything-else-using-twitters-api/?utm\\_source=Daily+Lab+email+list&utm\\_campaign=86bc6e883c-dailylabemail3&utm\\_medium=email&utm\\_term=0\\_d68264fd5e-86bc6e883c-396240277](https://firstdraftnews.org:443/latest/how-to-investigate-health-misinformation-and-anything-else-using-twitters-api/?utm_source=Daily+Lab+email+list&utm_campaign=86bc6e883c-dailylabemail3&utm_medium=email&utm_term=0_d68264fd5e-86bc6e883c-396240277). Consulté le : 10/3/2020.



## 9. Remerciements

Le Conseil supérieur de l'audiovisuel tient à remercier l'ensemble des personnes rencontrées dans le cadre de la préparation de cette étude et qui ont permis, par leurs retours constructifs, de contribuer à l'améliorer. Les échanges n'impliquent cependant pas validation ou approbation des travaux par ces personnes.

Ont ainsi pris le temps de recevoir les services du Conseil supérieur de l'audiovisuel :

- AFP Factuel, représentée par Grégoire Lemarchand, rédacteur en chef investigation numérique à l'AFP.
- Romain Badouard, maître de conférences en sciences de l'information et de la communication à Paris-II, chercheur au sein du laboratoire CARISM.
- Les Décodeurs du *Monde* : Maxime Ferrer, Assma Maad, Jonathan Pariente et Adrien Sénécat.
- Twitter France, représentée par Audrey Herblin-Stoop, Public Policy Director, France et Russie.