



HAL
open science

Identification et reconfiguration des communautés politiques durant la campagne présidentielle française 2017

Noe Gaumont, Mazyar Panahi, David Chavalarias

► **To cite this version:**

Noe Gaumont, Mazyar Panahi, David Chavalarias. Identification et reconfiguration des communautés politiques durant la campagne présidentielle française 2017. 2017. hal-01575456v1

HAL Id: hal-01575456

<https://hal.science/hal-01575456v1>

Preprint submitted on 20 Aug 2017 (v1), last revised 21 Sep 2018 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

IDENTIFICATION ET RECONFIGURATION DES
COMMUNAUTÉS POLITIQUES DURANT LA CAMPAGNE
PRÉSIDENTIELLE FRANÇAISE 2017
analyse socio-sémantiques des réseaux de
militants politiques sur Twitter

Noé Gaumont¹, Maziyar Panahi², and David Chavalarias ^{*1}

¹CNRS, EHESS-CAMS / ISC-PIF

²CNRS, Institut des Systèmes Complexes Paris Île-de-France (ISC-PIF)

Titre anglais : *Identification and reconfiguration of political communities during 2017 French presidential elections - socio-semantic networks of Twitter political activists*

Résumé

À l'occasion de l'élection présidentielle française de 2017, nous avons développé le *Politoscope*, une plateforme ouverte d'analyse des tweets politiques qui a collecté et analysé près de 60 millions d'échanges Twitter entre plus de 2,4 millions d'utilisateurs interagissant avec les personnalités politiques françaises. Dans cet article, nous nous intéressons à ce que révèlent ces données quant à l'évolution du paysage politique lors de ces élections. Parmi les nombreux types d'interactions qu'offre Twitter, nous nous concentrons sur l'action de retweet. Nous caractérisons les réseaux socio-sémantiques de l'espace politique français et leur évolution pendant les onze mois précédents l'élection présidentielle. Nous montrons que l'analyse du graphe des retweets permet de suivre précisément et de manière automatique les reconfigurations d'un paysage politique multipolaire, en caractérisant les différentes communautés politiques, leurs stratégies d'alliances et leurs spécificités sémantiques.

Cette première étude de l'évolution socio-sémantique d'un paysage politique multipolaire contribue à l'émergence de nouvelles méthodes issues des sciences sociales computationnelles, qui modifient en profondeur les pratiques de la recherche en sciences sociales, et en sciences politiques en particulier.

Mots-clés : opinion dynamics, political activism, elections, socio-semantic networks, networks dynamics, Twitter.

*Corresponding author : david.chavalarias at ehess dot fr

1 Introduction

Les espaces numériques sont de plus en plus présents et influents dans le fonctionnement de nos démocraties, au point que ce qui s’y passe ne peut plus être ignoré lorsqu’il s’agit d’analyser l’évolution du paysage politique. Après chacun des grands séismes politiques ayant secoué les démocraties occidentales ces derniers mois (ex. le Brexit, l’élection de Trump), les commentateurs se sont tournés vers les réseaux sociaux pour y trouver des éléments de compréhension.

Les réseaux sociaux ont également été, en quelques années, totalement intégrés aux stratégies des partis politiques. Barack Obama a été qualifié de “*game changer*” en 2012 alors qu’il était le premier candidat à la présidentielle à utiliser massivement les réseaux sociaux pour sa campagne. La campagne présidentielle USA 2016 a confirmé cette tendance avec un budget cumulé dédié aux réseaux sociaux estimé à près de 500M€ [Gre15].

La France a suivi cette même tendance. Les partis politiques les plus importants ont aujourd’hui recours aux services de sociétés spécialisées pour mieux comprendre leurs électeurs potentiels via l’analyse des masses de données issues des réseaux sociaux. Parmi ceux-ci, Twitter détient une bonne place dans les sources disponibles. C’est par ailleurs un réseau social utilisé par la plupart des personnalités politiques pour réagir à l’actualité ou donner de la visibilité à leurs opinions et leurs prises de paroles. Ce réseau social est ainsi devenu un média incontournable pour le monde politique lorsqu’il s’agit de faire passer des messages ou d’asseoir une légitimité.

À l’occasion de l’élection présidentielle française de 2017, nous avons développé le *Politoscope*¹, une plateforme ouverte d’analyse des tweets politiques qui a collecté et analysé près de 60 millions d’échanges Twitter entre plus de 2,4 millions d’utilisateurs. Avec ces données comme cas d’étude, nous présentons une méthode permettant de suivre automatiquement l’évolution socio-sémantique d’un paysage politique multipolaire.

2 Travaux existants

Twitter fait l’objet d’une littérature académique assez importante, car cette plateforme cumule deux avantages. D’une part, elle est accessible au grand public, même si elle n’est pas représentative de l’ensemble de la population mondiale ou même Française [BR15], et ses contenus sont massivement repris par la presse traditionnelle ou la télévision. D’autre part, la collecte de données y est relativement aisée grâce aux différentes API proposées par Twitter².

Il existe de nombreuses méthodologies pour aborder les productions de Twitter dans le cadre de l’étude de l’action politique. Certaines méthodes [RY10, PP11, CGR⁺11] s’attachent à tirer profits du contenu des tweets pour déduire l’orientation politique d’un tweet ou d’un utilisateur. Cependant, les contraintes sur le format des tweets semblent limiter leur efficacité. Par exemple, d’après Cohen et Ruths [CR13], les hashtags utilisés pour parler de politique varient en fonction du militantisme des personnes et ne semblent que peu différents pour les personnes les moins militantes.

Afin de contourner les problèmes liés au traitement automatique du langage, les recherches se sont tournées vers des méthodes analysant la structure des échanges ayant lieu sur Twitter au lieu de leur contenu. Elles s’inspirent ainsi d’une étude pionnière sur les blogs politiques aux États Unis [AG05], qui a démontré qu’il y a plus de liens entre deux blogs s’ils sont du même bord politique.

Dans le cadre de Twitter, il existe plusieurs définitions possibles d’un lien entre deux comptes. Un compte peut suivre ou mentionner un autre compte. Il peut répondre, citer, commenter ou simplement reprendre à son compte (retweet) le tweet d’un autre compte. Toutes ces notions de liens sont pertinentes et dépendent de l’objectif de l’étude [DOG15]. Cependant pour classer les comptes en fonction de leur orientation politique, il semble que l’utilisation du retweet comme définition d’un lien entre deux comptes soit la plus utilisée. Garimella *et al.* [GDGM16] et Connover *et al.* [CGR⁺11] ont notamment observé que pour étudier les débats politiques sur Twitter l’utilisation du retweet est plus pertinente que la mention.

Connover *et al.* [CGR⁺11] ont été parmi les premiers à exploiter les retweets en étudiant la structure de l’espace politique aux USA. Ils ont retrouvé sa nature bipartite et ils ont démontré qu’il existe significativement moins de retweets entre deux comptes lorsqu’ils sont de partis différents. De nombreuses autres études ont par la suite utilisé ce type de relation. Boutet *et al.* [BHY12] ont appliqué la même méthode pour étudier la structure de l’espace politique en Angleterre et analyser les différences de pratiques dans la manière dont les groupes politiques utilisent les mentions, hashtags et les liens URL.

Garimella *et al.* [GDGM16] et Morales *et al.* [MBLB15] ont utilisé la notion de retweet pour mesurer l’existence de polarisation et de controverses sur Twitter. Amor *et al.* [AVC⁺16] s’en servent pour caractériser le débat sur *care.data* du *National Health System* anglais et Guerrero *et al.* [GS15] pour étudier le mouvement *Catalan Process Toward Independence*.

1. <http://politoscope.org>

2. <https://dev.twitter.com/overview/api>

Ces approches ont permis de retrouver des structures politiques binaires mais la notion de retweet permet également de retrouver des structures multipolaires. C’est le cas de Cherepnalkoski *et al.* [CM15, CM16] qui ont retrouvé sur Twitter l’affiliation politique des députés européens via les retweets. Les auteurs notent cependant que l’approche se généralise mal à l’ensemble des utilisateurs car en plus des accointances politiques il existe des préférences nationales.

Le retweet n’est pas la seule relation exploitable pour caractériser la structure d’un espace politique. Les travaux de Barberá [Bar15] ont utilisé les relations suiveur/suivi (follower/followee) entre comptes. L’originalité de leur travaux est d’assigner à chaque compte une position entre -1 et 1 sur un axe idéologique gauche-droite. Ainsi, un compte n’est pas assigné au groupe démocrate mais à une position proche d’autres démocrates. Cette approche permet de comparer un ensemble de comptes sur un continuum idéologique. Ils ont été répliqués avec succès sur des données Twitter en France par Briatte et Gallic [BG15].

En utilisant leur méthode d’estimation idéologique, Barberá *et al.* [BJN⁺15] ont également montré qu’il y a une forte polarisation des retweets lorsque le sujet évoqué est politique, c’est-à-dire qu’il existe de nombreux retweets entre personnes proches idéologiquement et peu lorsqu’elles sont éloignées.

Enfin il est important de noter que malgré l’existence de ces nombreux travaux évaluant l’alignement idéologique, l’une des applications des sciences politiques consistant à effectuer des prévisions de vote reste problématique [Hub15], notamment à cause de la représentativité de Twitter par rapport à la population réelle.

En quelques années, Twitter est donc devenu un terrain d’étude fructueux pour les sciences politiques, avec cependant certaines limites concernant les types de populations qu’il y est pertinent d’étudier.

Dans cet article, nous reprenons et développons des méthodes d’analyse de retweets en les appliquant à un corpus généré par des militants politiques pendant près d’un an. Cette durée étendue nous permet d’étudier pour l’une des premières fois la dynamique socio-sémantique d’un paysage politique multipolaire.

3 Identification des groupes sociaux de la twittosphère politique française

En 1890, Gabriel Tarde [Tar90] proposait une définition du groupe social comme “*une collection d’êtres en tant qu’ils sont en train de s’imiter entre eux ou en tant que, sans s’imiter actuellement, ils se ressemblent et que leurs traits communs sont des copies anciennes d’un même modèle*”. La notion d’imitation que mobilisait Tarde à l’époque était relativement générale et recoupe en grande partie celles plus larges d’*apprentissage social* ou d’*influence sociale* (voir [MGFL16, Tom09] pour une contextualisation de ces notions). Plus de cent ans après, Twitter offre un terrain privilégié pour adapter cette proposition à l’espace numérique informationnel en prenant pour traits les manières de s’exprimer sur Twitter. Cette plateforme associe en effet trois éléments qui favorisent le suivi de la diffusion de ces traits :

- Chacun peut publier, sans filtre, des contenus visibles par tous,
- Le système de follower/followee permet à quiconque de s’abonner aux publications de n’importe quel utilisateur,
- Le système de retweet permet, sans effort, de copier un contenu (tweet) en le relayant à l’identique à tous ses abonnés.

Ainsi, en se concentrant sur les événements de type retweet (relais d’information sans modification), nous disposons de marqueurs permettant de définir, pour une période donnée, un ensemble de groupes d’utilisateurs au sein desquels les contenus Twitter ont tendance à circuler en subissant des modifications moindres (reprise à l’identique) que lorsqu’ils circulent entre ces groupes (transitions pendant lesquelles ils peuvent être détournés, commentés de manière ironique ou humoristique, contredits ou tout simplement non repris). Nous définissons ainsi les groupes sociaux de la twittosphère politique comme les réseaux denses de comptes relayant entre eux, de manière récurrente, des contenus sans modification.

Pour cette étude, nous avons développé un dispositif de capture temps-réel de l’activité de la twittosphère politique qui a collecté en continu, durant les 10 mois précédents l’élection présidentielle, près de 60 millions de tweets émis autour des principaux acteurs politiques français par plus de 2,4M de comptes uniques (cf. annexes B). Nous noterons \mathcal{D} ce jeu de données.

Nous avons estimé à moins de 5% la proportion de retweets émis par des bots dans \mathcal{D} (cf. annexes C), ce qui nous garantit une certaine représentativité de nos résultats au regard du militantisme politique en ligne.

Ce jeu de données s’est révélé rétrospectivement d’une qualité exceptionnelle pour tester des méthodes de reconstruction de communautés. Le caractère exceptionnellement mouvementé de ces élections (cf. annexe A) dans le paysage politique multipolaire constituent un défi pour la question de l’identification dynamique de communautés. Ils permettent de tester les méthodes en vérifiant qu’elles sont bien capables de restituer ces évolutions du paysage politique.

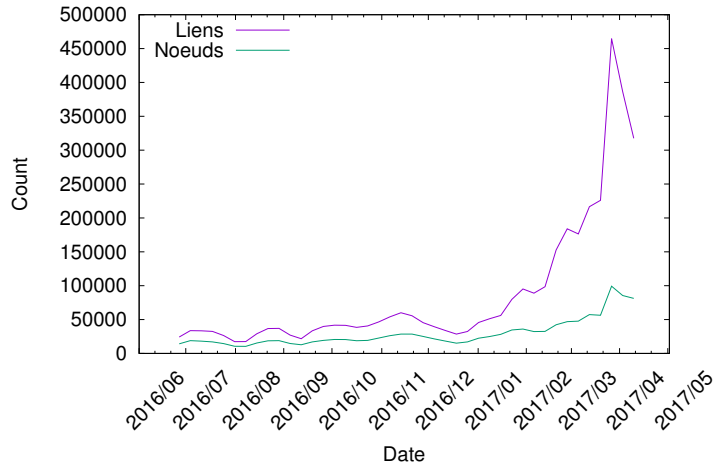


FIGURE 1 – Évolution du nombre de nœuds et de liens pour les graphes G_{T_i} .

3.1 Les communautés de retweets

Sur Twitter, les marqueurs d’influence sociale portant sur les formes d’expression ou les opinions dépassent largement le cadre du retweet. Les utilisateurs peuvent par exemple s’influencer les uns les autres dans leur utilisation d’expressions ou de hashtags sans nécessairement retweeter. Par ailleurs l’action de retweet peut avoir plusieurs interprétations, une interprétation négative étant par exemple fréquente lorsque le retweet est tardif (cf. annexes I). Cependant, dans la grande majorité des cas, lorsque l’action de retweet est répétée dans un contexte politique, le retweet contribue à la diffusion des idées et des opinions dans l’ensemble du réseau [BGL10, KLPM10, CHBG10, BHMW11] et peut être assimilé à une forme d’adhésion à des idées, un candidat ou un parti politique.

Le suivi de compte (*follow*) est une autre fonctionnalité très étudiée de Twitter qui pourrait aussi servir de proxy pour identifier une forme d’influence sociale d’un compte sur un autre. Il est cependant peu fréquent d’arrêter de suivre un compte et par conséquent l’action “follow” est moins informative de ce point de vue que le retweet, qui est intrinsèquement temporel et indique une proximité de l’usager pour ce qui est dit par un compte à un instant donné. La fonction “follow” peut par ailleurs être utilisée dans le simple but de se tenir informé des propos tenus par un compte, tout en gardant une certaine distance à leur égard.

Le retweet a enfin l’avantage de quantifier l’influence sociale entre deux comptes puisque sur une période donnée, il peut exister aucun, peu ou beaucoup de retweets entre deux comptes. Ces relations de retweets peuvent de plus être uni ou bi-directionnelles, information que nous n’exploiterons pas ici.

Afin de définir des groupes d’utilisateurs sur la base des retweets, nous avons sélectionné les retweets dans notre jeu de données \mathcal{D} et défini un graphe $G_T = (V, E)$ où chaque nœud, $n \in V$, est un compte Twitter et chaque lien pondéré et non dirigé, $(u, v, w) \in E$, représente les w retweets ayant eu lieu entre les comptes u et v au cours de la période temporelle T .

Pour détecter les groupes politiques dans G_T , nous adoptons une approche similaire à celles proposées par [HKKS04, FSB06, PBV07] en appliquant un algorithme de détection de communautés sans recouvrement sur G_T . Parmi l’ensemble des algorithmes existants et régulièrement cités dans les états de l’art [XKS13, Kan14, FH16], nous appliquons l’algorithme de Louvain [BGLL08] qui, étant donné la définition adoptée pour la notion de lien, calcule des communautés de nœuds qui correspondent à l’interprétation de groupes sociaux que nous recherchons : ce sont des sous-groupes de comptes qui échangent beaucoup plus d’information sans modification à l’intérieur du groupe qu’avec le reste du réseau.

Nous adopterons dans la suite les définitions suivantes :

Définition : lien w -fort : deux comptes A et B sont reliés par un lien fort sur la période T s’ils se sont retweetés au moins w fois quelle que soit la direction.

Définition : w -communautés : communautés détectés sur des graphes où n’ont été conservés que les liens au moins w -forts.

Dans ce qui suit, les analyses portent sur les communautés calculées sur des périodes T de deux semaines (cf. annexes D pour la justification de cette période temporelle) et des graphes ne contenant que des liens au moins 3-forts. Ce seuil permet de garder les 10% des relations qui sont les plus fortes. L’évolution du nombre de liens et de comptes du graphe ainsi filtré est présentée sur la figure 1. L’annexe F propose quant à elle la reconstruction des 5–communautés sur la période août 2016 - décembre 2016 qui restitue une image de la twittosphère politique avant le début de la campagne présidentielle.

Les communautés politiques $\mathcal{C}^T = \{C_i^T\}_{1 \leq i \leq N}$ trouvées par cette méthode ont été labellisées avec les noms des principales personnalités politiques présentes dans ces groupes, ceci afin de proposer une lecture

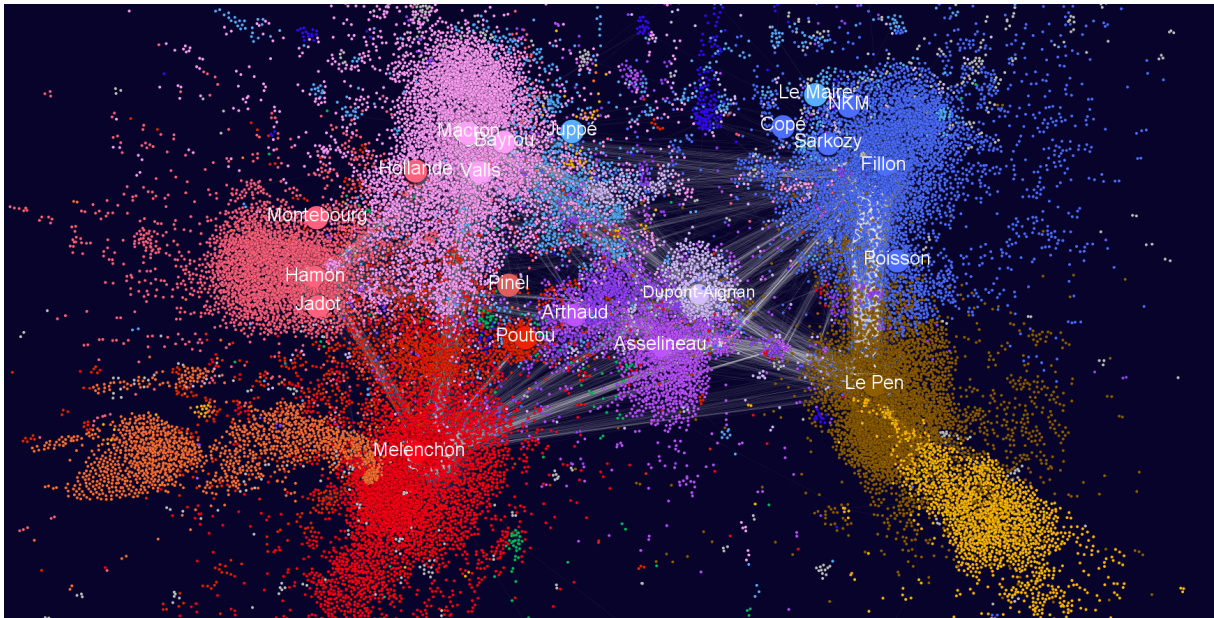


FIGURE 2 – 3–communautés du paysage politique français le jour du 1er tour (T=[9 avril 2017 - 23 avril 2017]). Le paysage est très morcelé et multipolaire, les principales forces politiques étant représentées (les pourcentages indiqués correspondent aux scores obtenus au premier tour) : Mélenchon et la *France Insoumise* (rouge vif - 19.58%), Hamon et le *Parti Socialiste* (rouge clair - 6.36%), Macron et *En Marche!* (rose - 24.01%), Fillon et *Les Républicains* (bleu - 20.01%), Le Pen et le *Front National* (marron - 21.30%), Dupont-Aignan et *Debout la France* (lilas - 4.7%). Pour les “petits” candidats, Asselineau (0.92%) et Arthaud (0.64%) se retrouvent dans la même communauté (violet), Poutou (1.09%) se retrouve avec Mélenchon mais n’a pas vraiment de communauté Twitter. On remarquera que non seulement Bayrou est dans la communauté Macron, ce qui est normal après son ralliement le 23 février 2017, mais que Juppé, ex-candidat à la primaire *Les Républicains*, est beaucoup plus proche de la communauté Macron que de celle de Fillon et Sarkozy. Son “lieutenant” Edouard Philippe, deviendra premier ministre de Macron. Deux communautés non labellisées par un leader politique sont néanmoins importantes. En orange à gauche, proche de la communauté Mélenchon on retrouve une communauté identifiée comme proche de la *Discorde Insoumise*, communauté du jeu vidéo qui a joué un rôle important dans la campagne de Mélenchon. En jaune en bas à droite, proche de la communauté Le Pen, nous retrouvons une communauté de comptes populistes et nationalistes anglophones qui a fortement soutenu Le Pen. Elle est elle-même composée de sous-communautés, en particulier des supporters anglophones de Le Pen, des supporters de Trump et des nationalistes UK pro-Brexit. On relèvera également des différences importantes avec la carte pré-campagne présidentielle de 2016 (figure 15) : en plus de l’axe “gauche-droite” qui était alors dominant, un nouvel axe semble important (“haut-bas” sur la carte) qui distingue les partis à tendance nationaliste ou protectionniste des partis adhérant à la mondialisation et à la globalisation des échanges économiques (le fait que les petits candidats soient au “centre” de la carte ne doit pas être interprété comme une forme de centralité. Seules les contiguïtés topologiques sont significatives sur ce type de graphe. Les communautés des “petits” candidats étant relativement peu connectées aux autres se retrouvent au centre du fait de l’algorithme de spatialisations du graphe).

de l’évolution des communautés politiques autour des principaux acteurs de la campagne présidentielle (cf. figure 2). Nous avons observé qu’il est possible d’étudier le paysage politique avec une granularité plus ou moins fine en choisissant une liste plus ou moins grande de comptes représentatifs et en ajustant la résolution de l’algorithme de Louvain afin de distinguer les communautés autour de ces comptes. Il est ainsi possible de révéler la structure multi-échelle de l’organisation d’un espace politique. Ceci sort cependant du cadre du présent article. Dans ce qui va suivre, nous avons pris les labels parmi la liste des candidats à la présidentielle, et ajouté quelques personnes politiques qui, bien que non candidates, ont joué un rôle important dans cette campagne (ex. : Président sortant, ancien Président, favoris des sondages, etc., voir les annexes B.3). La résolution de l’algorithme de Louvain est la résolution par défaut.

La méthodologie d’identification des groupes sociaux que nous proposons ne fait pas l’hypothèse que ceux-ci puissent être déterminés de manière exacte ou exhaustive. La méthode assume les imprécisions de qualification au niveau individuel (comme par exemple des comptes qui seraient catégorisés dans des groupes qui se révéleraient inappropriés après enquête auprès des propriétaires de ces comptes) et l’arbitraire de certains choix concernant le positionnement des frontières de ces groupes suite au paramétrage de l’algorithme de Louvain ou du seuil w . Les groupes sociaux, quelle qu’en soit leur définition, ont des frontières floues et mouvantes. Notre approche ne cherche pas tant la position absolue des frontières des groupes sociaux que la manière dont celles-ci fluctuent avec le temps, les bifurcations qu’elles mettent en évidence et les sous-corpus de tweets que cette approche permettrait de constituer (cf.). Nous nous sommes cependant assurés, par des études de sensibilité sur les principaux paramètres (granularité dans Louvain, période temporelle, niveau de filtre sur les liens) que les résultats présentés ici ne sont pas un artefact de choix particuliers dans la modélisation.

3.2 Période caractéristique de stabilité des communautés

Les réseaux socio-sémantiques co-évoluent : les acteurs sont influencés par les contenus auxquels ils sont exposés, ce qui les amène à faire évoluer leur réseau local et donc à changer la teneur des contenus auxquels ils seront exposés dans le futur. Ainsi, il est fréquent que des personnes changent de sphère d’influence politique au cours d’une campagne présidentielle ou deviennent influencées par de nouveaux acteurs, ce qui se traduit sur Twitter par une modification des motifs de retweets.

Pour tenir compte de ce phénomène, nous avons considéré des ensembles de graphes $\mathcal{G} = \{G_T\}_{T \in \mathcal{T}}$ où \mathcal{T} est un ensemble d’intervalles temporels, $\mathcal{T} = \{T_i\}_{1 \leq i \leq K}$, tel que $T_i = [t_{min}^i, t_{max}^i]$, t_{min}^i et t_{max}^i étant des entiers positifs représentant des dates dans le système de référence Unix (que nous mentionnerons en donnant la date usuelle). Les liens d’un graphe G_{T_i} ne sont définis que par les retweets ayant lieu dans l’intervalle T_i .

Le choix de \mathcal{T} est important. Si les intervalles sont trop étendus (plusieurs mois) il ne sera pas possible de suivre les reconfigurations des groupes sociaux à cause d’une agrégation temporelle trop importante. Au contraire, si les intervalles sont trop petits (ex. un jour), les graphes correspondants seront presque vides et statistiquement peu significatifs. Il faut donc faire un compromis entre l’adaptabilité temporelle de la méthode de reconstruction des groupes politiques et sa stabilité.

L’évolution des opinions politiques résultant de l’enchevêtrement de plusieurs processus dynamiques avec des constantes de temps très différentes (ex. reconfiguration des courants à l’intérieur d’un parti, reconfiguration des partis eux-mêmes, évolution du contexte politique européen, etc.), plusieurs échelles de temps peuvent être pertinentes à analyser. Pour cette étude, nous avons choisit une échelle de temps propre à Twitter et aux médias. En analysant la distribution temporelle des retweets sur notre corpus, nous avons observé que l’actualité politique ainsi que l’activité de Twitter ont une périodicité hebdomadaire et que la période de retweet d’un tweet original est en moyenne de l’ordre de la semaine, 98.13% des retweets ayant lieu dans les sept jours (cf. D).

Nous avons donc considéré des périodes de sept jours comme échelle de temps caractéristique pour l’activité politique sur Twitter et opté pour des intervalles T_i de deux semaines afin d’obtenir des périodes recouvrantes. Dans les analyses suivantes, nous avons considéré les communautés calculés tous les lundi avec $\mathcal{G} = \{G_T\}_{T \in \mathcal{T}}$ où \mathcal{T} est un ensemble d’intervalles temporels, tel que $\forall k \in [2, K - 1], |T_k| = 14 \text{ jours}$, $|T_{k-1} \cap T_k| = |T_k \cap T_{k+1}| = 7 \text{ jours}$, lorsque les intervalles sont ordonnées par leur date de début.

Ainsi, chaque tweet de notre jeu de données est pris en compte dans exactement deux graphes de retweets. Comme nous le montrons plus bas, ce choix permet de suivre de manière relativement continue les transformations des communautés politiques et de leur espace sémantique.

4 Résultats

4.1 Les communautés politiques comme support de diffusion des tweets politiques

Le but de notre étude était de démontrer que l’analyse des échanges sur Twitter permet de qualifier et quantifier l’activité des communautés politiques. Le premier résultat concerne les motifs de diffusion et la diversité de l’information politique sur Twitter que la reconstruction des communautés politiques permet d’analyser à un niveau meso.

Nous avons tout d’abord validé l’hypothèse H_0 , intuitive d’après notre méthode de reconstruction, selon laquelle la diffusion des retweets est fortement contrainte par la structure des communautés calculées (cf. annexes E). En analysant un sous corpus de tweets retweetés au moins 50 fois par les communautés³, nous avons montré que sur l’ensemble des périodes, pour 92% des diffusions, H_0 est vraie avec une p -value inférieure à 10^{-16} . En considérant un seuil de p -value de 1%, 99.8% des tweets émis par les communautés ont une diffusion de retweet très différente d’une diffusion aléatoire dans les communautés.

Afin d’avoir une vision plus précise de la manière dont les communautés structurent cette diffusion, nous avons calculé pour chaque jour j et chaque tweet $tw \in \mathcal{D}_{50}^j$ l’entropie \mathcal{H} de sa diffusion en fonction des communautés \mathcal{C}^j , ainsi que, à titre de comparaison, l’entropie de la diffusion de \mathcal{D}_{50}^j dans son ensemble : $\mathcal{H}(P^j(tw)) = - \sum_i P^j(tw)_i \log(P^j(tw)_i)$.

Comme le montre la figure 3, l’entropie de la distribution des retweets au sein des communautés est très faible avec une moyenne inférieure à 1, ce qui indique que la majorité des retweets ont lieu dans une ou deux communautés⁴ alors que pour l’ensemble du corpus de tweets, la distribution des retweets couvre un ensemble beaucoup plus large de communautés (entropie de 2.707 en moyenne correspondant à une diffusion dans 6.532 communautés).

L’analyse de la diffusion des retweets montre donc qu’une grande majorité des tweets sont retweetés à l’intérieur de leur communauté d’origine, comme c’est le cas par exemple pour le tweet de François

3. Ces tweets ont captés 31% de tous les retweets.

4. Exemple, si un tweet diffuse dans une seule communauté alors l’entropie est de 0. S’il diffuse dans 4 communautés de manière uniforme alors l’entropie est de 2.



FIGURE 3 – Motifs de diffusion des tweets. (a) à gauche : Boîte à moustache de l'entropie de la diffusion de chaque tweet dans les communautés et de l'ensemble des tweets dans le réseau ; (b) au centre : exemple de tweet de François Fillon et (c) à droite sa diffusion privilégiée dans certaines communautés.

Fillon présenté dans la figure 3 : “Il y a eu une étude Rexecode qui montre que mon programme permet la création d'un 1,5 million d'emplois. C'est celui qu'il faut appliquer. (779 retweets, entropie=0.245).

Nous retrouvons le rôle d'écho chamber des communautés politiques [CRA14], qui structurent la diffusion de l'information, avec, dans notre cas d'étude, une structuration multipolaire.

4.2 Spécificités thématiques des communautés politiques

Au delà du phénomène de circulation préférentielle des contenus à l'intérieur de leur communauté politique d'origine, les communautés politiques reconstruites par notre méthode ont aussi des spécificités thématiques qui sont cohérentes avec les spécificités thématiques attendues. Si une caractérisation exhaustive des communautés du point de vue des thématiques traitées sort du cadre de cet article, nous avons démontré qu'il y a hétérogénéité dans les thèmes traités et la manière de les traiter.

Nous avons constitué deux corpus représentatifs du discours politique de cette élection présidentielle : une liste de 980 mesures politiques extraites des programmes des 11 candidats à la présidentielle et un corpus de 25 883 tweets publiés par ces candidats et quelques autres personnalités politiques importantes entre août 2016 et janvier 2017.

Nous avons analysé ces deux corpus à l'aide du logiciel de text-mining Gargantext (<http://gargantext.org>) et identifié un vocabulaire de près de 1500 groupes d'expressions qualifiant les thèmes de la campagne présidentielle. Nous avons utilisé Gargantext pour regrouper ces expressions en thèmes afin de produire une carte thématique des sujets discutés par les candidats et leur communauté politique lors de cette campagne (cf. figure 16 de l'annexe G).

A l'aide de cette carte, nous avons défini 11 grands thèmes de cette campagne présidentielle, définis par un ensemble de mots-clés, auxquels nous avons associé 11 corpus de tweets en faisant une requête dans notre base de données de tweets à partir des mots-clés des thèmes (cf. annexes G).

L'analyse de la spécificité du discours politique en fonction de thèmes peut alors se faire à plusieurs niveaux :

- Diversité des partis politiques traitant chaque thème,
- Diversité des thèmes traités par un parti politique donné,
- Diversité des expressions utilisées par chaque communauté politique.

Le calcul pour chaque thème, de l'entropie des communautés traitant de ce thème montre qu'ils sont tous couverts par plusieurs communautés (autour de 6 en moyenne) avec une diversité moindre pour les thèmes *sécurité intérieure* et *immigration et multiculturalisme*, particulièrement investis par la droite et l'extrême droite (cf. table 6 de l'annexe G.2).

Le calcul pour chaque communauté politique, de la diversité des thèmes abordés (cf. table 1 de l'annexe G.2) montre très clairement une hétérogénéité des communautés dans la manière de traiter les thèmes sélectionnés⁵ et une mise à l'agenda politique différenciée des thèmes en fonction des moments de la campagne (cf. analyse de la dynamique des thèmes par communauté figures 25 à 30). Cette hétérogénéité se retrouve également au niveau du vocabulaire employé par les membres de chaque communauté politique (cf. table 2).

La plateforme compagnon de cet article, le *Politoscope*⁶, permet d'explorer l'évolution pour chacun des principaux candidats et chacun des onze thèmes des 15 mots-clés les plus significatifs pour le candidat. On y observe, en plus d'une spécificité sémantique des communautés vis-à-vis des thèmes, une évolution

5. Cette analyse ne recherchant pas l'exhaustivité des thèmes, il n'est pas possible de qualifier les candidats au regard de la diversité des thèmes qu'ils traitent sur la base de cette étude.

6. <https://presidentielle2017.politoscope.org/dashboard>

Nom	Entropy	Nb thèmes	Thèmes Dominant	%
Jean-Luc Mélenchon	3.10	8.6	démocratie, politique étrangère	37.3
François Asselineau	2.63	6.2	politique étrangère, démocratie	57.7
Marine Le Pen	2.78	6.9	immigration et multiculturalisme, politique étrangère	49.0
François Fillon	3.06	8.3	immigration et multiculturalisme, démocratie	33.5
Emmanuel Macron	3.03	8.2	politique étrangère, démocratie	44.2
Benoît Hamon	3.19	9.1	emploi, politique étrangère	39.3
Philippe Poutou	3.08	8.5	démocratie, emploi	38.8
Nathalie Arthaud	2.88	7.4	politique étrangère, démocratie	48.6
Tous candidats	3.08	8.4	politique étrangère, démocratie	36.9

TABLE 1 – Tableau de l’entropie des thèmes abordés par chaque candidats et sa 3–communauté. Les deux thèmes dominants ainsi que leur pourcentage sont également présentés. L’entropie maximale est de 3.46 lors que les 11 thèmes sont discutés de manière homogène. Relativement aux thèmes sélectionnés, Benoît Hamon est le candidat ayant traité les plus de thèmes de manière homogène alors que François Asselineau est celui qui s’est le plus concentré sur un sous-ensemble de thèmes (dont *politique étrangère* et *démocratie*). L’évolution de l’attention des candidats pour ces thèmes est donnée par les figures 25 à 30.

au cours du temps pour les différents candidats, l’usage de certains mot-clés apparaissant, se renforçant ou disparaissant au cours du temps (cf. annexes G.1 pour le détail de la méthode et des exemples).

En conclusion, l’analyse des profils sémantiques des communautés politiques reconstruites par notre méthode montre que celles-ci ont des profils très spécifiques et cohérents avec ce que l’on sait par ailleurs de leurs orientations politiques. Sur Twitter, l’information circule de manière privilégiée à l’intérieur de ces communautés et nous pouvons suivre l’évolution de leur agenda politique au cours du temps en observant les variations de l’attention accordée aux différents thèmes.

4.3 Stabilité des communautés

Les communautés politiques peuvent évoluer en gagnant ou perdant des membres qui ne comptaient pas auparavant parmi les comptes politisés sur Twitter ou en gagnant l’adhésion de membres appartenant à d’autres communautés. Il y a en effet co-évolution entre la structure des communautés et les opinions de leurs membres. Les communautés façonnent la circulation d’information et les membres d’une communauté donnée sont exposés à des contenus spécifiques, mais cette exposition peut participer à l’évolution de leur opinion et éventuellement les amener à changer de communauté, modifiant ainsi la structure globale des communautés.

Afin de qualifier la structure des communautés, nous avons analysé la stabilité des communautés sur \mathcal{G} . Nous avons calculé pour chacun des comptes ayant été au moins une fois dans l’une de ces communautés, la distribution cumulative du temps passé au sein d’une communauté (cf. tableau 3 et figures 31 et 32 en annexe).

Il y a une forte hétérogénéité dans la cohésion communautés, les communautés de droite (Le Pen et Fillon) étant celles qui possèdent les cœurs les plus importants et les plus fidèles (10% des comptes restant plus de 100 jours dans la communauté entre le 6 février et le 23 avril 2017), celles des personnalités qui n’ont pas passé les primaires étant, sans surprise les moins structurées et les moins pérennes.

L’analyse de la distribution cumulative du temps passé dans une communauté montre que pour la plupart des communautés plus de 20% des comptes étant présents moins de 4 jours, et plus de 20% des comptes étant impliqués plus de 40 jours.

Nous avons ainsi une structure générale des communautés politiques sur Twitter, constituées d’un cœur de militants restant plusieurs semaines au sein de la communauté, d’une périphérie beaucoup plus volatile et d’un ensemble intermédiaire couvrant un large spectre d’implication.

Par ailleurs, on constate (cf. table 3) que le temps moyen passé dans une communauté, toutes communautés confondues, est bien supérieur (55.8 j) au temps moyen passé à l’intérieur d’une même communauté (≤ 33.8 j). Cela indique qu’une part importante de comptes change de communauté au cours du temps, ce qui peut indiquer une évolution ou une instabilité de leurs opinions politiques.

4.4 Évolution des communautés et dynamiques d’opinions

Les opinions et les communautés politiques co-évoluent. Les campagnes politiques sont faites de retournements, coups de théâtre et reconfigurations des soutiens politiques. Les commentateurs s’efforcent

Nom	Entropie	Nb keyword	Keyword dominant	%
Jean-Luc Mélenchon	7.53	184.8	loi travail, peuple, travail, république, nuitdebout, sondage, élus, médias, Pénélope Fillon, chômeurs	24.3
François Asselineau	5.98	63.2	frexit, europe, sondage, médias, brexit, peuple, élus, démocratie, traités, l'euro	48.7
Marine Le Pen	7.82	226.2	peuple, islamistes, patriotisme, police, médias, sondage, chômeurs, migrants, élus, terrorisme	20.7
François Fillon	7.75	215.8	médias, sondage, république, redressement, élus, chômeurs, police, peuple, travail, sécurité	19.6
Emmanuel Macron	7.71	209.1	Pénélope Fillon, république, élus, travail, sondage, justice, chômeurs, europe, police, mise_en_examen	19.8
Benoît Hamon	7.77	218.6	république, travail, les jeunes, revenu universel, démocratie, europe, élus, écologie, peuple, enfants	21.1
Philippe Poutou	7.55	187.5	nuitdebout, loi travail, république, travail, Pénélope Fillon, élus, peuple, sondage, chômeurs, médias	26.1
Nathalie Arthaud	7.32	159.2	frexit, sondage, europe, élus, chômeurs, médias, travail, peuple, nuitdebout, république	29.8
All	8.23	301.23	peuple, république, sondage, élus, médias, chômeurs, travail, frexit, police, europe	17.5

TABLE 2 – Tableau de l'entropie des mots-clés utilisés par chaque candidat et sa 3-communauté. Les 10 mots-clés dominants ainsi que leur pourcentage sont également présentés. L'entropie maximale est de 10.73 lors que les 1700 mots-clés sont discutés de manière homogène.

Nom	Personnes présentes	Temps moyen en jours
Marine Le Pen	35 812	33.8
François Fillon	30 516	30.7
Jean-Luc Mélenchon	42 206	21.8
Emmanuel Macron	57 651	16.2
Benoît Hamon	37 748	17.7
François Asselineau	14 774	11.8
Philippe Poutou	51 908	11.6
Nicolas Dupont-Aignan	30 366	5.6
Nathalie Arthaud	45 450	4.5
Jacque Cheminade	29 153	4.4
Global	187 619	55.8

TABLE 3 – Tableau présentant pour chaque candidat-e le nombre de personnes ayant été au moins une fois dans sa communauté ainsi que le temps moyen passé l'intérieur de la communauté. Les communautés Le Pen et Fillon se détachent clairement comme étant les communautés les plus stables avec un temps moyen très important. Le fait que le temps moyen d'un compte dans une communauté (ligne global) soit très élevé indique qu'une proportion significative de comptes ont changé de communauté sans disparaître pour autant.

de les anticiper et de les analyser tandis que les partis politiques s'évertuent à masquer ce qui se trame en coulisse pour prendre leurs adversaires de court. Le principal résultat de notre étude est qu'une analyse automatique des motifs de retweets est en mesure de révéler, parfois avant que la presse n'en parle, l'évolution des communautés politiques face à ces événements et, dans certains cas, face à leur anticipation.

Afin de suivre les reconfigurations des communautés politiques sur des intervalles de temps où les changements sont perceptibles, nous avons comparé deux à deux les communautés calculées sur des intervalles de temps consécutifs. Ce choix permet de suivre de manière continue les transformations des communautés politiques.

Entre chaque période, pour chaque communauté, nous avons calculé le nombre de comptes changeants de communauté et nous avons généré un graphe alluvial représentant ces changements (figure 4). Cette visualisation synthétise les reconfigurations de l'espace politique sur 11 mois au niveau de granularité des principaux acteurs de cette élection présidentielle. On peut constater qu'alors que la plupart des grands événements politiques peuvent se lire sur cette visualisation, celle-ci révèle également des rapports plus subtils entre les communautés (cf. analyses des figures 5, 6, 7, 8 et 9), comme par exemple la prise de distance de la communauté Sarkozy par rapport à la communauté Fillon peu avant le Pénélope Gate puis le ralliement d'un nombre important de militants sarkozistes à Dupont-Aignan. Ce ralliement n'a duré qu'un temps puisque la majorité de ceux-ci ne l'ont pas suivi lorsque Dupont-Aignan s'est lui-même rallié à Le Pen entre les deux tours. Le comportement des différents courants de la droite pendant les primaires des *Républicains* est également très révélateur des lignes de fractures de ce parti.

L'intérêt de cette reconstruction de communautés et de la visualisation qui en découle est qu'elle se situe à un niveau intermédiaire entre les leaders politiques et leur base électorale. En s'appuyant sur le comportement d'un ensemble de militants relevant d'un large spectre d'implication politique, elle donne accès à un niveau d'observation qui est rarement accessible dans la cadre d'études sociologiques ou en sciences politiques. Au total, près de 190.000 comptes uniques ont pu être identifiés comme ayant pris part à une communauté politique, la plupart des communautés ayant impliqué plusieurs dizaines de milliers de militants (cf. table 3). Ces communautés sont constituées d'un nombre important de leaders d'opinion et de relais politiques. Au vu de nos résultats, nous pouvons raisonnablement faire l'hypothèse que les tribulations politiques de sous-populations au sein de cet ensemble sont des indicateurs et parfois des précurseurs au changement et reconfigurations politiques.

5 Conclusions

En s'inspirant d'une définition du groupe social donnée par Gabriel Tarde, nous avons démontré que des méthodes d'analyse des échanges à contenu politique sur de réseaux sociaux tels que Twitter permettaient de reconstruire automatiquement des communautés politiques qui, par leur profil sémantique et les motifs de leur évolution reflètent fidèlement les dynamiques à l'œuvre dans l'espace politique militant.

En appliquant ces méthodes aux données de la présidentielle française 2017, nous avons ainsi reconstruit l'espace politique français multi-polaire, retrouvé les principaux événements de la campagne et mis en évidence les multiples reconfigurations des communautés politiques, dont certaines se sont produites de manière relativement confidentielle.

Utilisant exclusivement des données d'interaction, notre approche a l'avantage d'être indépendante de la langue utilisée pour ces échanges pour peu que l'activité sociale étudiée implique un certain niveau d'engagement des participants, comme c'est le cas en politique.

Cette étude, de part sa durée (11 mois), sa granularité temporelle (capture temps réel des données à l'échelle de la seconde) et son ampleur ($\geq 1,7$ millions de comptes Twitter dont ≥ 180.000 ont interagi suffisamment avec le *Politoscope* pour être catégorisés dans une communauté politique), pose les fondements d'une série d'analyses qui apporteront un éclairage inédit sur les dynamiques d'opinion et les reconfigurations des communautés politiques (voir par exemple [CGP17]). Elle donne accès à un niveau de granularité intermédiaire entre les enquêtes sociologiques de terrain et les études statistiques globales (type INSEE) dont nous anticipons qu'il sera, à l'avenir, une source importante de connaissances politiques et sociologiques venant compléter des approches plus traditionnelles.

Nous concluons cet article par une remarque sur les implications politiques et éthiques du développement de ces méthodes. Depuis quelques années, la frontière entre vie privée et vie publique n'a cessé de s'estomper à mesure que les traces numériques laissées par les citoyens ont été analysées de manière systématique, en croisant un nombre croissant de bases de données (achats, historiques de navigation web, votes, cartes de fidélité, transports, GPS, etc.), en considérant un historique de plus en plus conséquent portant sur des tailles d'échantillon de la population qui tendent à être de l'ordre de grandeur de la population elle-même. Le marketing et la sphère économique ont été les premiers domaines d'activité à s'emparer de ces nouvelles opportunités à des fins de valorisation. Plus récemment, la sphère politique a pris la mesure de ces nouvelles opportunités et a cherché à les mettre à profit afin de mieux connaître les préférences

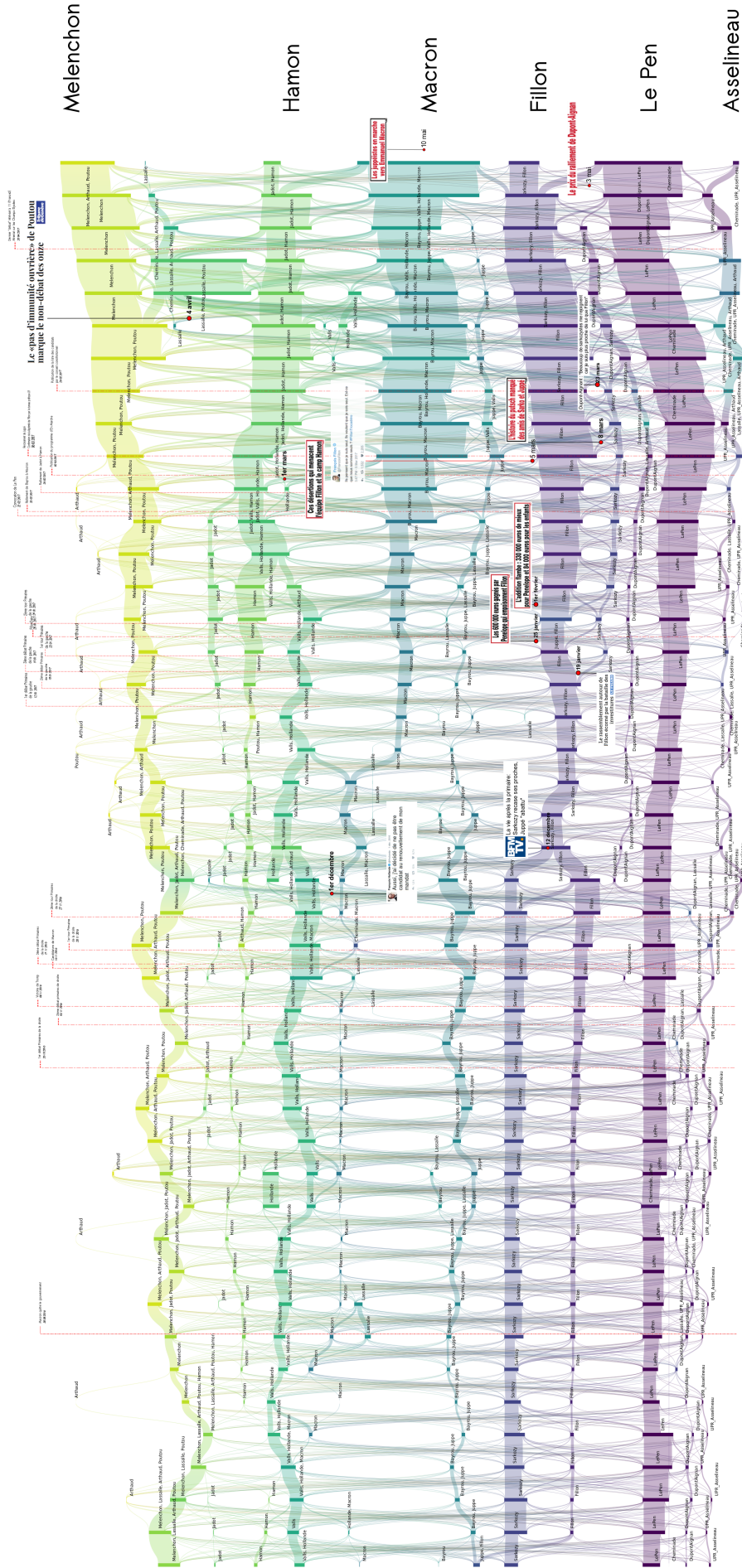


FIGURE 4 – Reconfigurations des communautés politiques entre juin 2016 et avril 2017. Certains grands titres de journaux ont été ajoutés pour apporter un éclairage sur les bifurcations observées. Les labels des communautés encore actives lors du second tour ont été ajoutés en fontes larges pour une meilleure visibilité. Chaque barre verticale correspond à une communauté politique sur Twitter labellisée par les comptes des candidats qui y participent. La hauteur est proportionnelle au nombre de comptes Twitter faisant partie de cette communauté. Celles-ci sont calculées tous les lundis. **A remplacer par le pdf pour le doc final.**

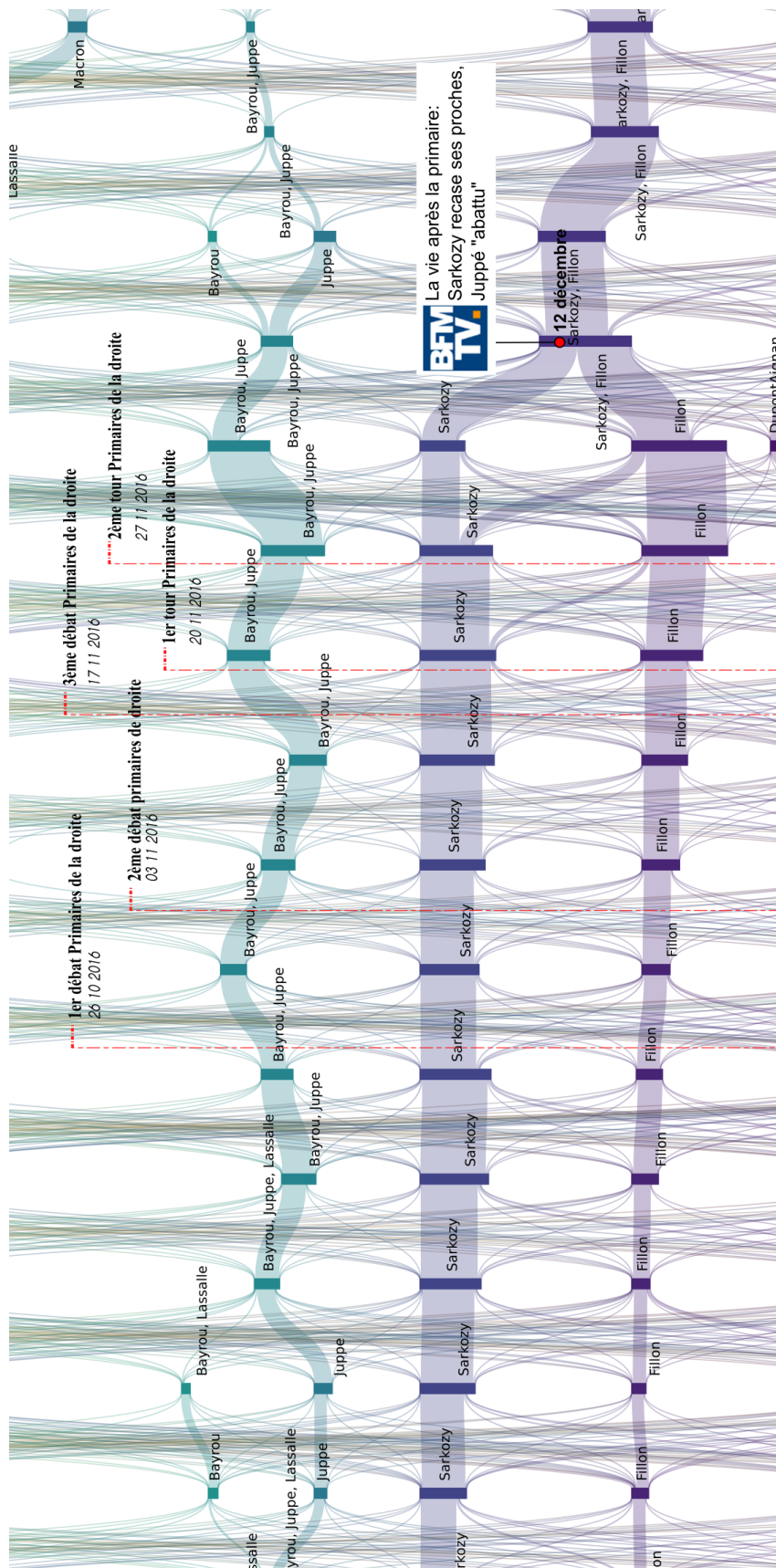


FIGURE 5 – Reconfiguration des communautés politiques au moment des primaires de la droite. Les primaires de la droite ont été rythmées par trois débats et deux votes. On peut observer les renforcements progressifs et réguliers de la communauté Fillon semaines après semaines, alors que personne ne le pronostiquait au second tour de la primaire, Juppé étant le favori des sondages. A l'issue du premier tour, de la victoire surprise de Fillon et de l'élimination également surprenante de Sarkozy (président du parti *Les Républicains* et ex-Président de la République), seule une infime partie de la communauté Sarkozy a rejoint la communauté Fillon. La communauté Sarkozy traînera des pieds pour se rallier à la communauté Fillon et Sarkozy négociera pendant plusieurs jours la place de ses membres dans le dispositif Fillon. La communauté Juppé, en revanche, prendra immédiatement ses distances par rapport aux nouvelles orientations politiques prises par le parti. Cette ligne de fracture de la droite était apparente dans le paysage politique pré-électoral (cf. figure 15). Après l'échec du retour de Juppé en mars 2017 (cf. figure 8), une partie de la communauté Juppé rejoindra Macron. Les bifurcations observées ici sont bien résumées par le titre d'un article de BFM TV du 12 décembre 2017 : *La vie après la primaire : Sarkozy recase ses proches, Juppé "abattu"*. Chaque barre verticale correspond à une communauté politique sur Twitter labellisée par les comptes des candidats qui y participent. La hauteur est proportionnelle au nombre de comptes Twitter faisant partie de cette communauté. Celles-ci sont calculées tous les lundis.

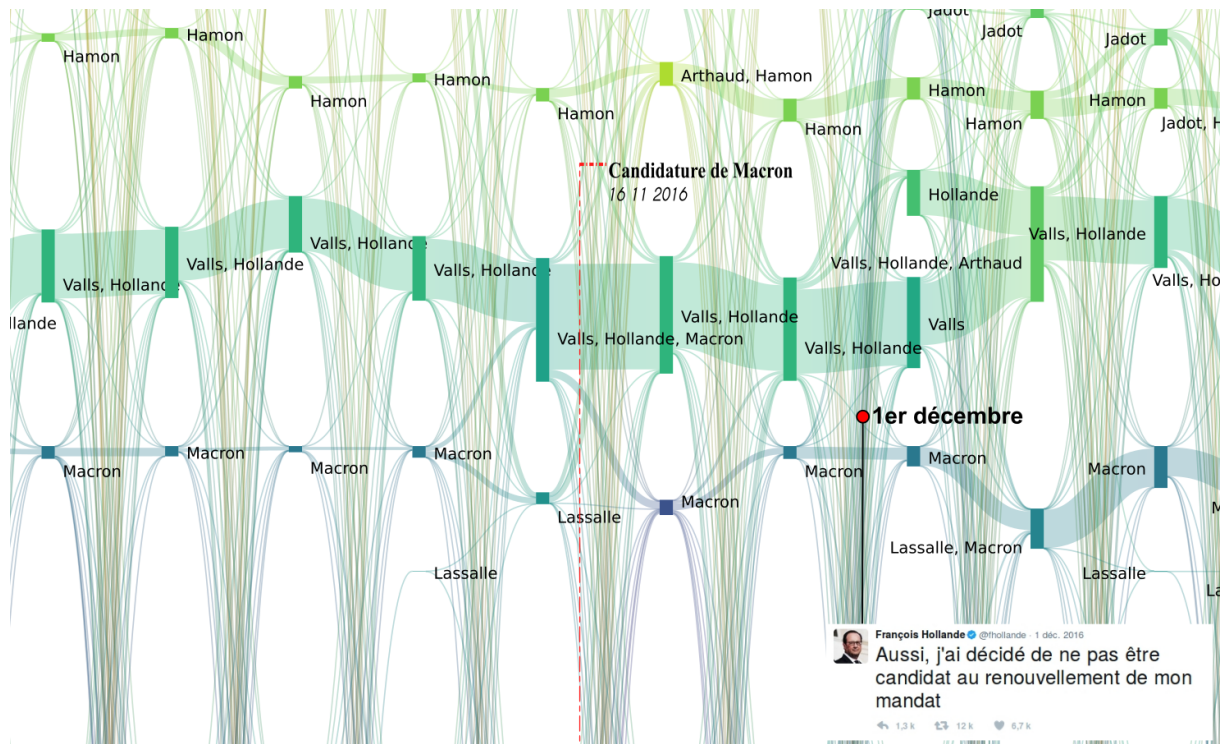


FIGURE 6 – Reconfiguration des communautés politiques au moment où le Président de la République François Hollande annonce qu’il ne se représentera pas. Cette déclaration a eu lieu peu après que Macron ait annoncé sa candidature. Il apparaît clairement sur le graphe alluvial que la non candidature de Hollande a dynamisé la communauté de Macron au détriment de celle de Valls. Alors que la taille de la communauté Macron était restée stable au moment de l’annonce de sa candidature (427 comptes la semaine précédente, 561 la semaine de l’annonce, 472 la semaine suivante), elle a bondit à 749 comptes la semaine du renoncement de Hollande, puis 1,479 comptes la semaine suivante. Elle a dès lors été en croissance régulière jusqu’aux élections. Cette évolution est concomitante avec une perte de vitesse de la communauté Valls-Hollande, aile droite du *Parti Socialiste* la plus proche de Macron d’un point de vue idéologique. La communauté Hamon, plus à gauche, est en revanche restée relativement stable sur cette période. Chaque barre verticale correspond à une communauté politique sur Twitter labellisée par les comptes des candidats qui y participent. La hauteur est proportionnelle au nombre de comptes Twitter faisant partie de cette communauté. Celles-ci sont calculées tous les lundis.

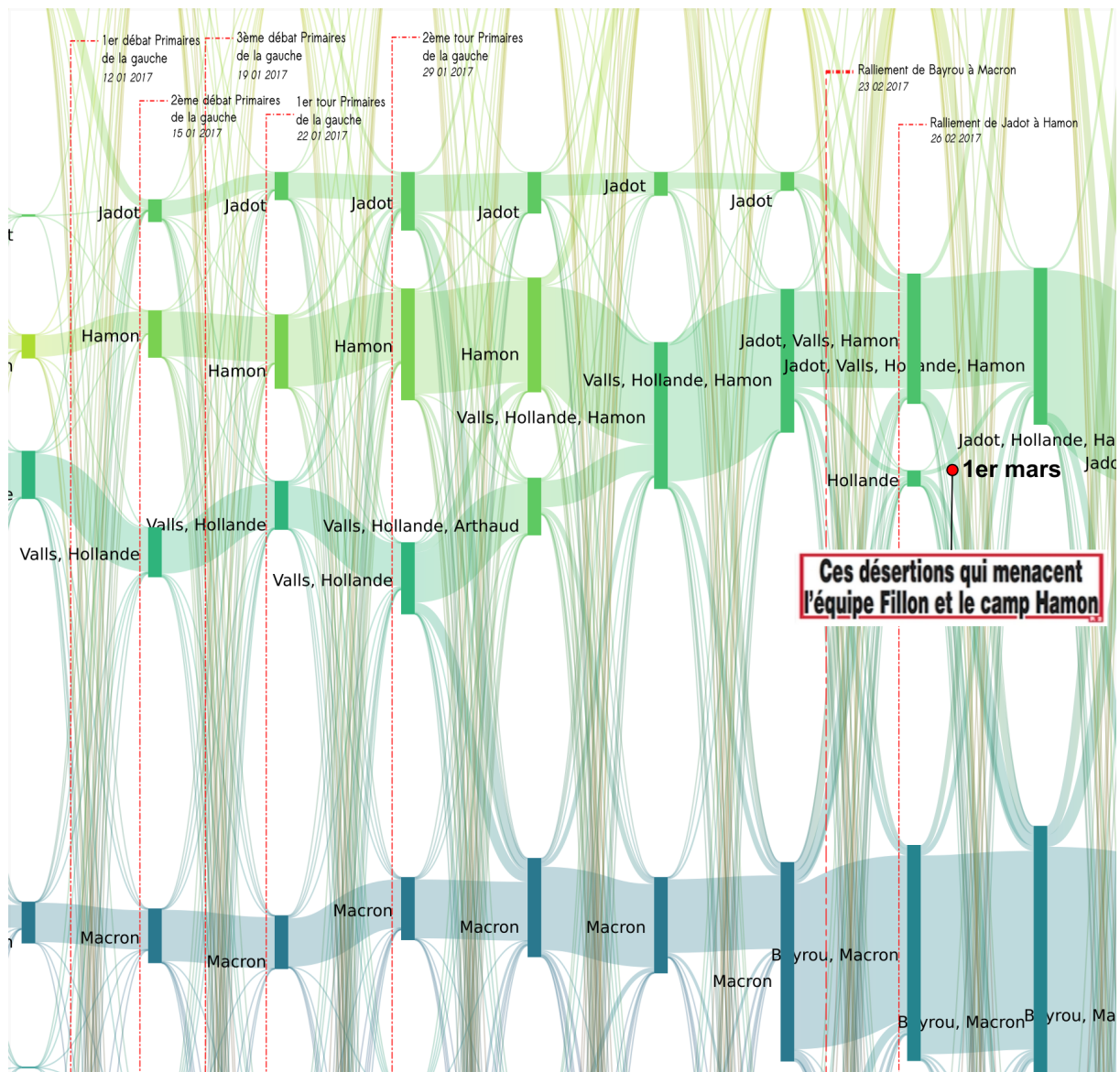


FIGURE 7 – Reconfiguration des communautés politiques au moment des primaires de la gauche. Benoît Hamon (36,03% au 1er tour et 58,69% au second tour) et Manuel Valls (31,48% au 1er tour et 41,31% au second tour) étaient les deux favoris de ces primaires et représentaient deux courants du parti socialiste qualifiés par Valls de “gauches irréconciliables” lors de sa campagne. Bien que François Hollande ne se soit pas exprimé officiellement en faveur d’un candidat particulier, Valls affirmait que celui-ci le soutenait (cf. par exemple *Le Monde* du 4 janvier 2017 “Selon Valls, Hollande le soutient”). En effet, sur notre reconstruction, la communauté Hamon est, pendant toute notre période d’observation, bien distincte de la communauté Valls. Cette dernière est quant à elle, la plupart du temps également la communauté de Hollande, notamment au moment des primaires, indiquant une proximité idéologique entre les deux personnalités politiques. Après la victoire de Hamon, une polémique éclata sur l’engagement pris par les candidats à cette primaire de se ranger derrière le vainqueur alors que certains affichaient clairement leur préférence pour Macron. Ce flottement est clairement visible sur le graphe alluvial avec des mouvements importants de militants de la communauté Valls-Hollande vers la communauté Macron dès le lendemain du deuxième tour de la primaire. Sur les 2075 comptes que comportait la communauté Valls-Hollande juste avant sa fusion avec la communauté Hamon le 13 février (Valls rejoignant également la communauté Hamon), seuls 959 comptes resteront. Des mouvements importants de militants continueront d’avoir lieu entre ces deux communautés jusqu’au premier tour de la présidentielle. Si Valls finira par soutenir officiellement Macron le 29 mars, il n’apparaissait déjà plus dans la communauté Hamon dès le 13 mars. Il apparaît dans la communauté Macron à partir du 17 avril et jusqu’au deuxième tour, indiquant un soutien actif de sa part. On remarquera également sur ce zoom le ralliement de l’écologiste Jadot à Hamon le 23 février 2017, qui se manifeste immédiatement par la fusion de sa communauté avec celle de Hamon. Chaque barre verticale correspond à une communauté politique sur Twitter labellisée par les comptes des candidats qui y participent. La hauteur est proportionnelle au nombre de comptes Twitter faisant partie de cette communauté. Celles-ci sont calculées tous les lundis.

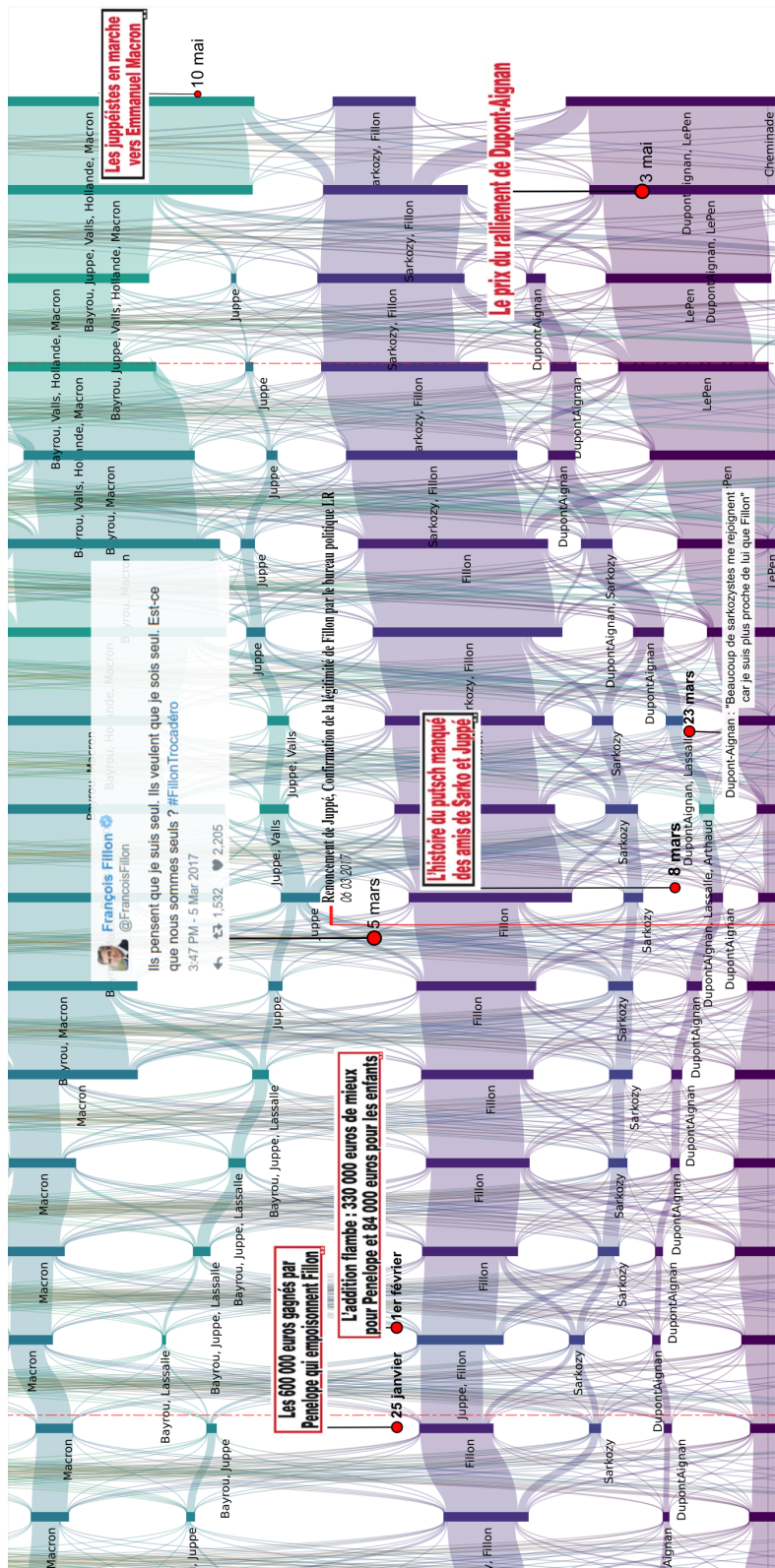


FIGURE 8 – Reconfiguration des communautés politiques au moment du Pénélope Gate. Le Pénélope Gate arrive à un moment où le rassemblement autour de Fillon est fragilisé en raison d’une polémique sur les investitures aux législatives, mal vécues par les sarkozystes. Une partie de ceux-ci avait déjà commencé à se détacher de la communauté *Les Républicains* dès le 23 janvier 2017 et formeront une communauté à part pendant toute la polémique. Les différentes révélations du *Canard Enchaîné* provoquent de larges mouvements de militants entre différentes communautés, les communautés autour de Juppé (potentielle alternative à la candidature Fillon ou “plan B”) et Sarkozy se renforçant progressivement. Au 1er mars, jour de la convocation de Fillon en vue d’une mise en examen, une grande partie des Sarkozystes soutiennent le retour de Juppé dans un ultime effort pour pousser Fillon à la démission, ce qui se manifeste par un flux inhabituel de comptes de la communauté Sarkozy vers la communauté Fillon. Le rassemblement du Trocadéro le 5 mars met fin aux perspectives de plan B. La communauté Fillon gagne beaucoup de nouveaux militants suite à ce coup de force mais une partie des sarkozistes et des juppéistes restent à l’écart et rejoindront respectivement, Dupont-Aignan et Macron. Sarkozy quant à lui, restera plusieurs semaines en retrait, ce qui se traduit par son absence dans la plupart des communautés Fillon entre le 16 janvier et le 10 avril. Ce n’est qu’à partir du 7 avril, qu’il soutiendra à nouveau activement la candidature Fillon, comme le mentionne par exemple le titre du *Parisien* de ce jour “Présidentielle : Sarkozy soutient (enfin) Fillon” et ce qui se manifeste par sa présence permanente dans la communauté Fillon à partir du 10 avril 2017. Comme on peut le constater, les bifurcations observées dans le Politoscope sont concomitantes et même parfois anticipent les commentaires de la presse sur l’activité des communautés politiques. Chaque barre verticale correspond à une communauté politique sur Twitter labellisée par les comptes des candidats qui y participent. La hauteur est proportionnelle au nombre de comptes Twitter faisant partie de cette communauté. Celles-ci sont calculées tous les lundis. 15

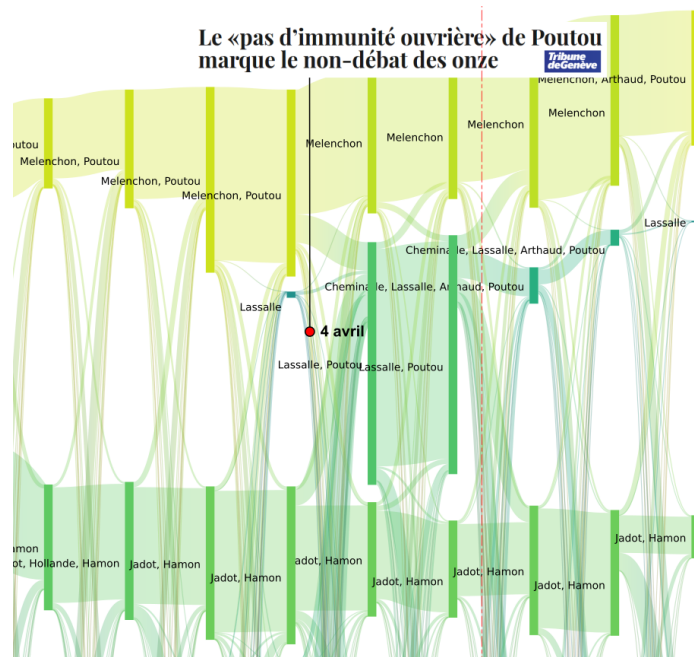


FIGURE 9 – Reconfiguration des communautés politiques au moment du “Grand Débat” télévisé du 4 avril 2017. Philippe Poutou a créé le buzz lors de cette confrontation en dénonçant ouvertement la triche et la corruption des candidats Fillon et Le Pen et en critiquant les positions anti-système : “Quand nous on est convoqué par la police, on n’a pas d’immunité ouvrière, on y va.” Son intervention a été relayée par la plupart des médias et a recueilli plus d’un million de vues en ligne. Sa communauté Twitter, quasi inexistante jusque là et souvent fusionnée avec celle de Mélenchon, s’est développée considérablement (> 13,000 comptes) sur une courte période suite à cette intervention.

et orientations politiques des électeurs. En témoigne la fuite survenue en juillet 2017 de données personnelles sensibles de près de 200 millions d’américains compilées pour le compte du parti Républicain ($\approx 99\%$ du corps électoral)⁷. Nous avons pris soin sur l’ensemble du projet *Politoscope* d’anonymiser les données traitées et de ne livrer que des données agrégées, rendant impossible l’identification d’individus particuliers mis à part quelques personnalités politiques connues. Les travaux présentés dans cet article démontrent cependant qu’il est possible, à partir de traces laissées ostensiblement dans l’espace public et avec des moyens computationnels relativement accessibles, de catégoriser une proportion significative de l’électorat d’un pays en fonction de ses orientations politiques, même si, comme nous l’avons souligné, une certaine marge d’erreur persiste. Dès lors, il est important de prendre conscience que toute trace laissée sur le web engage son auteur-e de ce point de vue. La plateforme ouverte *politoscope.org* a notamment pour but d’accompagner cette prise de conscience. Nous ne pouvons compter que sur l’honnêteté et la bonne foi des dirigeants politiques actuels et futurs pour ne pas utiliser, comme cela a pu se produire maintes fois par le passé et encore actuellement dans certains pays, des données de ce type à des fins de manipulation, de discrimination, voire de répression ou de persécution.

7. voir <https://www.upguard.com/breaches/the-rnc-files>

Remerciements

Ces recherches ont reçu le soutien de l’Institut des Systèmes Complexes de Paris Île-de-France (<http://isc-pif.fr>), du projet *Algodiv* (ANR-15-CE38-0001), du projet *Science en Poche* (programme *Emergence(s)*, Ville de Paris) et du DIM de la Région Île-de-France *Problématiques Transversales aux Systèmes Complexes*.

Références

- [AG05] Lada A. Adamic and Natalie Glance. The Political Blogosphere and the 2004 U.S. Election. *Proceedings of the 3rd international workshop on Link discovery - LinkKDD '05*, pages 36–43, 2005.
- [AVC⁺16] Benjamin R C Amor, Sabine I Vuik, Ryan Callahan, Ara Darzi, Sophia N Yaliraki, and Mauricio Barahona. Community detection and role identification in directed networks : understanding the Twitter network of the care. data debate. *Dynamic Networks and Cyber-Security*, 1 :111, 2016.
- [Bar15] Pablo Barberá. Birds of the same feather tweet together : Bayesian ideal point estimation using twitter data. *Political Analysis*, 23(1) :76–91, jan 2015.
- [BG15] François Briatte and Ewen Gallic. Recovering the French Party Space from Twitter Data. In *Science Po Quanti*, pages 1–23, Paris, France, 2015.
- [BGL10] Danah Boyd, Scott Golder, and Gilad Lotan. Tweet, tweet, retweet : Conversational aspects of retweeting on twitter. In *Proceedings of the Annual Hawaii International Conference on System Sciences*, 2010.
- [BGLL08] Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics : Theory and Experiment*, 2008(10) :P10008, oct 2008.
- [BHMW11] Eytan Bakshy, Jake M Hofman, Winter A Mason, and Duncan J Watts. Everyone’s an influencer : quantifying influence on twitter. In *Proceedings of the fourth ACM international conference on Web search and data mining*, pages 65–74. ACM, 2011.
- [BHY12] A. Boutet, Hyoungshick Kim, and E. Yoneki. What’s in Twitter : I Know What Parties are Popular and Who You are Supporting Now! In *2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, pages 132–139. IEEE, aug 2012.
- [BJN⁺15] Pablo Barberá, John T Jost, Jonathan Nagler, Joshua A Tucker, and Richard Bonneau. Tweeting From Left to Right : Is Online Political Communication More Than an Echo Chamber? *Psychological science*, 26(10) :1531–42, 2015.
- [BR15] Pablo Barberá and Gonzalo Rivero. Understanding the Political Representativeness of Twitter Users. *Social Science Computer Review*, 33(6) :712–729, 2015.
- [CGP17] David Chavalarias, Noé Gaumont, and Maziyar Panahi. Sur Twitter, lepénistes prosélytes et hamonistes idéalistes. *Libération.fr*, April 2017.
- [CGR⁺11] Michael D. Conover, Bruno Gonçalves, Jacob Ratkiewicz, Alessandro Flammini, and Filippo Menczer. Predicting the political alignment of twitter users. In *Proceedings - 2011 IEEE International Conference on Privacy, Security, Risk and Trust and IEEE International Conference on Social Computing, PASSAT/SocialCom 2011*, pages 192–199, 2011.
- [CHBG10] Meeyoung Cha, Hamed Haddai, Fabricio Benevenuto, and Krishna P Gummadi. Measuring User Influence in Twitter : The Million Follower Fallacy. *International AAAI Conference on Weblogs and Social Media*, pages 10–17, 2010.
- [CM15] Darko Cherepnalkoski and Igor Mozetic. A retweet network analysis of the European Parliament. In *2015 11TH INTERNATIONAL CONFERENCE ON SIGNAL-IMAGE TECHNOLOGY & INTERNET-BASED SYSTEMS (SITIS)*, pages 350–357, 2015.
- [CM16] Darko Cherepnalkoski and Igor Mozetič. Retweet networks of the European Parliament : evaluation of the community structure. *Applied Network Science*, 1(1) :2, 2016.
- [CR13] Raviv Cohen and Derek Ruths. Classifying Political Orientation on Twitter : It’s Not Easy! In *International AAAI Conference on Web and Social Media*, pages 91–99, 2013.
- [CRA14] Elanor Colleoni, Alessandro Rozza, and Adam Arvidsson. Echo chamber or public sphere? predicting political orientation and measuring political homophily in twitter using big data. *Journal of Communication*, 64(2) :317–332, 2014.
- [DOG15] David Darmon, Elisa Omodei, and Joshua Garland. Followers are not enough : A multifaceted approach to community detection in online social networks. *PLoS ONE*, 10(8), 2015.

- [FH16] Santo Fortunato and Darko Hric. Community detection in networks : A user guide. *Physics Reports*, 2016.
- [FSB06] Tanja Falkowski, Myra Spiliopoulou, and J Bartelheimer. Community dynamics mining. In *Proceedings of 14th European Conference on Information Systems (ECIS 2006)*. Citeseer, 2006.
- [GDGM16] Kiran Garimella, Gianmarco De Francisci Morales, Aristides Gionis, and Michael Mathioudakis. Quantifying Controversy in Social Media. In *Proceedings of the Ninth ACM International Conference on Web Search and Data Mining - WSDM '16*, pages 33–42, New York, New York, USA, 2016. ACM Press.
- [Gre15] Kay Green. The Game Changer : Social Media and the 2016 Presidential Election. *The Huffington Post*, November 2015. 00002 bibtext : green_game_2015.
- [GS15] F. Guerrero-Sole. Community Detection in Political Discussions on Twitter : An Application of the Retweet Overlap Network Method to the Catalan Process Toward Independence. *Social Science Computer Review*, 35(2) :244 – 261, 2015.
- [GSAM17] Pedro Calais Guerra, Roberto C. S. N. P. Souza, Renato M. Assunção, and Wagner Meira. Antagonism also Flows through Retweets : The Impact of Out-of-Context Quotes in Opinion Polarization Analysis. *arXiv preprint arXiv :1703.03895*, mar 2017.
- [HKKS04] John Hopcroft, Omar Khan, Brian Kulis, and Bart Selman. Tracking evolving communities in large linked networks. *Proceedings of the National Academy of Sciences of the United States of America*, 101 Suppl :5249–5253, apr 2004.
- [Hub15] Mark Huberty. Can we vote with our tweet ? On the perennial difficulty of election forecasting with social media. *International Journal of Forecasting*, 31(3) :992–1007, 2015.
- [Kan14] Rushed Kanawati. Seed-centric approaches for community detection in complex networks. In Gabriele Meiselwitz, editor, *6th international conference on Social Computing and Social Media*, volume 8531, pages 197–208. Springer International Publishing, 2014.
- [KLPM10] Haewoon Kwak, Changhyun Lee, Hosung Park, and Sue Moon. What is Twitter, a social network or a news media ? In *Proceedings of the 19th international conference on World wide web - WWW '10*, page 591, 2010.
- [MBLB15] A. J. Morales, J. Borondo, J. C. Losada, and R. M. Benito. Measuring political polarization : Twitter shows the two sides of Venezuela. *Chaos*, 25(3), 2015.
- [MGFL16] Angélique Martin, Nicolas Guéguen, and Jacques Fischer-Lokou. L'imitation humaine : Une synthèse de 50 années de recherche en psychologie sociale. *Canadian Psychology/Psychologie canadienne*, 57(2) :101, 2016.
- [MMW⁺15] Panagiotis Metaxas, Eni Mustafaraj, Kily Wong, Laura Zeng, Megan O’Keefe, and Samantha Finn. What Do Retweets Indicate ? Results from User Survey and Meta-Review of Research. In *Ninth International AAAI Conference on Web and Social Media*, 2015.
- [PBV07] Gergely Palla, Albert-László Barabási, and Tamás Vicsek. Quantifying social group evolution. *Nature*, 446(7136) :664–667, 2007.
- [PP11] Marco Pennacchiotti and Ana-Maria Popescu. A Machine Learning Approach to Twitter User Classification. *Icwsn*, 11(1) :281–288, 2011.
- [RY10] Delip Rao and David Yarowsky. Detecting Latent User Properties in Social Media. *Proc. of the NIPS MLSN Workshop*, pages 1–7, 2010.
- [Tar90] Gabriel Tarde. *Les lois de l’Imitation*. Jean-Marie Tremblay, Les classiques des Sciences Sociales edition, 1890.
- [Tom09] Michael Tomasello. *The cultural origins of human cognition*. Harvard university press, 2009.
- [XKS13] Jierui Xie, Stephen Kelley, and Boleslaw K. Szymanski. Overlapping community detection in networks. *ACM Computing Surveys*, 45(4) :1–35, aug 2013.

ANNEXES

Identification and reconfiguration of political communities during 2017 French presidential elections - socio-semantic networks of Twitter political activists

NOÉ GAUMONT, MAZIYAR PANAHİ & DAVID CHAVALARIAS

Note : Certaines figures de cet article sont disponibles en haute définition ou en version interactive sur <https://iscpif.fr/projects/politoscope>

Table des matières

1	Introduction	2
2	Travaux existants	2
3	Identification des groupes sociaux de la twittosphère politique française	3
3.1	Les communautés de retweets	4
3.2	Période caractéristique de stabilité des communautés	6
4	Résultats	6
4.1	Les communautés politiques comme support de diffusion des tweets politiques	6
4.2	Spécificités thématiques des communautés politiques	7
4.3	Stabilité des communautés	8
4.4	Évolution des communautés et dynamiques d’opinions	8
5	Conclusions	10
A	Principaux événements politiques de la campagne présidentielle française de 2017	SI 2
B	Dispositif de collecte des données	SI 3
B.1	Collecte centrée sur les comptes des leaders politiques	SI 3
B.2	Collecte centrée sur les termes	SI 3
B.3	Liste des principales personnalités politiques suivies	SI 5
B.4	Architecture	SI 5
C	Détection de bots	SI 7
D	Temps de vie des tweets et patterns d’activité de Twitter	SI 7
E	Correlation entre structure de communautés et motifs de retweets	SI 9
F	État des lieux du paysage politique avant le début de la campagne présidentielle 2017	SI 9
G	Caractérisation des thèmes de la campagne présidentielle 2017	SI 11
G.1	Évolution du vocabulaire des communautés	SI 11
G.2	Définition des thèmes par requêtes	SI 11
G.3	Distribution des thèmes par communautés politiques	SI 16
H	Stabilité des communautés	SI 20
H.1	Vocabulaire du discours politique de l’élection présidentielle 2017	SI 21
I	Analyse des retweets tardifs	SI 30

A Principaux événements politiques de la campagne présidentielle française de 2017

Les élections présidentielles françaises de 2017 ont été exceptionnellement mouvementée avec de nombreux retournements de situation en des laps de temps très courts et l'effondrement des deux principaux candidats (Fillon et Hamon) et partis ayant dominé la scène politique depuis plus de quarante ans. Nous en présentons ici les principaux jalons. Deux événements internationaux majeurs se sont également déroulés de manière concomitante qui ont surpris tous les analystes et ouvert le champ des possibles pour les futures élections : le Brexit et l'élection de Donald Trump.

Au printemps 2016, beaucoup anticipaient une alternance gauche-droite, Alain Juppé étant le grand favori de tous les sondages.

- **2017/06/12** : Interview de Juppé sur TF1 déplorant son surnom "Ali Juppé" (action de dénigrement lancée par l'extrême droite puis reprise par certains membres des communautés Sarkozy et Fillon pendant les primaires de la droite),
- **2017/06/23** : Le Royaume-Uni vote le Brexit,
- **2016/07/14** : Attentat de Nice,
- **2016/08/28** : François Fillon, candidat à la primaire de la droite, a attaqué son rival Nicolas Sarkozy sur le volet judiciaire en posant la question "Qui imagine un seul instant le général de Gaulle mis en examen?". Cette phrase sera ressortie à plusieurs reprises pendant le Pénélope Gate,
- **2016/08/30** : Macron quitte le gouvernement
- **2016/10/13** : 1er débat primaires de la droite, Sarkozy et Juppé sont les favoris des sondages,
- **2016/11/08** : Élection de Donald Trump ,
- **2016/11/03** : 2ème débat primaires de la droite,
- **2016/11/16** : Annonce de la candidature Macron,
- **2016/11/17** : 3ème débat primaires de la droite,
- **2016/11/20** : Premier tour primaires de la droite,
- **2016/11/27** : Deuxième tour primaires de la droite,
- **2016/12/01** : Hollande renonce à se représenter,
- **2017/01/12** : 1er débat primaires de la gauche,
- **2017/01/15** : 2ème débat primaires de la gauche,
- **2017/01/17** : Frictions chez les *Républicains* après des investitures aux législatives contestées, dont celle de Nathalie Kosciusko-Morizet (NKM), parachutée chez dans la circonscription de Fillon réputée imperdable,
- **2017/01/19** : 3ème débat primaires de la gauche. La présidente du Front national est créditée de 25% à 26% d'intentions de vote au premier tour, selon un sondage Ipsos publié par Le Monde. De son côté, Emmanuel Macron se rapproche d'un François Fillon en perte de vitesse,
- **2017/01/20** : Prise de fonction de Trump. Charge de Rachida Dati (ex. Garde de Sceaux sous Sarkozy) contre NKM,
- **2017/01/21** : Maurice Leroy (UDI et soutien de Sarkozy pendant la primaire) met en garde François Fillon,
- **2017/01/22** : 1er tours primaires de la gauche,
- **2017/01/25** : Le *Canard Enchaîné* dévoile les premières informations sur ce qui sera appelé le *Pénélope Gate*,
- **2017/01/29** : 2ème tours primaires de la gauche,
- **2017/02/21** : Fillon annonce son plan santé, revirement par rapport à ses déclarations précédentes,
- **2017/02/22** : Annonce de la convocation de Marine Le Pen par les policiers de l'Office anti-corruption de la police judiciaire (OCLCJFF), elle ne s'y rendra pas,
- **2017/02/23** : Ralliement de Bayrou à Macron,
- **2017/02/26** : Ralliement de Jadot à Hamon,
- **2017/03/01** : Convocation de Fillon pour une mise en examen, plusieurs ténors des *Républicains* appellent à un plan B. Juppé revient sur le devant de la scène comme possible remplaçant de Fillon,
- **2017/03/02** : Publication du programme d'*En Marche!*,
- **2017/03/03** : Convocation de Marine Le Pen en vue d'une possible mise en examen (affaire des assistants parlementaires),
- **2017/03/05** : François Fillon organise un rassemblement au Trocadéro. Il annonce qu'il y avait 200.000 personnes (dont des ténors soutiens de Sarkozy) et présente cela comme un soutien au maintien de sa candidature,
- **2017/03/06** : Juppé annonce qu'il renonce à toute perspective d'être candidat à la présidentielle. Confirmation de la légitimité de la candidature de Fillon par le bureau politique des *Républicains*,
- **2017/03/20** : Publication de la liste officielle des candidats par le Conseil Constitutionnel,
- **2017/03/20** : Débat TF1 entre les cinq "principaux" candidats à la présidentielle (Mélenchon, Hamon, Macron, Fillon, Le Pen),
- **2017/03/29** : Valls annonce qu'il votera Macron,
- **2017/04/04** : Grand débat sur BFMTV à onze candidats. Poutou se fait remarquer avec sa charge contre Le Pen et Fillon à propos de leurs "affaires",
- **2017/04/16** : Une commission du Parlement européen pourrait convoquer Marine Le Pen avant le second tour de la présidentielle,
- **2017/04/20** : Attentat des Champs Élysées pendant l'émission politique sur la chaîne France 2 alors que les onze candidats étaient invités à s'exprimer. Le Pen radicalise son discours avec une polémique : "la colonisation a beaucoup apporté, notamment à l'Algérie",
- **2017/04/23** : Premier tour de la présidentielle et victoire de Macron et Le Pen. Mélenchon refuse de donner des consignes de vote à ses électeurs, contrairement à ce qu'il avait fait en 2002 où il avait appelé à faire barrage contre le Front National,
- **2017/04/28** : Dupont-Aignan annonce son ralliement à Le Pen,
- **2017/05/03** : Contre-performance de Le Pen lors du débat télévisé Macron-Le Pen de l'entre deux tours.

B Dispositif de collecte des données

Notre objectif étant d'étudier l'organisation de l'espace politique à l'occasion de la présidentielle française de 2017, nous avons choisi de collecter les données Twitter via deux méthodes complémentaires.

B.1 Collecte centrée sur les comptes des leaders politiques

Nous avons établi une liste de plus de 3700 comptes de personnalités politiques françaises (candidats aux primaires, députés, sénateurs, personnes ayant un mandat important au sein d'un parti). Nous avons suivi chacun de ces comptes grâce à l'API 'follow' de Twitter (voir encadré).

La documentation de l'API donne la description exacte des contenus collectés pour le suivi du compte Twitter d'un utilisateur particulier ('user') :

*From Twitter **follow** API documentation*^a. For each user specified, the stream will contain :

- Tweets created by the user.
- Tweets which are retweeted by the user.
- Replies to any Tweet created by the user.
- Retweets of any Tweet created by the user.
- Manual replies, created without pressing a reply button (e.g. "@twitterapi I agree").

The stream will not contain :

- Tweets mentioning the user (e.g. "Hello @twitterapi!").
- Manual Retweets created without pressing a Retweet button (e.g. "RT @twitterapi The API is great").
- Tweets by protected users.

^a. <https://dev.twitter.com/streaming/overview/request-parameters>

Ainsi, si l'on note $T(X, A)$ le fait que le compte X tweete le contenu A , nous avons récolté, dans la limite des volumes autorisés par les API de Twitter, pour chaque compte X suivi, et tout compte Y les contenus $T(X, A)$, $T(Y, T(X, A))$ et, dans le cas où A contient une réponse à X ("@X ...") tous les $T(Y, A)$ (cf. fig. 10).

B.2 Collecte centrée sur les termes

La méthode de collecte centrée sur comptes via l'API 'follow' ne permet pas de collecter les tweets faisant référence à des personnes en dehors des cas mentionnés par Twitter. Pour pallier à cette limitation et obtenir les tweets qui citent un candidats important dans le corps du texte, nous avons constitué une liste de termes liés aux candidats (screenname, nom de parti, etc.) et utilisé l'API 'track' qui permet de capturer tous les tweets mentionnant une expression donnée. Plusieurs hashtags ont été créés spécifiquement à l'occasion de la présidentielle pour catégoriser les tweets pertinents pour cet événement. Nous en avons recensé la plupart et les avons ajouté à la liste d'expressions suivies afin de collecter les tweets les mentionnant (ex. #legranddebat [1,360,502 mentions], #2017ledebat [761,400 mentions]).

*From Twitter **track** API documentation*^a.
A comma-separated list of phrases which will be used to determine what Tweets will be delivered on the stream. A phrase may be one or more terms separated by spaces, and a phrase will match if all of the terms in the phrase are present in the Tweet, regardless of order and ignoring case. By this model, you can think of commas as logical ORs, while spaces are equivalent to logical ANDs (e.g. 'the twitter' is the AND twitter, and 'the,twitter' is the OR twitter).
The text of the Tweet and some entity fields are considered for matches. Specifically, the text attribute of the Tweet, expanded_url and display_url for links and media, text for hashtags, and screen_name for user mentions are checked for matches.

^a. <https://dev.twitter.com/streaming/overview/request-parameters>

La liste

des expressions suivies par l'API track est la suivante :

- | | | |
|-----------------------|---------------|-------------------|
| — présidentielle2017 | — frexit | — #Macron |
| — présidentielle2017 | — Nuit Debout | — #PenelopeGate |
| — #presidentielle2017 | — nuitdebout | — #Hamon2017 |
| — primaires2016 | — #nuitdebout | — #Valls2017 |
| — loitravail | — #Fillon | — #Montebourg2017 |

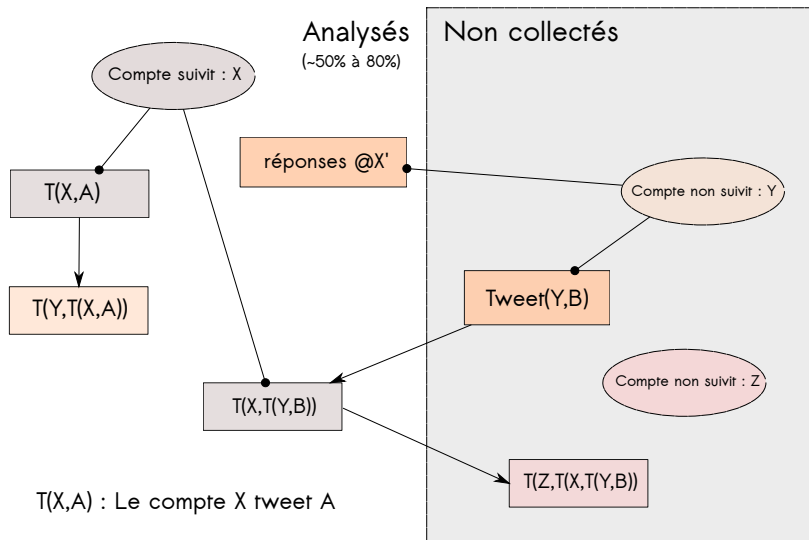


FIGURE 10 – Description du dispositif de capture centré sur les leaders politiques. Celui-ci est complété par un dispositif de capture centré sur les mots-clé (Twitter track API).

- | | | |
|-------------------------|-----------------------|-------------------|
| — #Marine2017 | — n.arthaud | — avote |
| — #Fillon2017 | — montebourg | — avoté |
| — #Peillon2017 | — UPR_Asselineau | — sansmoiLe7mai |
| — #Asselineau2017 | — jfpoisson78 | — Debat2017 |
| — #Macron2017 | — JLM2017 | — 2017LeDebat |
| — #Marinelepen2017 | — #NDA2017 | — jamaismacron |
| — #Mlp2017 MLP_officiel | — dupontaignan | — jamaislepen |
| — FrancoisFillon | — jeanlassalle | — jamaismarine |
| — fhollande | — PhilippePoutou | — jamaislefn |
| — manuelvalls | — n.arthaud | — JeVoteMarine |
| — alainjuppe | — manuelvalls | — JeVoteMacron |
| — nk_m | — #legranddebat | — ChoisirLaFrance |
| — JLMelenchon | — presidentielles2017 | — Dupont-Aignan |
| — NicolasSarkozy | — présidentielles2017 | — enmarche |
| — benoithamon | — présidentielle | — frontnational |
| — EmmanuelMacron | — présidentielle | — #mfga |
| — yjadot | —jevoté | |

L'API de Twitter ne donne accès qu'à un échantillon de la totalité des tweets. Mais chaque retweet contient dans ses méta-données le nombre total de retweets jusqu'à la date du retweet considérée. Nous avons ainsi pu estimer la proportion de retweets capturés par notre dispositif en comparant le nombre total de retweets capturés au nombre maximum de retweets mentionné dans les méta-données.

Plus précisément, pour chaque jour t et pour chaque tweet original Tw émis ce jour là, si l'on note $n_t^c(Tw)$ le nombre de retweets de Tw captés et $n_t^m(Tw)$ le nombre maximum de retweets mentionnés dans les méta-données des retweets captés de Tw , nous estimons la proportion de retweets captés par :

$$\nu(t) = \frac{\sum_{\{Tw\}} n_t^c(Tw)}{\sum_{\{Tw\}} n_t^m(Tw)}$$

La figure 11 présente l'évolution de la proportion journalière de retweets capturée ainsi que le nombre total de tweets. Cette proportion est très variable en fonction du jour de collecte et dépend de l'ensemble du trafic Twitter d'une manière qui n'est pas spécifiée par l'API. En moyenne, sur l'échantillon testé, cette proportion est égale à 34%.

Twitter ne publiant pas le code source de ses APIs, nous ne savons pas comment sont produits les échantillons fournis par les APIs. En particulier, nous ne savons pas si le nombre de retweets mentionnés dans les méta-données d'un retweet entre dans les paramètres de sélection des tweets de l'échantillon. Néanmoins, nous pouvons raisonnablement penser que la proportion de tweets captés (mentions, citations, etc) par notre dispositif est de l'ordre de 34% du total des tweets correspondant à nos critères de capture.

Nos résultats montrent cependant que cette incertitude sur la précision de l'échantillon intrinsèque aux limites imposées par les API Twitter n'est pas un frein à l'analyse de la dynamique des communautés politiques.

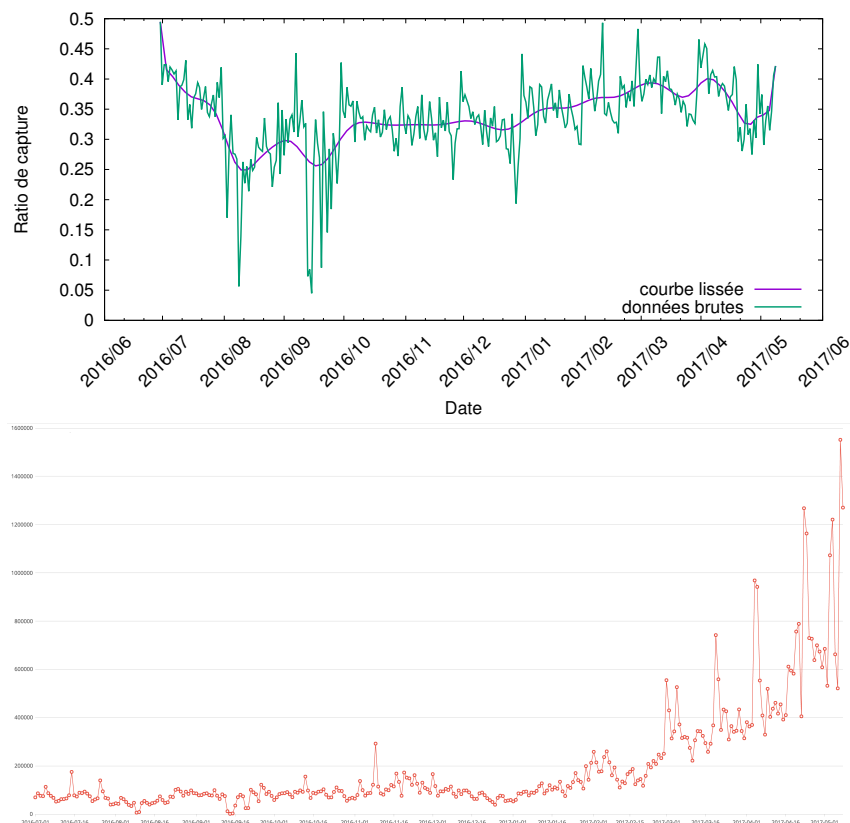


FIGURE 11 – *En haut.* Évolution du taux de capture des retweets au cours du temps. Certains points bas correspondant à des pannes serveurs. *En bas.* Nombre quotidien de tweets collectés.

B.3 Liste des principales personnalités politiques suivies

La campagne présidentielle française a été rythmée par deux primaires organisées par les deux partis ayant occupés le pouvoir ces dernières années, le *Parti Socialiste* (gauche) et *Les républicains* (droite). Les principaux candidats à ces deux primaires ainsi que l'ensemble des candidats à la présidentielle ont fait l'objet d'un suivi complémentaire au suivi via l'API "follow" à partir de l'API "track" *Politoscope*.

Nous avons ainsi suivi les personnalités suivantes :

Candidats à la présidentielle (par ordre de score)

1. Emmanuel Macron (EM)
2. Marine Le Pen (FN)
3. François Fillon (LR)
4. Jean-Luc Mélenchon (FI)
5. Benoît Hamon (PS)
6. Nicolas Dupont-Aignan (DF) (UPR)
7. Jean Lassalle
8. Philippe Poutou (NPA)
9. François Asselineau
10. Nathalie Arthaud (LO)

Principaux candidats aux primaires

- Yves Jadot (EEV - rallié à Hamon)
- Emmanuel Valls (candidat à la primaire de la gauche et premier ministre sortant)
- Arnaud Montebourg (candidat à la primaire de la gauche, ancien ministre)
- Vincent Peillon (candidat à la primaire de la gauche et ancien ministre)
- François Hollande (Président sortant)
- Alain Juppé (candidat à la primaire de la droite, favori des sondages jusqu'à la primaire de la droite)
- Nathalie Kosciusko-Morizet (candidate à la primaire de la droite)
- Nicolas Sarkozy (ex Président, candidat à la primaire de la droite)
- Jean-Frédéric Poisson (candidat à la primaire de la droite)

B.4 Architecture

Ce dispositif de collecte a été mis en place le 1er juin 2016 à partir de l'infrastructure Multivac du CNRS/ISC-PIF (<http://multivac.iscpif.fr> cf. figure 12) et a opéré en continu jusqu'au lendemain de l'élection présidentielle à l'exception de rares interruptions pour maintenance de nos serveurs (entre le 13-09-2016 et 15-09-2016 et entre le 21-09-2016 et le 22-09-2016). Entre le 1 juin 2016 et le 7 mai 2017 (deuxième tour de l'élection présidentielle), nous avons ainsi collecté 60 326 216 tweets émis par 2 477 910 comptes uniques.

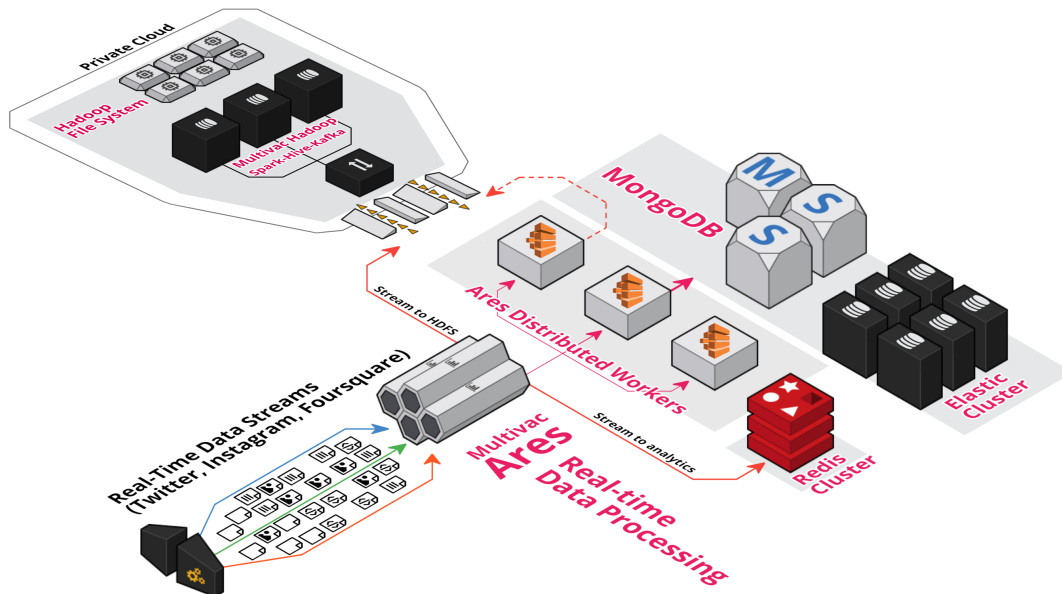


FIGURE 12 – Description synthétique de l’infrastructure Multivac utilisée pour mettre en place le dispositif d’analyse en temps réel des tweets. Cette infrastructure mobilise des technologies telles que Elasticsearch, Apache Hadoop, RabbitMQ, MongoDB ou Redis.

TWITTER DATA ANALYTICS ARCHITECTURE (see also figure 12)

Real-time data processing and micro services

- Elasticsearch Cluster (highly available and distributed)
 - Primary database with over 4 billion indexed documents
 - Full-text search
 - Indexing keywords
 - Data aggregation (date histogram, significant, terms, nested, etc.)
- MongoDB
 - Secondary database with over 8 billions of data
- RabbitMQ
 - Distributing Tweets to workers for parallel computations
- Node.js
 - Streaming Tweets by tracking and following Twitter accounts
 - Scalable workers : annotating and tagging Tweets in realtime
 - Extracting Tweets content
- Redis
 - Caching
 - In-memory analysis

Offline and batch processing : Hadoop cluster

- HDFS (Hadoop distributed file system)
- Apache Hadoop YARN : Managing cluster resources
- Apache kafka : Streaming data for batch processing
- Apache Spark
 - ETL
 - Pre-processing data
- Apache Hive : Run SQL queries in Hadoop

C Détection de bots

Il est fréquent d'utiliser des robots ou "bots" pour animer des comptes Twitter et démultiplier une présence en ligne. Cette utilisation, si elle était trop massive, pourrait aller à l'encontre de l'objet de notre étude qui est d'analyser le comportement des militants politiques. Pour s'assurer que la proportion des comptes animés par des bots restait à un niveau raisonnable dans notre base de données, nous avons utilisé l'API BotOrNot pour détecter les comptes qui pourraient ne pas être animés par de "vrais" utilisateurs.

En utilisant le même seuil de 0.5 préconisé par les concepteurs de l'API pour assigner la catégorie "bot" à un compte, nous obtenons qu'environ 5.3% des comptes ayant appartenu à une communauté sont des bots. Ces comptes ont produits moins de 5% des retweets enregistrés dans les communautés (voir tables 4 et 5).

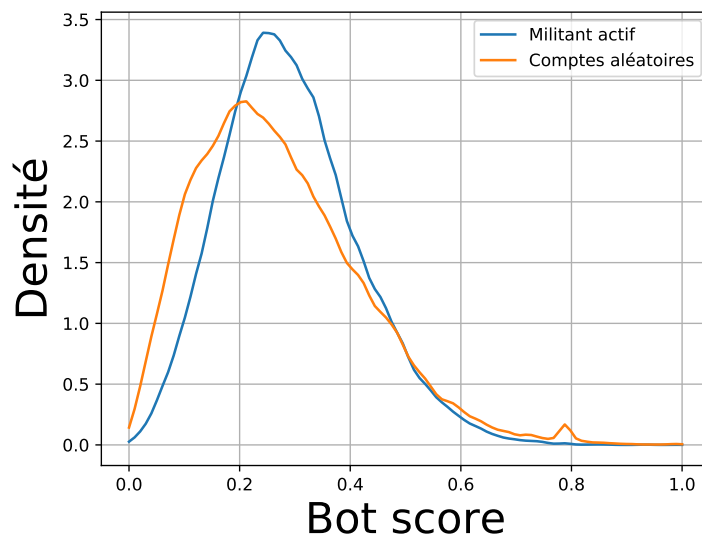


FIGURE 13 – Distribution du bot score des comptes ayant été au moins une fois dans une communauté (courbe bleue) et de comptes pris aléatoirement (courbe orange). Cette détection a été effectuée en juin 2016.

Nom	Nb moyen de comptes/période	Nb moyen de bots/période	% de bot
Jean-Luc Mélenchon	804	30	3.7 %
Emmanuel Macron	994	44	4.4 %
Nathalie arthaud	192	10	5.4 %
Marine Le Pen	1560	70	4.5 %
François Asselineau	213	14	6.5 %
Nicolas Dupont-Aignan	210	9	4.1 %
Manuel Valls	890	46	5.1 %
Yan Jadot	528	30	5.6 %
alainjuppe	432	23	5.4 %
Benoit Hamon	561	31	5.6 %
Jacque Cheminade	109	7	6.1 %
Nicolas Sarkozy	959	46	4.8 %
Philippe Poutou	478	20	4.2 %
François Fillon	1340	68	5.1 %
François Hollande	904	45	5.0 %
François Bayrou	932	40	4.3 %

TABLE 4 – Tableau présentant le nombre moyen de comptes par période et par communauté ainsi que le nombre moyen de bots détectés par période par communauté.

D Temps de vie des tweets et patterns d'activité de Twitter

L'activité politique sur Twitter connaît une certaine régularité qui est le couplage entre une activité endogène (jour/nuit, semaine/w.e.) et une activité exogène qui dépend de l'actualité politique. La

Type de tweet	Communautés politiques	Compte aléatoire
Original	7.2%	7.4%
Retweet sec	4.8%	3.7%
Réponse	0.9%	1.8%
Citation	3.4%	1.5%
Total	4.7%	3.9%
Référence	5.1%	6.8%

TABLE 5 – Proportion de contenus postés par les bots appartenant à des communautés politiques et par les bots détectés parmi les comptes choisis aléatoirement. Par exemple, les bots présents dans les communautés politiques ont émis 7.3% des tweets originaux émis par l'ensemble des comptes étant dans une communauté politique.

moyenne sur l'ensemble de notre corpus fait apparaître clairement une régularité hebdomadaire (cf. figure 14).

Concernant le temps caractéristique de retweet, 89.005% des retweets ont lieu en moyenne dans les 24h suivant la publication du tweet original et 98.977% ont lieu dans les 14 jours.

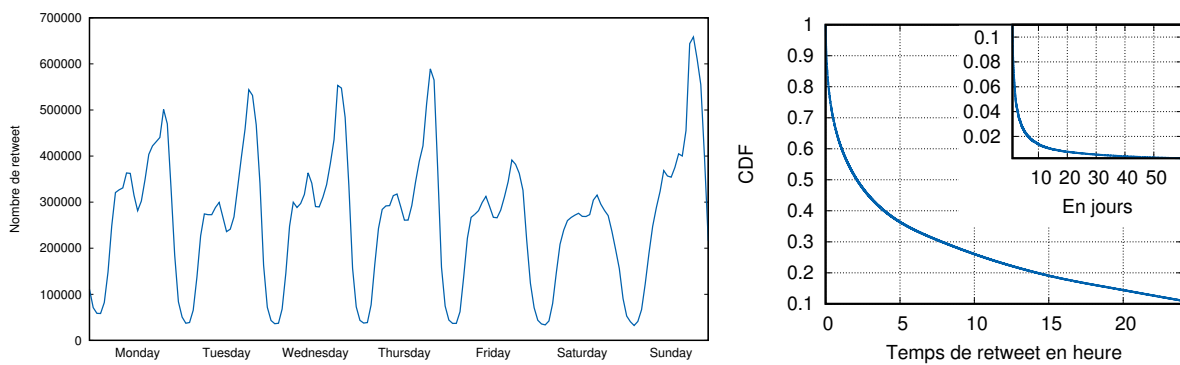


FIGURE 14 – *A gauche* : Nombre de retweets observés par jour et par heure entre le 1er juillet 2016 et 2 mai 2016; *à droite* : Distribution cumulative du temps pour retweets.

E Correlation entre structure de communautés et motifs de retweets

Nous avons analysé le détail des motifs de retweets afin de valider l'hypothèse H_0 , intuitive d'après notre méthode de reconstruction, selon laquelle la diffusion des retweets est fortement contrainte par la structure des communautés calculées. Nous vérifions pour cela que les motifs de retweets par les communautés sont très différents de ce qu'ils seraient s'ils avaient lieu de manière aléatoire au sein des communautés en analysant les motifs de diffusion pour des tweets ayant un nombre suffisant de retweets (≥ 50). Via notre dispositif, nous avons capté 4 124 888 tweet originaux qui ont totalisé 39 284 569 retweets. Parmi ces 4 millions de tweets, seulement 118 967 ont 50 retweets ou plus mais ils totalisent 21 288 537 des retweets soit 54% des retweets. En imposant que les retweets proviennent de comptes affiliés à une communauté, nous obtenons 82 112 tweets originaux totalisant 12 076 600 retweets soit 31% des retweets. Entre le 1er juin 2016 et le 7 mai 2017, nous avons calculé pour chaque jour j , les N^j communautés politiques fournies par notre méthodologie, notées $\mathcal{C}^j = \{C_i^j\}_{1 \leq i \leq N^j}$, et constitué un sous-corpus de tweets originaux \mathcal{D}_{50}^j composé des tweets ayant été retweetés au moins 50 fois par des personnes dont la communautés est connue dans les 14 jours suivants j , c'est à dire sur un jeu de données indépendant du jeu de données ayant servi à calculer les communautés. Sur cette période, nous avons calculé le nombre q_i^j de retweets de \mathcal{D}_{50}^j produits par l'ensemble des C_i^j . Nous obtenons ainsi la distribution $\mathcal{Q}^j = \{q_i^j\}_{1 \leq i \leq N^j}$ des retweets de \mathcal{D}_{50}^j par les communautés C_i^j . À partir de cela, nous calculons la distance χ^2 entre cette distribution et la distribution de retweets $\mathcal{P}^j(tw) = \{p(tw)_i^j\}_{1 \leq i \leq N^j}$ de chacun des tweets tw de \mathcal{D}_{50}^j :

$$\chi^2(P(tw), Q) = \sum_i \frac{(P(tw)_i - Q_i)^2}{Q_i}$$

D'après cette analyse sur l'ensemble des périodes, pour 92% des diffusions, H_0 est vraie avec une p -value inférieure à 10^{-16} . En considérant un seuil de p -value de 1%, 99.8% des tweets émis par les communautés ont une diffusion de retweet très différente d'une diffusion aléatoire dans les communautés.

F État des lieux du paysage politique avant le début de la campagne présidentielle 2017

Afin de proposer une visualisation de l'état de la twittosphère politique avant la campagne présidentielle, nous avons calculé le graphe de retweet sur la période août 2016 - décembre 2017, mis en évidence les 5-communautés (parties denses et colorées de la carte) et spatialisé avec le logiciel Gephi le graphe en résultant (cf. figure 15). Bien que sur cette période, les primaires de la droite aient pu commencer à modifier le paysage politique, cette carte que nous avons publiée sur le site du Politoscope et qui a été reprise comme image de la semaine dans le journal *Le Monde* (Supplément *Science et Médecine*) propose une représentation du paysage politique dont la pertinence a été confirmée par les alliances ultérieures qui se sont opérées en respectant les proximités des personnalités politiques sur la carte.

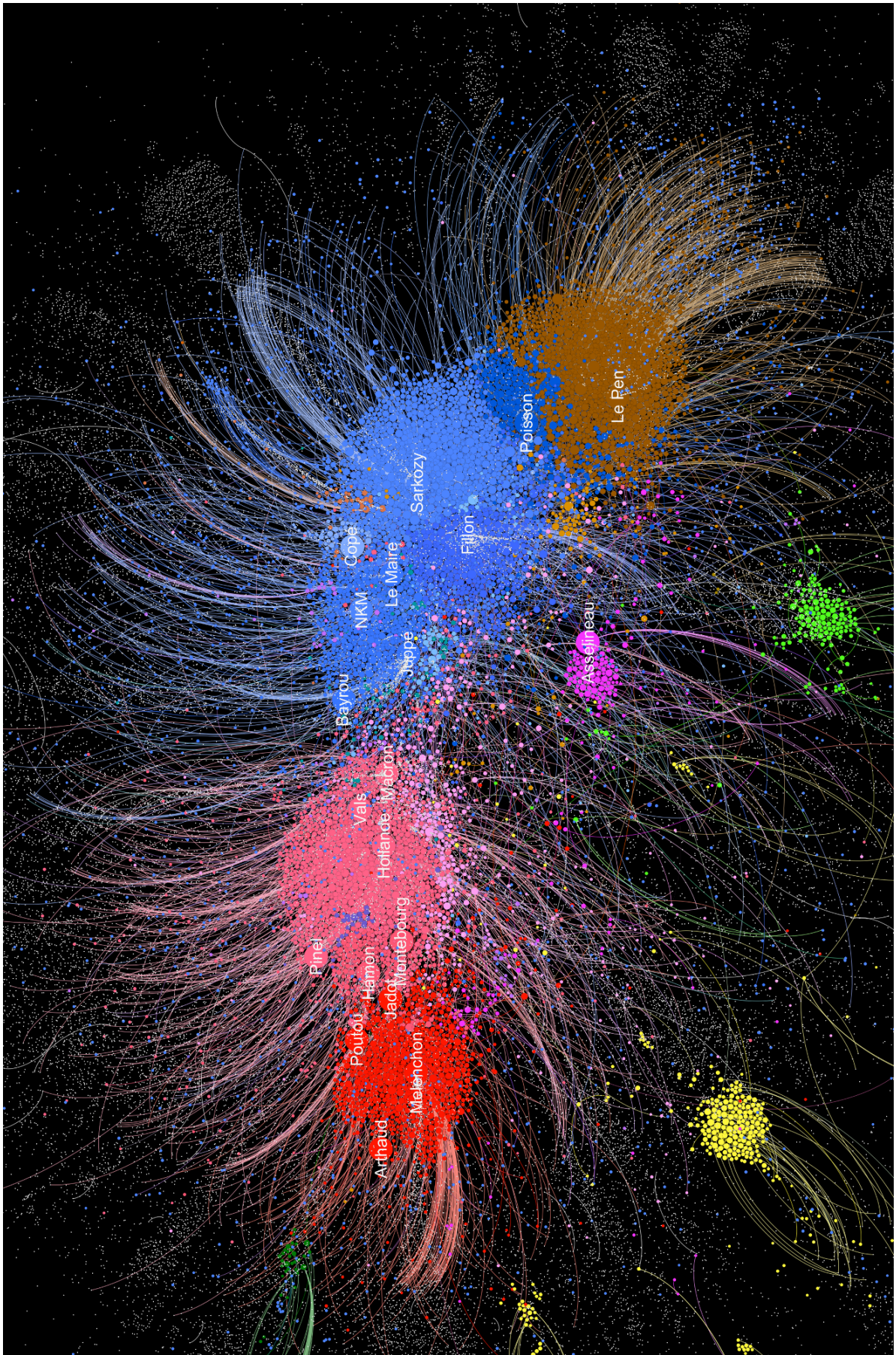


FIGURE 15 – Graphe des 5-communautés de la twittosphère politique française calculée sur la période 1 août 2016 - 31 décembre 2016. Le paysage politique multipolaire est bien visible avec un positionnement sur un axe gauche-droite pour la plupart des partis politiques. On notera que certaines proximités dans le graphe de personnalités politiques sont annonciatrices d’alliances futures : Hamon-Jadot puis Macron-Bayrou. Le parti de droite *Les Républicains* (en bleu) est morcelé en plusieurs courants, les communautés Juppé étant les plus éloignées du cœur formé par les communautés Fillon et Sarkozy. Ces deux dernières fusionneront suite aux primaires de la droite alors que la communauté Juppé migrera progressivement vers la communauté Macron. Le *Parti Socialiste* est comporte également une “aile gauche” (Hamon, Montebourg, Pinel) et une “aile droite” (Hollande, Vals). Source : Chavalarias D., Gaumont G. & Panahi M., La drôle de galaxie politique de Twitter. *Supplément Science et Médecine du Monde*, 18 avril 2017.

G Caractérisation des thèmes de la campagne présidentielle 2017

Afin de caractériser les différents débats et enjeux de cette campagne présidentielle, nous avons effectué une analyse du discours politique afin d'en extraire les principales expressions-clés. Nous avons utilisé pour cela deux types de sources : les mesures déclarées dans les programmes de campagne, rédigées dans un langage relativement élaboré, et les tweets des candidats, rédigés dans un langage relevant fréquemment du langage parlé.

Pour l'analyse des programmes des candidats, nous avons tout d'abord contacté le 1er mars 2017 toutes les équipes de campagne en position de rassembler 500 candidatures (le Conseil Constitutionnel n'a publié la liste officielle des candidats que le 20 mars 2017) en leur proposant un protocole pour nous fournir leurs éléments de programme et les expressions-clés qu'elles jugeraient caractéristiques de leur programme. Sans surprise, nous n'avons eu quasiment aucune réponse, seule l'équipe d'un "petit" candidat nous a indiqué qu'étant donné la taille de l'équipe de campagne, elle n'était pas certaine de pouvoir répondre.

Nous avons alors utilisé les programmes mis en ligne par les onze candidats (téléchargés entre le 14 février et le 24 mars 2017) dont nous avons manuellement extrait 980 mesures politiques.

Pour compléter ce corpus, nous avons extrait à partir de la base de donnée Politoscope 25 883 tweets publiés par les onze candidats et quelques autres personnalités politiques importantes (cf. B) entre le août 2016 et janvier 2017. Ce second corpus a l'avantage de faire ressortir des thèmes qui ont émergés lors des débats politiques, indépendamment des orientations programmatiques des candidats.

Nous avons analysé ces deux corpus à l'aide du logiciel de text-mining Gargantext (<http://gargantext.org>) et identifié un vocabulaire de près de 1500 groupes d'expressions qualifiant les thèmes de la campagne présidentielle (cf. section H.1 pour la liste des labels). Gargantext est un logiciel libre d'analyse de corpus développé par le CNRS à l'Institut des Systèmes Complexes de Paris Île-de-France. Il permet à l'utilisateur d'interagir avec des algorithmes avancés de text-mining et des visualisations synthétisant le contenu du corpus afin d'y apporter sa propre expertise et fournir une représentation multi-échelle de son contenu (expressions importantes, documents importants, thèmes, articulation des thèmes et leur évolution).

En utilisant la carte thématique produite par Gargantext (figure 16), nous avons pu identifier les principaux sujets discutés par les candidats et leur audience lors de cette campagne.

Nous nous sommes appuyés sur cette carte pour sélectionner onze sujets qui nous ont semblé particulièrement importants et représentatifs des débats (cf. description en G.2). Notre objectif n'était pas d'être exhaustifs mais de mettre en évidence les hétérogénéités thématiques des communautés politiques et leurs spécialisations respectives sur certains grands thèmes.

G.1 Évolution du vocabulaire des communautés

La plateforme compagnon de cet article, le Politoscope (<https://presidentielle2017.politoscope.org/dashboard>) propose de suivre l'évolution des usages des 15 mots-clés les plus spécifiques à chacune des principales communautés politiques lorsqu'elles traitent chacun des onze thèmes sélectionnés. Ces mots clés ont été identifiés en utilisant l'algorithme d'agrégation basé sur le test de χ^2 implémenté dans Elasticsearch⁸.

On peut y observer qu'en plus d'une spécificité sémantique de traitement des thèmes, l'usage du vocabulaire évolue au cours du temps pour les différents candidats, certains thèmes apparaissant, se renforçant ou disparaissant en cours de route.

Nous en montrons ici quelques exemples, le lecteur pourra se référer à la plateforme pour une exploration exhaustive.

G.2 Définition des thèmes par requêtes

Nous nous sommes appuyés sur la carte présentée dans la figure 16 pour sélectionner onze sujets qui nous ont semblé particulièrement importants et représentatifs des débats. Notre objectif n'était pas d'être exhaustifs mais de mettre en évidence l'hétérogénéité thématique des communautés politiques et leurs spécialisations respectives sur certains grands thèmes. Les groupes de termes identifiés par Gargantext sur les mesures politiques ont été retravaillé afin de retirer les termes trop polysémiques et d'ajouter des expressions provenant de l'analyse des tweets. L'objectif était d'obtenir une requête permettant d'extraire de la base de données du Politoscope un ensemble de tweets cohérents relevant de ce thème.

Les requêtes suivantes ont été retenues :

- **Emploi** : travailleur* OR emploi* OR "apprentissage" OR "chômage" OR chômeur* OR "rsa" OR "smic" OR "flexisécurité" OR "formation professionnelle" OR "insertion professionnelle" OR "lycées professionnels" OR "formation initiale" OR "heures supplémentaires" OR "temps partiel" OR "contrat de professionnalisation" OR "conditions de travail" OR "emplois précaires" OR "enseignement professionnel et technologique" OR "compensation salariale" OR "39 heures" OR "charges sociales salariales" OR "employeurs" OR "filières professionnelles" OR "35 heures"

8. Voir <https://www.elastic.co/guide/en/elasticsearch/reference/current/search-aggregations-bucket-significantterms-aggregation.html> et Manning et al., Chapter 13.5.2 - <https://nlp.stanford.edu/IR-book/pdf/irbookonlinereading.pdf>

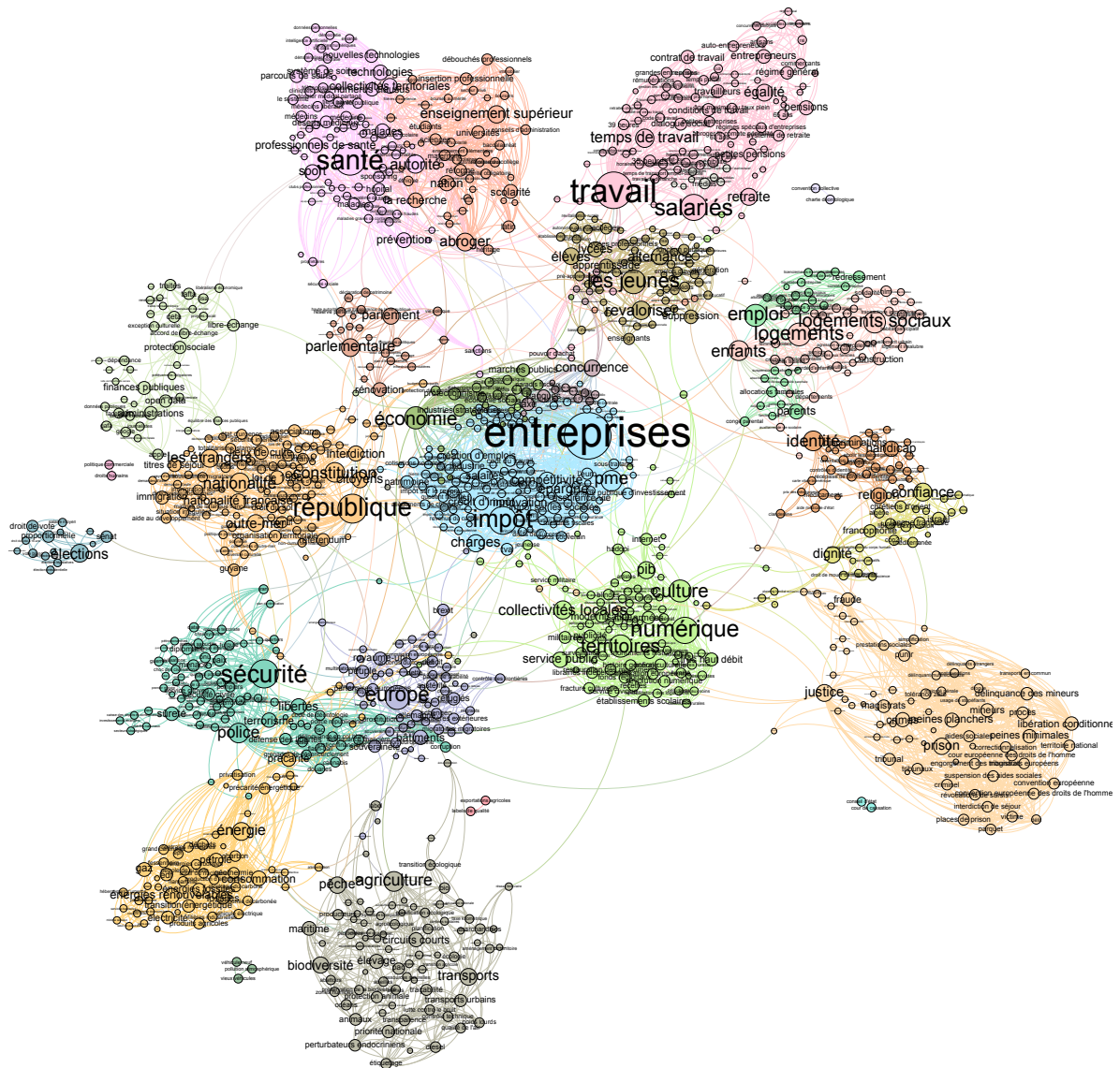


FIGURE 16 – Carte des thèmes abordés dans les programmes de la présidentielle 2017 réalisée à l’aide de Gargantext à partir de l’ensemble des mesures politiques extraites des programmes des candidats. Les nœuds de la carte sont des labels pour des groupes de termes jugés équivalents dans le discours politique. Les liens entre un label A et un label B indiquent que la probabilité que A et B soient mobilisés conjointement dans une même mesure politique est forte. Gargantext applique l’algorithme de Louvain pour identifier des clusters de termes ayant une forte interaction entre eux et les affiche dans la même couleur. Pour améliorer la visibilité, la carte a été éditée dans le logiciel Gephi (<http://gephi.org>) afin de fixer la taille des nœuds et des labels suivant une fonction monotone de leur PageRank (Page, L., Brin, S., Motwani, R., & Winograd, T. (1999). The PageRank citation ranking : Bringing order to the web. Stanford InfoLab..

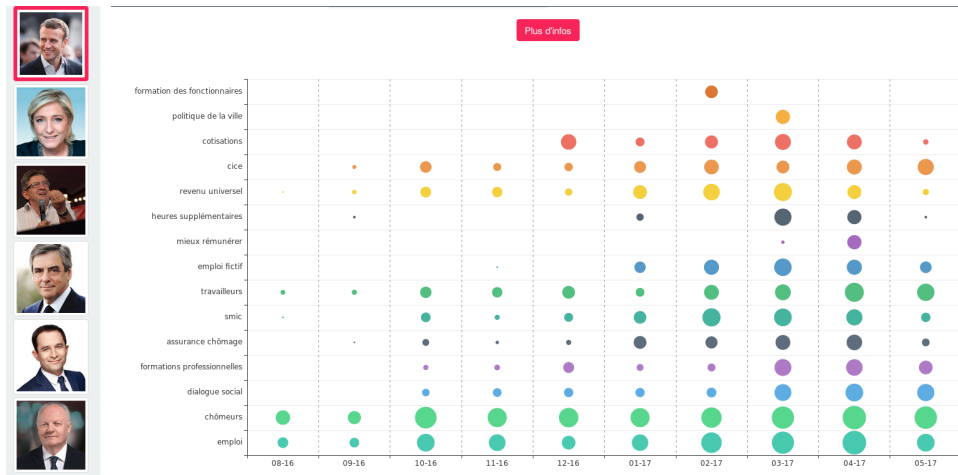


FIGURE 17 – Évolution mensuelle de l'usage des 15 termes les plus spécifiques à la communauté Macron sur le thème de l'emploi.

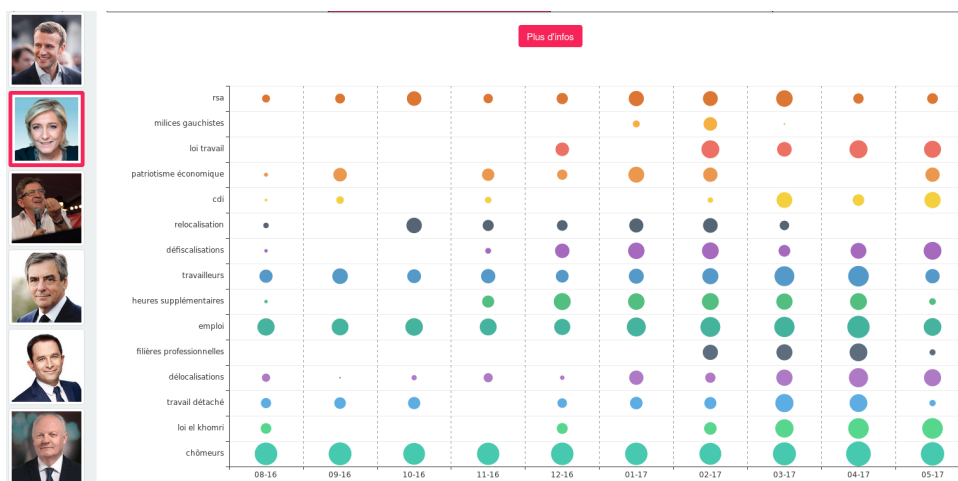


FIGURE 18 – Évolution mensuelle de l'usage des 15 termes les plus spécifiques à la communauté Le Pen sur le thème de l'emploi.

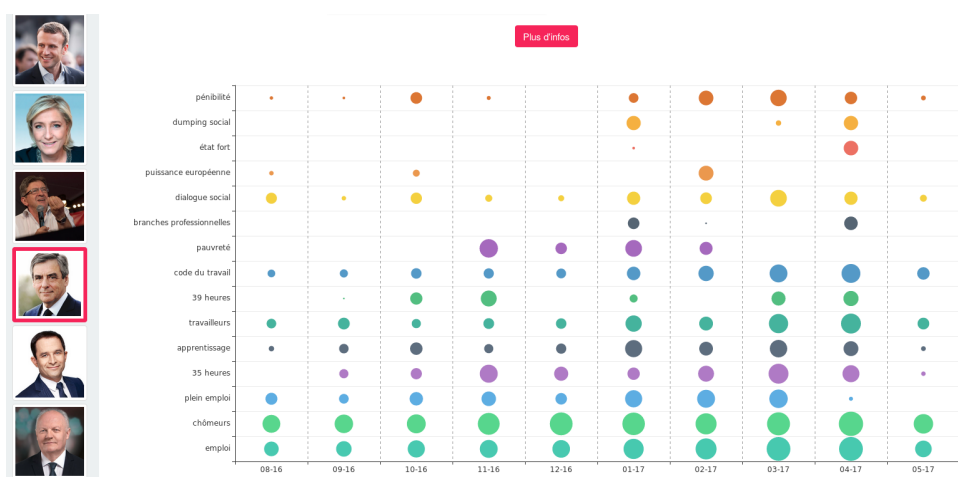


FIGURE 19 – Évolution mensuelle de l'usage des 15 termes les plus spécifiques à la communauté Fillon sur le thème de l'emploi.

OR "tutorat" OR "pôle emploi" OR "création d'entreprise" OR "salaire horaire" OR "négociation salariale" OR "exonérations de cotisation sociale" OR "insertion des jeunes" OR "parcours professionnels" OR "licenciement*" OR "travail obligatoires" OR "jeune en alternance" OR "37 heures" OR "contrat de génération" OR "loi travail" OR "augmenter le salaire" OR "prime pour l'emploi" OR "code du travail" OR "formation des fonctionnaires" OR "dialogue social" OR "salaire des fonctionnaires" OR "loi el khomri" OR "sous-traitants" OR "devenir entrepreneurs" OR "souffrance au travail" OR "gel des traitements" OR "obligation pour eux d'accepter un emploi" OR "maîtrise de la masse salariale" OR "non remplacement de un fonctionnaire sur deux" OR "suppressions de postes" OR "artisanat" OR "réduction du temps de travail" OR "RTT" OR "cice" OR "mobilité professionnelle" OR "âge

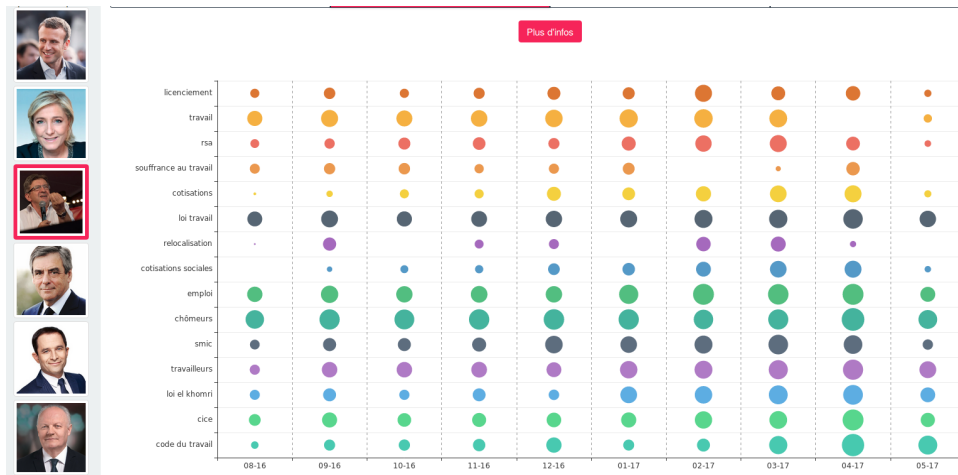


FIGURE 20 – Évolution mensuelle de l’usage des 15 termes les plus spécifiques à la communauté Mélenchon sur le thème de l’emploi.



FIGURE 21 – Évolution mensuelle de l’usage des 15 termes les plus spécifiques à la communauté Hamon sur le thème de l’emploi.

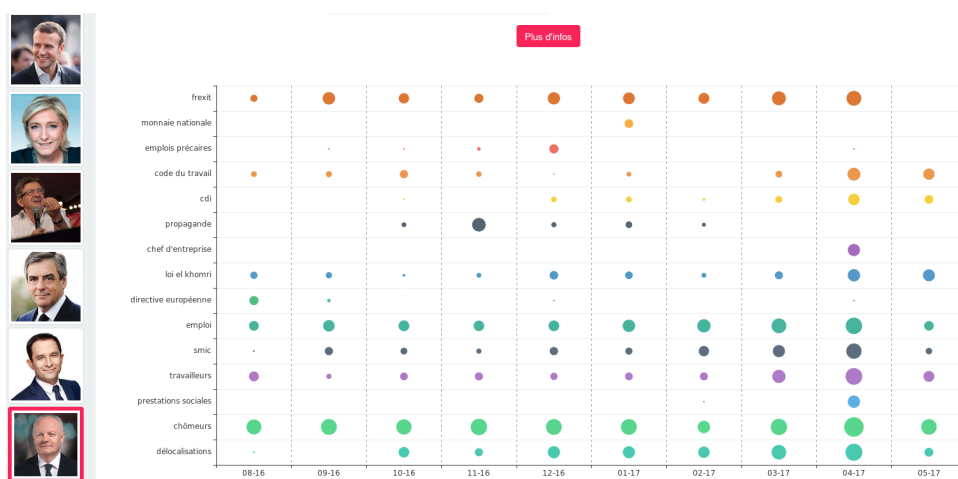


FIGURE 22 – Évolution mensuelle de l’usage des 15 termes les plus spécifiques à la communauté Asselineau sur le thème de l’emploi.

légal de départ à la retraite” OR ”retraite solidaire” OR ”pénibilité ” OR ”salaire moyen” OR ”cdi” OR ”cdd” OR ”délocalisations” OR ”relocalisation” OR ”revenu universel” OR ”revenuuniversel” OR ”contrat de travail” OR ”syndicalisation” OR ”cotisations sociales” OR ”détachement des travailleurs” OR ”travailleurs détachés” OR ”délocalisation” OR ”relocalisation” OR ”plans sociaux”

- **Économie & fiscalité :** révolution fiscale” OR ”coût de la vie” OR ”pouvoir d’achat” OR ”libre échange” OR ”libéralisme économique” OR ”économie sociale et solidaire” OR entreprise* OR ”pme” OR ”pmi” OR actionnaire* OR relocalisation* OR ”crédit impôt” OR nationalisation* OR commerçant* OR ”plus-values” OR délocalisation*

- OR "scop" OR "niches fiscales" OR "gafa" OR "cice" OR "taux d'imposition" OR "retenue à la source" OR "crédit impôt-recherche" OR "impôt sur les sociétés" OR "services innovants" OR "marchés publics" OR "livret d'épargne industrie" OR "exportation" OR "impôt sur la fortune" OR "conseils d'administration" OR "small business act" OR "surplus" OR "capital-risque" OR "comités de rémunération" OR "représentants des salariés" OR "harmonisation fiscale" OR "complexité administrative" OR "reprise d'entreprises" OR "stock options" OR "start-ups" OR "pme européennes" OR "commandes publiques" OR "fiscalité européenne" OR "impôt minimal" OR "chiffre d'affaires mondial" OR "prélèvement à la source" OR "échanges commerciaux" OR "travailleurs handicapés" OR "OR" OR "commerce indépendant" OR "politique commerciale" OR "développement économique" OR "politique commerciale" OR "charges patronales" OR "tva anti-délocalisations" OR "allègements de charges" OR "cice" OR "alléger les charges patronales" OR "allègement des charges patronales" OR "centrales d'achat" OR "concurrence internationale" OR "exploitations familiales" OR "financiarisation" OR "travailleurs indépendants" OR "liberté syndicale" OR "austérité" OR "fisc*" OR "impôt OR "imposition" OR "cotisations sociales" OR "retenue à la source" OR "tranche supplémentaire" OR "isf" OR "prélèvement à la source" OR "cotisations chômage" OR "exonérations de cotisation sociale" OR "tva" OR "charges patronales" OR "tva" OR "csg" OR "taxe" OR "allègements de charges" OR "cice" OR "trésor public" OR "droits de mutation" OR "dette des états" OR "dette de l'états" OR "dette publique" OR banque* OR financi* OR "subprimes" OR bancaire* OR "épargne" OR spécul* OR "livret A" OR "agios" OR "taux d'usure" OR emprunt* OR "agence de notation" OR "crédit à la consommation" OR "surendettement" OR "assurance-vie" OR "loi sapin" OR "banque centrale européenne" OR "euro-obligations" OR "pacte de responsabilité" OR "taxe mondiale sur les transactions financières" OR "taxe Tobin" OR "dépôts des particuliers" OR "crise bancaire" OR "crise financière" OR trader* OR "ras-le-bol fiscal"
- **Démocratie** : front républicain" OR "ni-ni" OR "non-cumul des mandats" OR "conseil constitutionnel" OR "ve république" OR "6e république" OR "6eme république" OR "5e république" OR "vie politique" OR "données publiques" OR "%penelope%" OR "%pénélope%" OR "%pénélope%" OR "démocratie sociale" OR "cour des comptes" OR "proportionnelle" OR "parité" OR "assemblée nationale" OR "élections législatives" OR "fonction publique territoriale" OR "formation des fonctionnaires" OR "durée du quinquennat" OR "préfet" OR "établissements publics" OR "cour européenne des droits de l'Homme" OR "préférence nationale" OR "nationalistes" OR "patriotes" OR "progressistes" OR "%penelope%" OR "parlement*" OR élu* OR "corruption" OR "ena" OR "révisions constitutionnelles" OR "réserve parlementaire" OR "statut pénal du chef de l'état" OR scandale OR (marine AND 300000) OR (marine AND "300 000") OR ("Le Pen" AND 300000) OR ("le pen" AND "300 000") OR "inéligibilité" OR "fonction présidentielle" OR "e-démocratie" OR "49.3" OR "déclaration de patrimoine" OR "égalité républicaine"
- **Recherche & Éducation** : lycée* OR "pisa" OR "éducation" OR apprentissage* OR "maternelle" OR "décrochage scolaire" OR "scolarité" OR "scolaire" OR "service civique" OR "savoirs fondamentaux" OR "méthodes pédagogiques" OR "enseignement professionnel et technologique" OR "établissements scolaires" OR "tutorat" OR "échec scolaire" OR "écoles d'ingénieurs" OR "financer leurs études" OR "banque de la jeunesse" OR "écoles de commerce" OR "e-learning" OR "éducation artistique" OR "enseignement primaire" OR universit* OR "la recherche" OR chercheur* OR "R&D" science* OR scientifique* OR enseignant* OR professeur* OR "loi lru" OR "enseignants-chercheurs" OR "enseignement supérieur" OR "grandes écoles" OR "investissements d'avenir" OR "loi firaso" OR "crédit impôt-recherche" OR "conseil national des universités" OR "CNU" OR "autonomie des universités" OR "universités" OR "ZEP"
- **Immigration et multiculturalisme** : migr* OR immigr* OR "nationalité" OR clandest* OR "droit de séjour" OR "schengen" OR "pays d'origine" OR "contrôle des frontières" OR "réadmission" OR "régularisations" OR "regroupement familial" OR naturalisation* OR "reconduites à la frontière" OR "titres de séjour" OR "procédures d'expulsion" OR "séjour des étrangers" OR "libre circulation des personnes" OR "connaissance préalable de la langue française" OR "intégration des étrangers" OR "citoyenneté française" OR "droit du sol" OR "droit du sang" OR islam* OR musul* OR muzz* OR salaf* OR "burkini" OR "barbarie" OR "laïcité" OR juif* OR juive* OR chrétien* OR "communautarisme" OR "multiculturalisme" OR "modèle républicain" OR "identité nationale" OR "civilisation française" OR "port du voile" OR "préférence raciale" OR race OR "demande d'asile" OR "submersion migratoire"
- **Santé** : santé" OR médecin* OR médic* OR hôpit* OR "sécurité sociale" OR "autisme" OR handicap* OR "mdph" OR "désamiantage" OR "cancer" OR "assurance-maladie" OR "hébergement en établissement" OR "maintien à domicile" OR "soins d'urgence" OR "urgences sanitaires" OR "maladies" OR "perturbateurs endocriniens" OR "accès aux soins" OR "hospitalisation" OR "soins dispensés" OR "parcours de soins" OR "centre de soin" OR "malentendants" OR "malvoyants" OR "dépassements d'honoraires" OR "transports sanitaires" OR "arrêt de travail" OR "numerus clausus" OR "carte vitale" OR "perte d'autonomie" OR "ehpad" OR "agences de sécurité sanitaire et alimentaire" OR "tiers payant" OR "recours aux génériques" OR "lunettes"
- **Écologie et environnement** : écolo* OR "biodiversité" OR "protection environnementale" OR "protection de l'environnement" OR pollution* OR "diesel" OR "ogm" OR "changement climatique" OR "fessenheim" OR "flamanville" OR "gaz de schiste" OR "pesticides" OR "gaz à effet de serre" OR "ressources naturelles" OR "quotas de pêche" OR "agroécologique" OR "énergies marines renouvelables" OR "pêche illégale" OR "performance énergétique" OR "rénovation thermique" OR "isolation de l'habitat" OR "précarité énergétique" OR "isolation de l'habitat" OR "forêts" OR "sylvicultures" OR "changements environnementaux" OR "services écosystémiques" OR "déforestation" OR "aires protégées" OR "émissions de carbone" OR "émission de CO2" OR "empreinte carbone" OR "gaz à effet de serre" OR "politique climatique" OR "durabilité" OR "obsolescence programmée" OR "niveau des mers" OR "niveau de la mer" OR "vagues de chaleur"
- **Politique énergétique** : politique énergétique" OR "électricité" OR "nucléaire" OR "énergies renouvelables" OR "éolien" OR "énergies marines renouvelables" OR "énergie renouvelable" OR "énergie solaire" OR "ep*" OR "flamanville" OR "charbon" OR "centrales thermiques" OR "diesel" OR "fessenheim" OR "flamanville" OR éol* OR "performance énergétique" OR "fukushima" OR "rénovation thermique" OR "isolation de l'habitat" OR "précarité énergétique" OR "gaz de schiste" OR "énergies marines renouvelables" OR "isolation de l'habitat" OR "sources d'énergie" OR "transition énergétique" OR "conversion écologique" OR "énergies alternatives" OR "dépendance énergétique" OR "combustibles fossiles" OR "énergies fossiles" OR "efficacité énergétique"
- **Politique étrangère** : europ* OR souverain* OR "CETA" OR "TAFTA" OR "bruxelles" OR "russie" OR "Poutine" OR "les russes" OR "pacte de stabilité" OR "directive européenne" OR "traité franco-allemand" OR "trump" OR "washington" OR "maison blanche" OR "moscou" OR "crise syrienne" OR "libye" OR "iran" OR "irak" OR "syrie" OR "OTAN" OR "traité transatlantique" OR "traités internationaux"
- **Sécurité intérieure** : terror* OR attentat" OR "cyber-sécurité" OR "guerre civile" OR "surveillance" OR délinqu* OR espion* OR loi taubira" OR "vidéo-surveillance" OR "vidéo-protection" OR "sécurité des français" OR "laxisme"

judiciaire " OR "prescription" OR "peines planchers" OR "remises de peine" OR "Daech" OR "cyber-attaques"
OR "cyber-attaque" OR "#cyberthreats" OR "#cyberthreat" OR "cyber-guerre" OR "cyber guerre" OR "cyber-
menace" OR "#cybersecurity" OR "cyber-sécurité" OR "cyber sécurité" OR "cybersécurité" OR "cyberdéfense"
OR "cyberarmée" OR "daech" OR "daesh"

- **Agriculture** : agricult* OR agro* OR "pac" OR "alimentation" OR "engrais" OR "circuits courts" OR "transition agricole" OR "surplus" OR intrant* OR "bio" OR "pesticides"

G.3 Distribution des thèmes par communautés politiques

Nous avons étudié la diversité des communautés discutant de chaque thème. Nous avons calculé l'entropie des communautés prenant la parole sur un thèmes (cf. tableau 6). Bien que la plupart des thèmes aient une entropie assez similaire, il ressort que les thèmes "*immigration et multiculturalisme*" et "*sécurité intérieure*" sont discutés par un nombre plus faible de communautés car leur entropie est basse. À l'inverse, le thème "*emploi*" est celui qui est discuté le plus uniformément par l'ensemble des communautés. En plus de l'entropie, nous montrons également pour chaque thème les deux communautés prenant particulièrement la parole sur ce thème. Comme les communautés ont des tailles très diverses, nous montrons les communautés ayant le plus haut score de *tf-idf* sur un thème au lieu des communautés les plus présentes. Ainsi, si une petite communauté ne parle que d'un thème, elle pourra tout de même faire partie des communautés principales sur ce thème même si en valeur absolue elle participe moins que d'autres. Le score *tf-idf* d'une communauté c_j pour un thème t_i est calculé de la manière suivante :

$$tf-idf(c_j, t_i) = \frac{tweets(c_j, t_i)}{\sum_k tweets(c_k, t_i)} \log \left(\frac{\sum_{k,l} tweets(c_k, t_l)}{\sum_l tweets(c_j, t_l)} \right),$$

où $tweets(c_j, t_i)$ est le nombre de tweets de la communauté c_j sur le thèmes t_i .

Thème	Entropie	Nb communautés	Communautés principales
agriculture	2.63	6.2	François Fillon, Marine Le Pen
économie	2.63	6.2	François Fillon, Jean-Luc Mélenchon
emploi	2.74	6.7	Jean-Luc Mélenchon, Benoît Hamon
démocratie	2.63	6.2	Emmanuel Macron, Marine Le Pen
écologie	2.52	5.7	Benoît Hamon, Jean-Luc Mélenchon
immigration et multiculturalisme	1.77	3.4	Marine Le Pen, François Fillon
recherche et éducation	2.70	6.5	François Fillon, Emmanuel Macron
santé	2.68	6.4	François Fillon, Benoît Hamon
sécurité intérieure	2.30	4.9	Marine Le Pen, François Fillon
politique énergétique	2.56	5.9	Jean-Luc Mélenchon, Benoît Hamon
politique étrangère	2.69	6.4	Marine Le Pen, Emmanuel Macron
Tous thèmes	2.61	6.1	Marine Le Pen, François Fillon

TABLE 6 – Tableau de l'entropie des 3-communautés discutant chaque thème. Les 2 communautés principales selon le $TF-IDF$ sont également présentées pour chaque thème. L'entropie maximale est de 3 lorsque les 8 candidats ont le même volume de parole sur ce thème.

Nous avons également étudié la diversité des thèmes discutés dans chaque communauté en calculant l'entropie des thèmes dans chaque communauté (cf. tableau 1). Sur ce point, il y a des disparités entre les candidats. Cependant, les thèmes que nous avons définis ne sont pas exhaustif et ne permettent pas de conclure sur la diversité des thèmes abordés par les candidats.

Afin de mieux comprendre la représentativité de ces thèmes et la présence de chaque candidat, la proportion de tweets appartenant à un thème et les proportions de chaque thèmes sont présentées dans les figures 23 à 30.

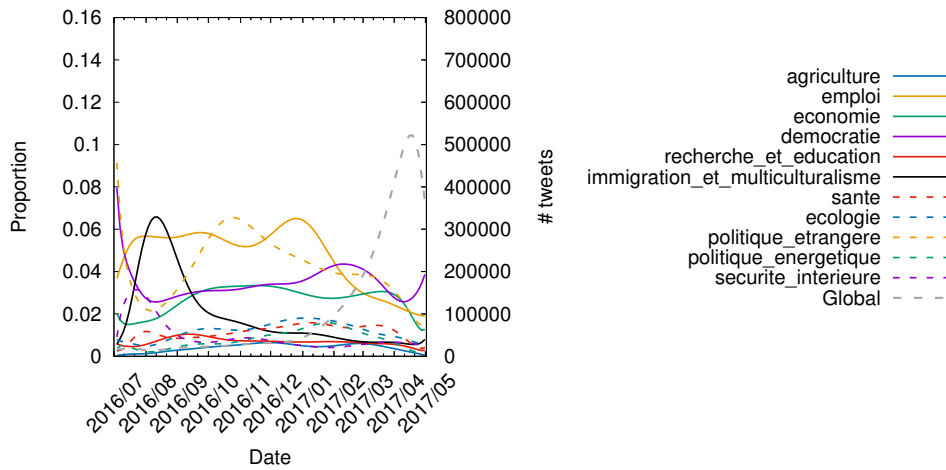


FIGURE 23 – Évolution des thèmes discutés par la communauté Jean-Luc Mélenchon entre le 01 Juin 2016 et le 08 Mai 2017.

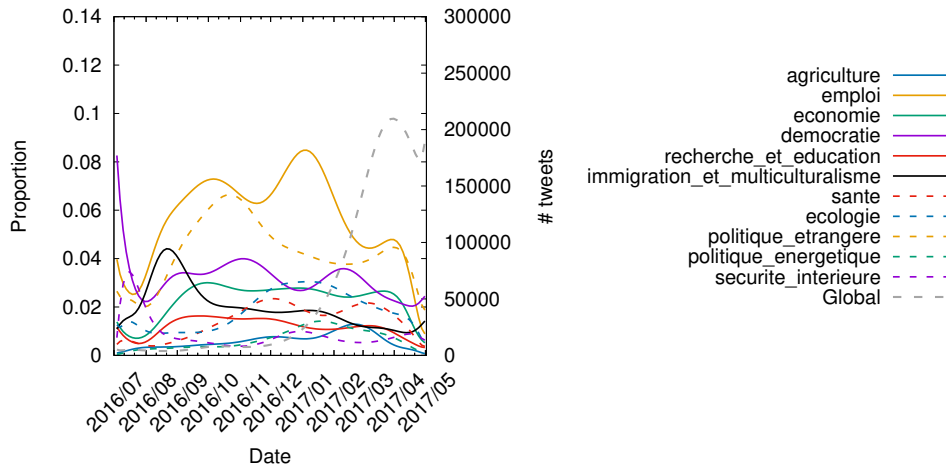


FIGURE 24 – Évolution des thèmes discutés par la communauté Benoit Hamon entre le 01 Juin 2016 et le 08 Mai 2017.

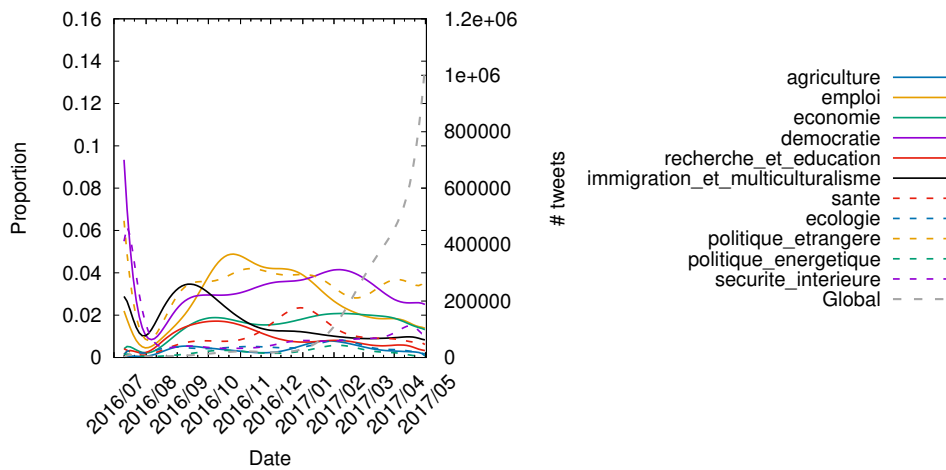


FIGURE 25 – Évolution des thèmes discutés par la communauté Emmanuel Macron entre le 01 Juin 2016 et le 08 Mai 2017.

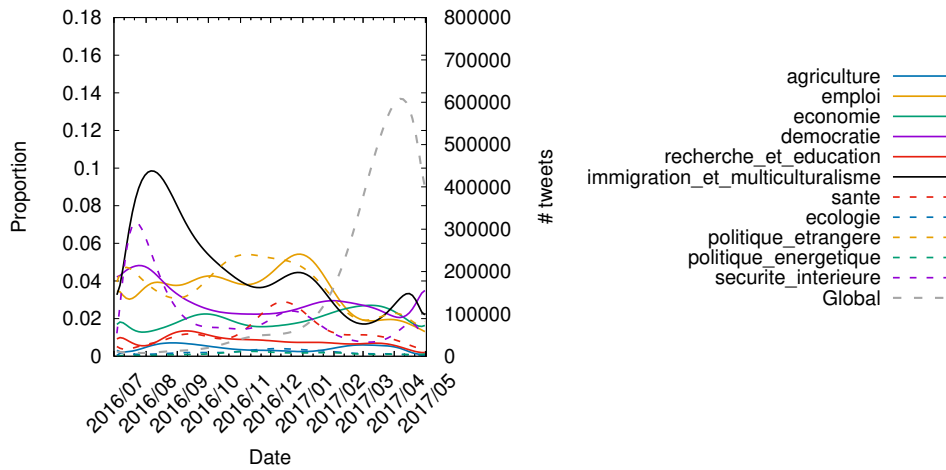


FIGURE 26 – Évolution des thèmes discutés par la communauté François Fillon entre le 01 Juin 2016 et le 08 Mai 2017.

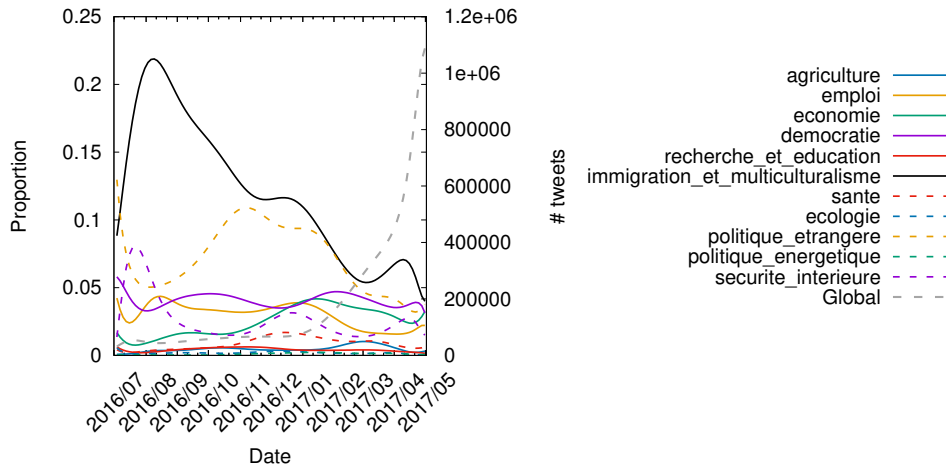


FIGURE 27 – Évolution des thèmes discutés par la communauté Marine Le Pen entre le 01 Juin 2016 et le 08 Mai 2017.

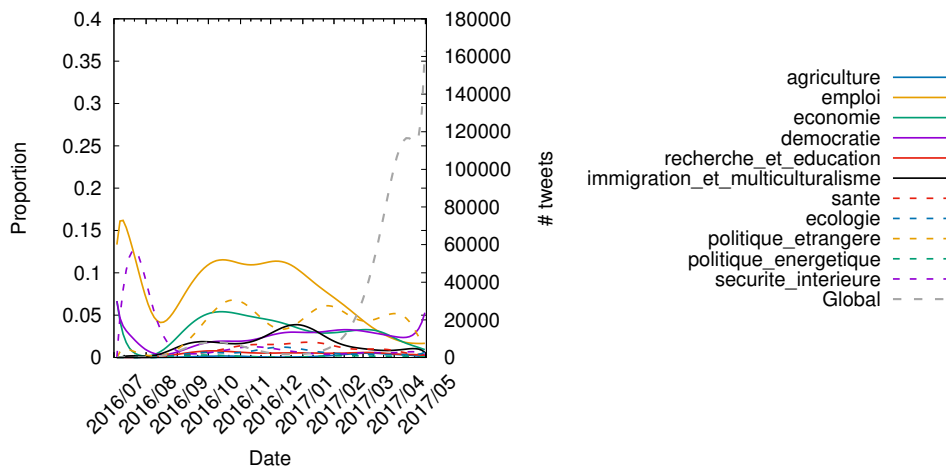


FIGURE 28 – Évolution des thèmes discutés par la communauté Nathalie Arthaud entre le 01 Juin 2016 et le 08 Mai 2017.

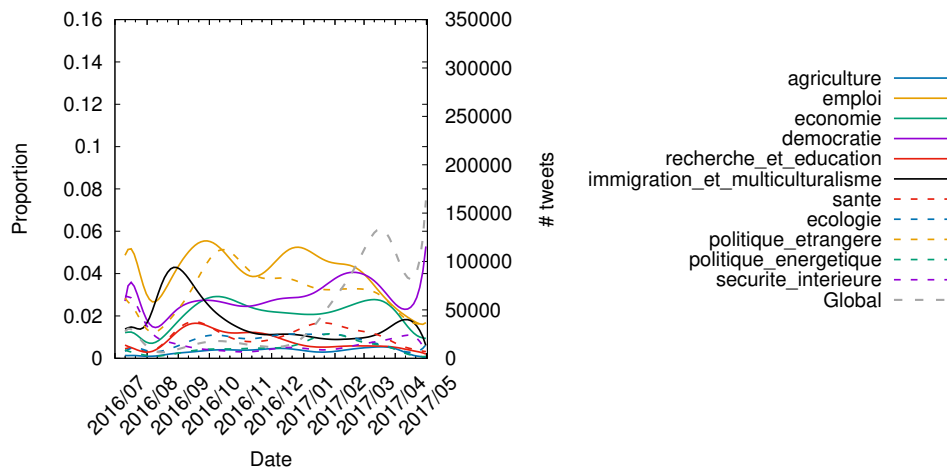


FIGURE 29 – Évolution des thèmes discutés par la communauté Philippe Poutou entre le 01 Juin 2016 et le 08 Mai 2017.

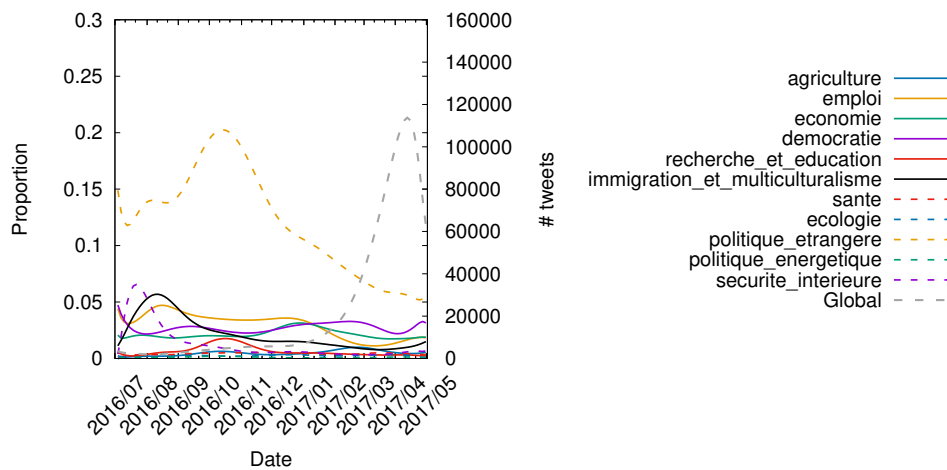


FIGURE 30 – Évolution des thèmes discutés par la communauté François Asselineau entre le 01 Juin 2016 et le 08 Mai 2017.

H Stabilité des communautés

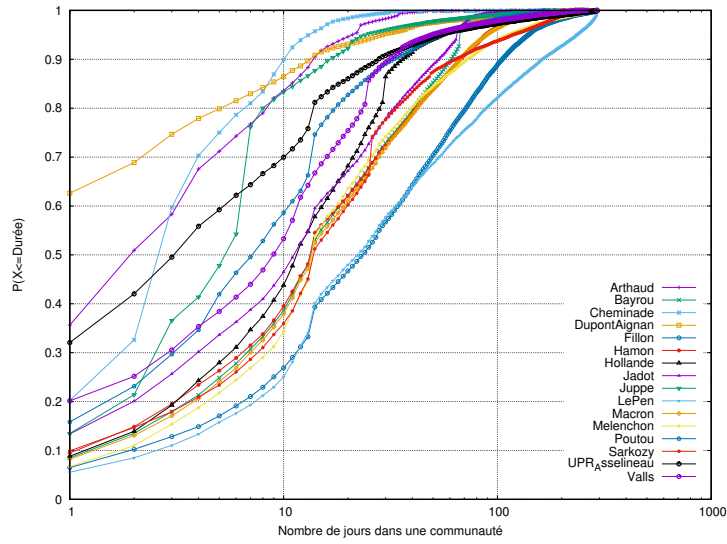


FIGURE 31 – Distribution cumulative du temps passé dans une communauté entre juillet et le 1er tour.

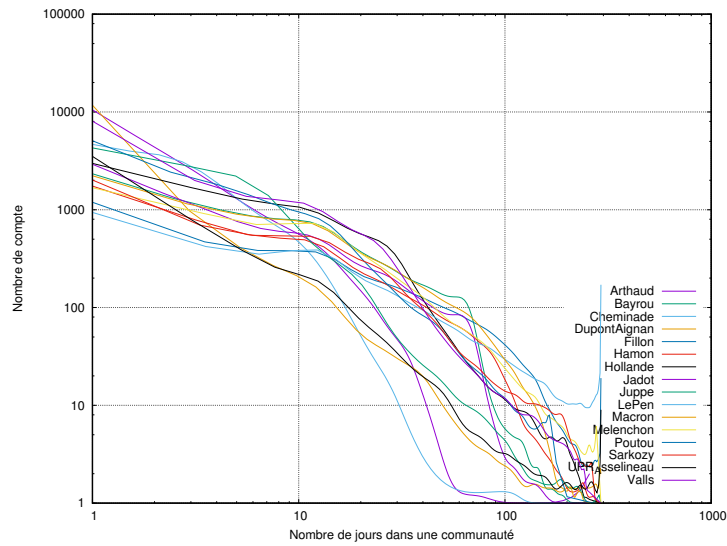


FIGURE 32 – Distribution (lissée) du temps passé dans une communauté entre juillet et le 1er tour.

H.1 Vocabulaire du discours politique de l'élection présidentielle 2017

Les termes jugés équivalents ont été regroupés et apparaissent dans la suite des analyses en un label qui est le premier terme du groupe.

Liste des groupes de termes extraits via le logiciel Gargantext (<http://gargantext.org>)

- #frexit OR frexit
- 18-25 ans
- 300 000 euros OR 300 000 € OR 300 000€
- 35 heures OR 35h
- 37 heures
- 39 heures
- 6ème république OR 6emerepublique OR 6erepublique OR 6erépublique
- abattements fiscaux
- abattoirs
- abroger OR abrogation
- abus de position dominante
- accès à la propriété
- accès aux soins
- accessibilité des personnes handicapées
- accord de libre-échange
- accords bilatéraux
- accords de schengen
- accords du touquet
- accueil des réfugiés
- accusé
- actifs
- action économique
- action publique
- actionnaires
- actionnariat salarié
- actions d'entreprises
- activités de conseil
- administration centrale
- administration fiscale
- administration pénitentiaire
- administratives OR administration OR administrations
- aéroports
- affaires criminelles
- afghanistan
- afp
- afrique
- âge légal de départ à la retraite
- agence publique européenne de notation
- agences de sécurité sanitaire et alimentaire
- agents publics
- agglomération
- agios
- agriculture OR agriculteurs OR agricole OR agricoles OR agriculteur
- agriculture biologique OR agriculture bio
- agroécologique
- aide à la construction
- aide médicale d'état OR aide médicale d'état
- aide sociale à l'enfance
- aides publiques
- aides sociales OR aide sociale
- airfrance
- algérie
- alimentation
- allègement
- allègements de charges
- allègements fiscaux
- alléger les charges patronales
- allemagne
- allocation adulte handicapé OR aah
- allocation aux adultes handicapés
- allocations OR allocs
- allocations chômage
- allocations familiales OR allocation familiale
- allocations sociales
- alpes-maritimes
- alstom
- alternance
- alternance forte
- aluminium
- alzheimer
- aménagement du territoire
- amendes
- anciens combattants
- animaux
- ankara
- apa
- apl
- apologie du terrorisme
- apologie du terrorisme
- apple
- apprentis
- apprentissage OR apprentissages
- l'arabie saoudite OR l'arabie saoudit
- argent public
- armée d'afrique
- armées OR armée
- arrangements fiscaux
- arrêt de travail
- artisanat local
- artisans OR artisanat
- artistes OR artiste
- aspa
- assange
- assassinat politique
- assassiner OR assassins
- assemblée constituante
- assemblée nationale
- assimilation
- assistance médicalisée
- assistant parlementaire OR assistante parlementaire
- associations OR association
- assurance-chômage
- assurance-maladie OR assurance maladie
- assurance-vie
- atteindre l'équilibre budgétaire
- attentatnice
- attentats OR attentat
- augmenter
- augmenter le salaire
- austérité
- autisme
- autonomie aux communes
- autonomie fiscale
- autorité
- autorité de l'état
- autorité hiérarchique
- autorité indépendante
- autorités françaises
- autoroutes
- auxiliaire de vie scolaire
- avions ravitailleurs
- avocats
- avortement
- avs
- ayant-droits
- baccalauréat
- bachar al-assad
- baisse d'impôts OR baisses d'impôts OR baisse d'impôt
- baisse des dépenses interventions de l'état
- baisser
- baisser les prix
- balkany
- banlieues OR banlieue
- banque centrale
- banque centrale européenne OR bce
- banque de détail
- banque de données
- banque de la jeunesse
- banque postale
- banque publique d'investissement
- banques OR banque
- banques d'affaires
- banques de dépôt OR banques de dépôts
- barbarie OR barbare OR barbares
- barbarie terroriste
- bas revenus
- bâtiments
- bâtiments publics OR bâtiments nationaux
- bayrou
- bénéficiaires du rsa
- bénévoles
- benoît hamon
- bi-langues
- bibliothèques
- bien-être animal
- bien-être des animaux
- biens communs

- bio
- biodiversité
- biogaz
- blessés
- blindés
- blocages
- bois
- bonus
- bonus-malus
- bonus/malus
- branches professionnelles
- brexit OR #brexit
- bricolage
- bruxelles
- budget européen
- budgets participatifs
- bureaucratie
- burkini OR #burkini
- burnout
- burqa
- buy european act
- bygmalion
- cadeaux fiscaux
- caf
- caisse des dépôts et consignations
- caissière
- calais
- cancers OR cancer
- cannabis
- cantine
- cantines bio
- cantines scolaires
- capital-risque
- caricature OR caricatures
- carnées
- carrières judiciaires
- carte de séjour
- carte vitale biométrique
- casier judiciaire
- caution solidaire
- caution universelle
- cavous
- cdd
- cdi OR contrat à durée indéterminée
- censure
- centrales OR central OR centrale
- centrales d'achat
- centrales nucléaires
- centres éducatifs fermés
- ceseda
- ceta
- chaînes publiques de télévision
- chambres régionales des comptes
- champ d'honneur
- changement climatique
- charges OR charge
- charges patronales
- charges salariales
- charges sociales salariales
- chef d'entreprise OR chefs d'entreprise OR chefs d'entreprises
- chef de l'état
- chefs d'état OR chef d'état OR chef d'état OR chefs d'état
- chiffre d'affaires mondial
- chine
- chirurgie ambulatoire
- chômeur de longue durée
- chômeurs OR chômeur OR chômage
- chrétiens d'orient OR chretiens-dorient
- cice
- circuits courts
- citoyenneté française
- citoyens
- civilisation européenne
- civilisation française
- civilisations OR civilisation
- clandestine
- clandestins OR clandestin
- classe politique
- classes moyennes
- clause générale de compétence
- clients des banques
- climat
- clubs professionnels
- cnspf
- cocaïne
- code civil
- code de l'entrée et du séjour des étrangers
- code de la nationalité française
- code du travail
- cohésion nationale
- cohésion sociale
- collectivités OR collectivité
- collectivités d'outre-mer
- collectivités locales
- collectivités territoriales OR collectivité territoriale
- collègues OR collègue
- combattants
- comités de rémunération
- commandes publiques OR commande publique
- commerçants
- commerçants indépendants
- commerce
- commerce indépendant
- commission européenne
- communautarisme OR communautarismes
- communautés urbaines
- communiste OR communistes
- compassion
- compensation salariale
- compétitivité
- complexité administrative
- complot OR complotiste
- compte bancaire
- compte individuel de formation
- comptes publics
- concertation
- concurrence OR concurrents
- concurrence déloyale
- concurrence internationale
- concurrence loyale
- concurrence sociale déloyale
- condamnations OR condamnation OR condamnés
- conditions de travail
- confession juive OR juif OR juifs
- confiance confiance
- conflits d'intérêt
- congé parental
- conjoints survivants
- connaissance préalable de la langue française
- conseil constitutionnel
- conseil de sécurité
- conseil national des universités
- conseil supérieur de l'audiovisuel
- conseil supérieur de la magistrature OR csm
- conseiller territorial
- conseils d'administration
- conseils généraux
- conseils régionaux
- consommations OR consommer OR consommation OR consommateurs
- consommation responsable
- constitution
- construction
- construction européenne
- contrat de génération
- contrat de professionnalisation
- contrats d'assurance
- contribution fiscale
- contrôle de légalité a priori
- contrôle des frontières OR contrôle de ses frontières
- contrôles d'identité
- convention de dublin
- convention européenne des droits de l'homme
- conventions citoyennes
- conversion écologique
- conversion écologique et énergétique
- convocation
- convocation judiciaire OR convocation des juges
- cop21
- corporatisme
- correctionnalisation
- corruption
- cotisation vieillesse
- cotisations OR cotisation OR cotiser
- cotisations chômage
- cotisations sociales
- coup d'État institutionnel
- coupables
- couples homosexuels
- cour des comptes
- cours d'assises
- coût de la vie
- coût du travail
- couverture mobile
- création artistique
- création d'entreprise
- crèches

- crédit à la consommation
- crédit impôt OR crédit d'impôt
- crédit impôt réparation
- crédit impôt-recherche OR crédit d'impôt recherche OR crédit impôt recherche
- crédits OR crédit
- crimes OR crime
- criminel OR criminels
- crise
- crise bancaire
- crise bancaire économique
- croissance
- csg
- culture OR cultures
- culture française
- cumul
- cyber-attaques OR cyber-attaque OR #cyberthreats OR #cyberthreat OR cyber-guerre OR cyber-guerre OR cyber-guerre OR cyber-menace
- #cybersecurity OR cyber-sécurité OR cyber sécurité OR cybersécurité
- cyberdéfense OR cyberarmée
- daech OR daesh
- décentralisation
- déchets
- décisions de remise en liberté
- déclassement
- découverts bancaires
- décrochage scolaire OR déscolarisé
- décrocheurs
- déductibilité fiscale des dons
- défense de la laïcité
- défense européenne commune
- défense nationale
- déficit
- défiscalisations OR défiscaliser OR défiscalisation OR défiscalisées
- dégradation du mobilier urbain
- délais de jugement
- délinquants étrangers
- délit OR délits OR délinquance OR délinquante OR délinquants
- délocalisations OR délocalisent OR delocalisees OR délocalisé
- demandes d'asile OR demandeurs d'asile
- dématérialisation OR dématérialisée
- démocratie OR démocraties OR democracy OR democracies
- démocratie sociale
- dentaires
- départements OR département
- dépassements d'honoraires
- dépendance
- dépendance au pétrole
- dépenses administratives
- dépenses culturelles
- dépenses d'investissement
- dépenses de l'état OR dépenses publiques OR dépenses de fonctionnement et d'intervention de l'état OR dépense publique
- dépôts des particuliers
- dérives jurisprudentielles
- désamiantage
- désendettement
- désengorger les urgences
- déserts médicaux
- déserts universitaires
- détachement des travailleurs
- détourné d'argent public OR détournement d'argent public
- dette OR dettes
- dette des états OR dette des états OR dette publique
- développement durable
- développement économique
- devenir entrepreneurs
- dialogue social
- dictatures OR dictature
- diesel
- dieu
- diffusion de l'art
- dignité
- diplomatie
- directive européenne
- directives de libéralisation
- discrimination
- discrimination positive
- discriminations
- dislocation
- dispositif de renseignement
- dissuasion nucléaire
- distribution d'électricité
- diversité culturelle
- dividendes
- djihadistes OR djihad
- dmto
- doctorants
- documents officiels
- donald trump OR d.trump OR donaldtrump OR d..trump OR d trump
- donations
- données publiques
- dons
- dotations de l'état OR dotations de l'état
- double nationalité
- drapeau français
- drogue OR drogues
- droit à l'oubli
- droit d'auteur OR droits des créateurs
- droit d'être syndiqué
- droit de faire appel des décisions
- droit de mourir dans la dignité
- droit de séjour
- droit de veto
- droit de vote
- droit de vote des étrangers
- droit du sol
- droit européen
- droit français
- droits à construire
- droits civiques
- droits culturels
- droits de douane
- droits de mutation OR droit de mutation
- droits de succession
- droits humains
- drones
- dumping social
- durée du quinquennat
- e-learning
- eau OR eaux
- eaux
- échanges commerciaux
- échec scolaire
- éco-conception
- école OR écoles
- école nationale de la magistrature OR ecole nationale de la magistrature OR ENM
- écoles d'ingénieurs
- écoles de commerce
- écoles primaires OR école primaire
- écologie OR écologique OR écologiques OR ecologie
- écologistes
- économie OR économies OR économie
- économie maritime OR economiedelamer
- économie numérique
- économie réelle
- économie sociale et solidaire
- économies sur les dépenses
- économiste OR économistes
- edf
- éducatives OR éducation
- éducation artistique
- éducation nationale
- eelv
- effectifs de l'état OR effectifs de l'état
- effort de défense
- effort national
- égalité OR egalite
- égalité républicaine
- égalité salariale
- ehpad
- el khomri
- électeurs
- élection présidentielle OR élections présidentielles
- élections OR élection
- élections législatives
- élections locales
- électricité
- élevage
- élevage intensif
- élèves
- éleveur OR éleveurs
- élu
- élucidation des affaires
- élus condamnés
- élus locaux
- embaucher
- émissions de gaz à effet de serre

- emotion
- empêcher les abus
- emploi OR emplois OR embauches OR embauche
- emploi fictif
- emploi présumé fictif
- emploi stable
- emplois précaires
- employés
- employeurs
- emprunts
- ena OR école nationale d'administration
- endettement de l'état
- énergie OR énergies
- énergies marines renouvelables
- énergies renouvelables OR énergie renouvelable OR éolien OR énergie solaire
- éolien
- enfants OR enfant
- enfants récidivistes
- engrais
- enmarche
- enquêtes OR enquête
- enr
- enseignants OR enseignant
- enseignants-chercheurs
- enseignement élémentaire
- enseignement primaire
- enseignement professionnel et technologique
- enseignement supérieur
- enseignements OR enseignement OR enseigner
- ententes
- entrepreneuriat culturel
- entrepreneurs
- entreprises OR entreprise
- entreprises innovantes
- entreprises publiques
- épargne
- épargne bancaire
- épargne populaire
- epr
- équilibre budgétaire
- erasmus
- espaces maritimes
- espaces naturels protégés
- établissements hospitaliers
- établissements privés
- établissements publics
- établissements scolaires
- établissements universitaires
- état de droit
- état fort
- état libre
- état palestinien
- état-major
- éthique OR éthiques
- étudiants
- étudiants étrangers
- euro-atlantistes
- euro-obligations
- europe OR union européenne OR ue
- exception culturelle
- exode des rapatriés
- exonérations OR exonération
- exonérations de cotisation sociale
- exonérer de charges patronales
- expérimentations animales
- expertise citoyenne
- exploitations familiales
- exportation
- exportations agricoles
- expulsions
- faceauxchrétiens
- facebook
- faciès
- falsification
- familles
- familles françaises
- fan-zones OR fan zone
- favoriser
- femme-homme OR femmes-hommes
- féodalités locales
- fermes-usines
- fessenheim
- fichés s
- fierté d'être français
- filières du travail clandestin
- filières industrielles
- filières professionnalisantes
- filières professionnelles
- fin de vie
- finance spéculative
- financement de la recherche
- financement étranger
- financement public
- financer leurs études
- finances publiques
- finances publiques locales
- financiarisation
- financiers
- fiscalité OR fiscal OR fiscales OR fiscale
- fiscalité européenne
- fiscalité locale
- fixation des tarifs
- flamanville
- florange
- flotte
- flux migratoires
- fnsea
- fonction présidentielle
- fonction publique OR fonctions publiques
- fonction publique hospitalière
- fonction publique territoriale
- fonctionnaires OR personnel de l'état
- fonds européen de défense
- fonds régionaux
- forces armées
- forêts
- formation OR formations
- formation de nos jeunes
- formation des fonctionnaires
- formation initiale
- formations politiques
- formations professionnelles OR formation professionnelle
- fossiles
- frais bancaires
- françafrique
- france du désastre
- france métropolitaine
- franchises OR franchise
- François Asselineau
- François Hollande
- francophonie
- frapor
- fraternité
- fraude OR frauduleux OR fraudes
- fraude aux prestations sociales
- fraude fiscale
- fraude sociale
- fraudes commerciales
- frenchtech
- frères musulmans
- frontière OR frontières
- frontières extérieures
- frustration
- fuite des cerveaux
- g20
- gafa
- gains de productivité
- garde à vue
- garde d'enfants
- garde nationale
- garde-frontières
- gauche irréconciliable OR gauches irréconciliables
- gaulois OR gaulois
- gaz
- gaz à effet de serre
- gaz de schiste
- gel des traitements
- gel des transferts
- gendarmes OR gendarmerie
- général de Gaulle
- génération
- gestion planifiée
- google
- gouvernement algérien
- gpa
- grand emprunt
- grand patronat
- grand remplacement
- grande Bretagne OR grande-Bretagne
- grande distribution OR grandes surfaces OR grands distributeurs
- grande guerre
- grande réforme fiscale
- grandes écoles
- grandes entreprises OR grands groupes OR grosses entreprises
- grèce
- gros patrimoines
- gros travaux

- guadeloupe
- guerre civile
- guerre mondiale
- guerres
- guyane
- gwada
- hacker
- hadopi
- harcèlement
- harki OR harkis
- harmonisation fiscale
- hébergement d'urgence
- hébergement en établissement
- héritage
- heures supplémentaires
- histoire nationale
- hlm
- hommage aux victimes
- hôpital OR hôpitaux hospital OR hopitaux
- hôpital public
- hôpitaux de proximité
- hospitalisation
- identité
- identité de genre
- identité nationale
- imams
- immigration de masse
- immigration illégale OR migrants clandestins OR immigration clandestine OR immigrés clandestins
- immigration légale
- immobilier
- importation OR importations
- impôt OR impôts OR imposition
- impôt minimal
- impôt sur la fortune
- impôt sur le revenu
- impôt sur les sociétés
- impôts locaux
- incivilités
- incubateurs
- indemnisation
- indemnités des élus locaux
- indemnités journalières
- indépendance de la justice
- indépendance stratégique
- industrie OR industriels
- industries stratégiques
- inéligibilité
- infirmières
- inflation
- ingénieurs OR ingénieur
- ingérence
- injustices OR injustice
- innovation OR innovations
- insertion des jeunes
- insertion professionnelle
- inspection générale de l'état
- inspections générales interministérielles
- insultes
- intégration des étrangers
- intercommunalités OR intercommunalité
- interdiction OR interdire
- intérêt général
- intérêts économique OR intérêts économiques
- intérêts nationaux OR intérêt national
- intermittent du spectacle
- internet
- intrants
- investissement privé
- investissement public
- investissements d'avenir
- investissements étrangers
- iran
- irlande
- isf
- islam OR l'islam
- islamistes OR islamiste OR islamique OR islamiques
- isolation de l'habitat
- israël
- istanbul
- ivg
- jaures OR jaurès
- jdd
- jean-luc mélenchon
- jeune en alternance
- jeunesse OR jeunesses
- journalistes
- juges OR juge
- jungle OR jungles
- juppé
- jury de cour d'assise
- jurys citoyens
- justice
- justice des mineurs
- justice fiscale
- justice sociale
- kinésithérapeutes
- égalité femmes-hommes OR égalité femmes/hommes
- état islamique OR état islamique OR ei
- les étrangers
- excellence environnementale
- extrême-droite OR extrême droite
- identité heureuse
- immigration
- islam politique
- facho OR fachos OR fachisme OR fachistes OR fachiste
- l'union européenne
- la poste
- la recherche OR chercheurs
- la république
- la russie
- labels de qualité
- lagarde
- laïcité
- langue française
- langues régionales
- latin
- laxisme
- laxisme judiciaire
- le pen
- légitime défense
- les collectivités
- les étrangers
- les jeunes OR jeunes OR la jeunesse
- libéralisation des services
- libéralisme économique
- libération conditionnelle
- liberonsnosconvictions
- liberté d'expression
- liberté religieuse
- liberté syndicale
- librairies indépendantes
- libre circulation
- libre circulation des personnes
- libre-échange
- libye
- licence globale
- licences professionnelles
- licenciement OR licenciements
- licenciements collectifs OR licenciement collectif
- lieux de culte
- litiges
- livret a
- livret d'épargne industrie
- locataires
- logements OR logement
- logements sociaux OR logement social
- loi d'orientation
- loi de 1905
- loi de moralisation de la vie publique
- loi el khomri
- loi fioraso
- loi hadopi
- loi lru
- loi religieuse
- loi sapin
- loi sru
- loi sur la protection des sources
- loi sur le renforcement de la démocratie et des libertés locales
- loिताubira OR loi taubira
- loitavail OR loi travail
- loyers
- lunettes
- lutte antiterroriste
- lutte contre l'évasion fiscale
- lutte contre l'immigration clandestine
- lutte contre l'optimisation fiscale
- lutte contre la contrefaçon
- lutte contre la discrimination
- lutte contre la fraude sociale
- lutte contre la précarité OR lutterai contre la précarité
- lutte contre les discriminations
- lutte contre les fraudes
- lycéens
- lycées OR lycée
- lycées professionnels
- machines OR machine

- madeinfrance
- magistrats OR magistrat OR magistrature
- maintien à domicile
- maisons de santé
- maisons de soins
- maîtrise de la langue française
- maîtrise de la masse salariale
- majorité silencieuse
- malades OR malade
- maladies OR maladie
- maladies graves ou contagieuses
- malentendants
- malvoyants
- mandat
- manifpolice
- marché commun
- marché unique européen
- marchés publics
- marges des producteurs
- mariage
- marine nationale
- maritime
- marseillaise
- massacre
- maternelle
- mdph
- mécénat populaire
- médecine
- médecins
- médecins généralistes
- medef
- médias
- médicaments
- méditerranée
- membres du parquet
- menace OR menaces
- menace terroriste OR menace terroristes
- ménages
- mensonges OR mensonge
- méritocratie
- mesures emblématiques
- mesures fiscales
- méthodes pédagogiques
- métiers de l'entrepreneuriat
- métiers de la finance
- métiers du droit
- microentreprise
- mieux rémunérer
- migrants
- migratoires
- militaires
- millefeuille administratif
- mineurs
- minimum vieillesse
- ministère de l'intérieur, de l'immigration et de la laïcité
- ministère des souverainetés
- ministres du culte
- mis en examen
- mixité sociale
- mlp-officiel
- mobilisation OR mobilise
- mobilité professionnelle
- modèle républicain
- modem
- modernisation
- moins de 18 ans
- monde carcéral
- monde judiciaire
- monde rural
- mondialisation
- monnaie
- monnaie nationale
- montauban
- montée de la délinquance
- moraliser
- moratoire
- mosquées OR mosquée
- moyen-orient OR moyen orient
- moyenne européenne
- moyens de nos armées
- multiculturalisme
- multilatéralisme
- multinationales
- musique OR musiques
- musulman OR musulmans OR musulmanes
- mutation digitale
- mutuelles
- nation OR nations
- nationalisation OR nationaliser
- nationalité
- nationalité française
- naturalisations OR naturalisation
- négociation salariale
- neutralité
- niceattentat OR attentat de nice
- niches fiscales OR niche fiscale
- nicolassarkozy
- non remplacement de un fonctionnaire sur deux OR non-remplacement d'un fonctionnaire sur deux
- non-cumul des mandats
- norme nationale
- normes OR norme
- normes environnementales
- normes européennes
- nos soldats
- notation sociale
- notre sécurité
- notredémocratie
- nouvelle calédonie OR nouvelle-calédonie
- nouvelles technologies
- nucléaire OR industrie nucléaire OR filière nucléaire OR part du nucléaire
- nuitdebout
- nullité de procédure
- numérique OR numériques OR numérisation
- numerus clausus
- obligation pour eux d'accepter un emploi
- obligations légales
- officines
- offre de logements
- offres en ligne
- ogm
- omc
- onema
- ong
- onu OR nations unies
- opa
- open data OR open-data
- opérateurs téléphoniques
- opérations extérieures OR opex
- opérations spéculatives
- opinion publique
- optimisation fiscale
- ordonnances
- organisations syndicales
- orthophonistes
- otan OR l'otan
- outil de défense
- outre-mer OR outre-mers
- ouvriers
- pac OR politique agricole commune
- pacs
- pacte de confiance
- pacte de responsabilité
- pacte de stabilité
- paix
- palestine
- paradis fiscaux
- parcours de soins
- parcours européens
- parcours professionnels
- parcours scolaire
- parents
- parisiens OR parisien
- parité
- parlement
- parlement européen
- parlementaire
- partenaires européens
- partenaires sociaux
- partenariats public-privé
- patrimoine
- patriotisme OR patriotes
- patriotisme économique
- patrons OR patron
- pays d'afrique
- pays d'origine
- pays de l'union OR état membre
- paysans OR paysanne OR paysannes
- pêche
- pêche illégale
- peines OR peine
- bi-nationalité OR binationalité OR binationaux
- déchéance de nationalité
- peines de prison OR peines d'emprisonnement
- peines de réclusion criminelle
- peines planchers
- penelope fillon OR fillongate
- pénibilité
- pensions OR pension

- performance énergétique
- période de sûreté
- periscope
- perpétuité
- personnels soignant OR personnel soignant OR personnels soignants
- personnes âgées
- personnes de plus de 55 ans
- personnes handicapées OR handicap OR handicaps OR handicapés OR personnes handicapés
- perte d'autonomie
- perturbateurs endocriniens
- pesticides
- petite enfance
- petites retraite OR petites retraites
- pétrodollars
- peuple
- peurs OR peur
- pib
- pisa
- places de prison
- places en crèche
- plan carcéral
- plan d'investissement
- plan de rénovation
- planète
- planification
- planification écologique
- planification stratégique
- plein emploi OR plein-emploi
- plus de 45 ans
- plus-value des terrains à bâtir
- plus-values OR plus-value
- pma
- pme
- pme européennes
- pmi
- pnru
- pôle emploi
- pôles de santé de proximité
- police OR policiers
- politique agricole commune
- politique commerciale OR politiques commerciales
- politique d'application des peines
- politique de coopération
- politique de défense
- politique de la ville
- politique étrangère
- politique industrielle de défense
- politique nataliste
- politiques culturelles
- politiques publiques
- pollution atmosphérique
- pollutions OR pollution OR polluants OR polluant
- polynésie
- pompiers
- porte-avions
- position dominante
- postes supplémentaires
- poursuites disciplinaires
- poutine OR putin
- pouvoir d'achat
- pré-apprentissage
- précarité
- précarité énergétique
- préfet
- prélèvement à la source
- prélèvements indexés sur le revenu
- prélèvements obligatoires
- premier ministre
- prescription à 10 ans
- prescriptions OR prescription
- présomption d'innocence
- prestations sociales
- prêts aidés
- prévention
- primaire à droite
- primaires citoyennes OR primairescitoyennes
- prime à la détention longue de terrains constructibles
- prime annuelle
- prime pour l'emploi
- primes OR prime
- principe de précaution
- priorité d'accès
- priorité nationale
- prison OR prisons
- prison ferme
- prison supplémentaires
- privatisation OR privatiser OR privatisations
- privilège
- prix des médicaments
- prix planchers
- procédures d'expulsion
- procès
- processus judiciaire
- producteurs
- production OR productions
- production agricole
- production d'électricité
- produit intérieur
- produits agricoles
- produits agricoles et alimentaires
- produits de santé
- produits dérivés spéculatifs
- produits financiers
- professionnels de santé
- professions judiciaires
- professions libérales
- programmes jeunesse
- progrès social
- progressistes OR progressiste
- proportionnelle
- propriétaires
- protection animale
- protection du territoire national
- protection sociale
- protectionnisme
- prothèses auditives
- ptz
- publicité
- puissance européenne
- puissances d'argent
- pupilles de la nation
- qatar
- qualité alimentaire
- quartiers OR quartier
- quartiers populaires
- quartiers prioritaires
- quinoa
- quotas OR quota
- quotas d'importation
- quotient familial
- racines chrétiennes
- racisme OR racism OR raciste
- anti-sémite OR antisémite OR antisémites OR antisémitisme
- xenophobe OR xénophobe
- racisme anti-français
- radicalisation
- rationalisation
- reinformation OR réinformation OR ré-information
- réadmission
- réarmement
- recettes
- recettes publiques
- recherche agronomique
- recherche publique
- récidivistes
- reconduites à la frontière
- reconnaissance
- recours aux génériques
- rectorats
- redistributif citoyen
- redressement OR redresser
- réduction
- réduction de la déductibilité des dépenses de gros travaux en revenus fonciers
- réduction de la dépense
- réduction des déficits
- réduction du taux d'endettement public OR réduction de la dette publique
- réduction du temps de travail
- réduire de moitié l'immigration
- réduire le nombre de parlementaires
- réduire les surcoûts
- référendum
- réforme
- réforme de la dépendance
- réforme de la fiscalité
- réfugiés
- régalien
- régime dérogatoire
- régime fiscal
- régime général
- régime social des indépendants
- régions OR région
- réglementation européenne
- règles du détachement
- regroupement familial
- régularisations

- régulariser des clandestins
- réhabilitation des logements
- réindustrialisation OR réindustrialisation OR réindustrialisation
- religion OR religions OR religieux
- relocalisation
- rembourser OR remboursements OR remboursable OR remboursement
- remises automatiques de peine
- remises de peine
- remplacement OR remplacer
- rémunération du président de la république
- rémunération forfaitaire
- rémunérations OR rémunération
- renégociation
- renouvelables OR renouvelable
- renouvellement urbain
- rénovation
- rénovation thermique
- rénovation urbaine
- renseignement intérieurs
- rentrées fiscales OR recettes fiscales
- représentants de l'état
- représentants des salariés
- reprise d'entreprises
- république
- république française OR république françaises
- réseaux sociaux
- réserve parlementaire
- réservée aux migrants clandestins
- réservistes
- résidences principales
- résistance OR résistances
- responsabilité pénale
- responsable politiques OR responsable politique OR responsables politiques
- ressources naturelles
- restaurants d'entreprise
- restauration collective OR restaurations collectives
- restauration scolaire
- rétention administrative
- retenue à la source
- retraite OR retraités OR retraites OR retrait
- retraite solidaire
- revalorisation OR revaloriser
- revenir à l'équilibre de nos comptes publics
- revenu universel
- revenus de l'assistance
- revenus du capital
- revenus modestes
- révisions constitutionnelles
- révolution citoyenne
- révolution fiscale
- révolution numérique
- rgpp OR révision générale des politiques publiques
- richesse nationale
- robots OR robotisation
- route migratoire
- royaume-uni
- rsa
- rsi
- ruralité
- russes OR russe
- russie
- rythmes scolaires
- sacrifices OR sacrifice
- salafisme OR salafiste OR salafistes
- salaire des fonctionnaires OR point d'indice
- salaire horaire
- salaire minimum
- salaire moyen
- salaires OR salaire
- salarié OR salariés OR salariées OR salarié.e.s
- sanctions
- santé
- santé publique
- savoirs fondamentaux
- schengen
- sciences OR science
- scolarité
- scolarité obligatoire
- scop
- secret de l'instruction
- secteur amateur
- secteur bancaire
- secteur privé
- secteur public
- secteurs régaliens
- secteurs stratégiques
- sécurité
- sécurité collective
- sécurité des français
- sécurité intérieure
- sécurité nationale
- sécurité sociale
- seniors OR senior OR séniors
- sens de l'état
- séparation des pouvoirs
- service civique
- service militaire
- service national
- service public OR services publics
- service sanitaire
- services de l'état
- services innovants
- services publics marchands
- simplification OR simplifier
- situation irrégulière
- small business act
- smart nation
- smic
- smicards
- sncf
- snowden
- société civile
- société du travail
- soins d'urgence
- soins dispensés
- solaire
- soldats
- solidarité OR solidarités OR solidaire OR solidaires
- sols
- solutions de garde
- sommetelevage
- sondage OR sondages
- sortir du nucléaire
- souffrance au travail
- souffrances OR souffrance
- sources d'énergie
- sous-traitants
- souveraineté
- souverainetés démocratiques
- souverainetés nationales OR souveraineté nationale
- spectacle vivant
- spéculation OR spéculateurs
- sport
- sportifs amateurs
- stages OR stage
- start-ups OR startups
- statut pénal du chef de l'état
- stock options OR stock-options
- stratégie de défense
- structures parapubliques
- subprimes
- subvention publique
- successions
- suicides OR suicide
- suppression
- suppression des cotisations minimales
- suppressions de postes
- supprimant OR supprimer
- surendettement
- sûreté
- surplus
- surveillance
- suspension des allocations
- sylvicultures
- syndicat OR syndicats OR syndicaliste OR syndical OR syndicalistes OR syndicalisation
- système de santé
- système scolaire
- système universel
- tafta
- tarification progressive
- tarifs réglementés
- taux d'élucidation
- taux d'imposition
- taux d'usage
- taux de chômage
- taxe OR taxes OR taxation OR taxer OR taxerai
- taxe d'habitation
- taxe foncière
- taxe mondiale sur les transactions financières
- technologies
- technologies numériques

- télé
- téléphones portables
- temps de travail
- temps partiel
- terrains constructibles
- territoire national
- territoires OR territoire
- territoires ruraux
- terrorisme
- terrorisme islamiste
- tgv
- thales
- théorie du complot
- tiers payant
- tisa
- titres de séjour
- tolérance zéro
- totalitarisme islamique
- tourisme
- toxiques
- tpe OR très petites entreprises
- traçabilité
- traité franco-allemand
- traités
- traités européens OR traité européen
- traités internationaux
- traités transatlantiques OR traité transatlantique
- tranche supplémentaire
- transactions financières
- mariage pour tous OR mariage gay
- transactions immobilières
- manif pour tous OR manifpourtous
- transformation économique
- transition agricole
- transition écologique
- transition énergétique
- transition numérique
- transparence
- transports OR transport
- transports régionaux
- transports sanitaires
- travail
- travail clandestin
- travail détaché OR travailleur détaché
- travail obligatoires
- travailleurs
- travailleurs handicapés
- travailleurs indépendants
- très haut débit
- trésor public
- tribunal OR tribune
- tribunal administratif
- tribunal de première instance
- tribunaux correctionnels
- tribunaux pénaux
- tribunaux pénaux de proximité
- tunisie
- turquie
- tutorat
- tva
- tva à 5,5
- tva anti-délocalisations
- uber
- ultra-marins
- souverainiste
- une france puissante
- une honte
- une nouvelle société
- une république forte
- une société du travail
- uniforme à l'école
- universités OR université
- urgences sanitaires
- urssaf
- usines
- vacances
- valeurs républicaines OR valeurs de la république
- valls OR vall
- ve république
- véhicule neuf
- vente à l'étranger
- victime OR victimes
- vidéo-surveillance
- vie politique
- vieux véhicules
- village OR villages
- violencesfaitesauxfemmes OR violences faites aux femmes
- vivre-ensemble OR vivre ensemble
- voile
- voitures OR voiture
- vote utile
- votées OR votés
- washington
- web
- zone euro
- zone inondable
- zone rurale
- zones de sécurité prioritaires
- zones désertifiées
- zones prioritaires OR zone prioritaire
- zones sans réseau

I Analyse des retweets tardifs

Plusieurs auteurs [MMW⁺15, GSAM17] conjecturent que les retweets peuvent être négatifs, notamment lorsqu'ils sont tardifs.

Nous avons étudié le temps de retweet de tout les tweets originaux des comptes que nous suivons⁹. Nous considérons qu'un retweet est tardif lorsqu'il a lieu plus de 14 jours après le tweet original.

Nous avons analysé les tweets originaux ayant au moins 10% de retweets tardifs et nous avons gardé les 100 tweets ayant collecté le plus de retweet (soit 109 retweets minimum). La liste ci-dessous présente pour ces 100 tweets le nom du compte émetteur, le nombre de retweet, la proportion de retweet tardif, le texte du tweet¹⁰, la date du tweet original, la date médiane des retweets tardifs et enfin l'url du tweet. Les retweets que nous avons catégorisés avec certitude comme étant négatifs sont signalés par un point rouge ●. La liste est trié par proportion de retweets tardifs décroissant.

Parmi ces 100 tweets, il y en a 30% qui ont généré des retweets négatifs ou ironiques. Cette proportion de retweets négatifs élevée corrobore donc le lien entre négativité et temps de retweet. Cependant, ceux-ci semblent être limités aux cas extrêmes des tweets ayant énormément de retweet tardif.

- ● **@Fillon2017_fr** 2848 RT (1) : *"Moi, si j'étais mis en examen, je ne serais pas candidat à l'élection présidentielle. Question de morale."* @FrancoisFillon #Bourdindirect - **29/08/2016** 01/03/2017 https://twitter.com/Fillon2017_fr/status/770150054480998400
- ● **@bayrou** 5864 RT (1) : *Emmanuel #Macron est le principal responsable de la politique économique de François #Hollande depuis quatre ans. Pour quel résultat ?* #QDI - **21/09/2016** 25/03/2017 <https://twitter.com/bayrou/status/778677784184971265>
- ● **@bayrou** 297 RT (1) : *La première mesure annoncée par E. #Macron est de proposer que les jeunes travaillent plus pour gagner moins.* #le79inter - **07/12/2016** 22/02/2017 <https://twitter.com/bayrou/status/806405379605274624>
- ● **@bayrou** 1252 RT (1) : *Je ne me reconnais pas ds ce qu'E. #Macron incarne. Son projet de société est proche de celui défendu par N. #Sarkozy en 2007.* #QuestionsPol - **11/09/2016** 22/02/2017 <https://twitter.com/bayrou/status/774926260266999808>
- ● **@bayrou** 116 RT (0.99) : *À propos d'Emmanuel #Macron : Il ne suffit pas de dire qu'on est ni de droite ni de gauche pour être centriste* #E1soir - **22/09/2016** 22/02/2017 <https://twitter.com/bayrou/status/779008508423471104>
- ● **@BrunoLeRoux** 214 RT (0.97) : *Il n'y a pas de complot, François #Fillon doit s'expliquer devant la justice* #LaMatinaleInfo #cnews <https://t.co/SVdDDqT2OU> - **02/03/2017** 21/03/2017 <https://twitter.com/BrunoLeRoux/status/837201126332194816>
- ● **@Fillon2017_fr** 686 RT (0.96) : *"Si j'étais mis en examen je ne me présenterais pas à l'élection présidentielle. C'est évident."* @FrancoisFillon #8h30Apathie - **02/09/2016** 01/03/2017 https://twitter.com/Fillon2017_fr/status/771600966655234048
- ● **@bayrou** 2782 RT (0.96) : *Derrière Emmanuel #Macron il y a des gds intérêts financiers incompatibles ac l'impartialité exigée par la fonction politique* #Bourdindirect - **07/09/2016** 21/04/2017 <https://twitter.com/bayrou/status/773413492963573760>
- **@JeanLucRomero** 1046 RT (0.93) : *Malgré les gesticulations de la #ManifPourTous, 65% des Français pour le maintien de la loi sur le #mariagepourtous L'amour triomphe tjrs !* - **14/09/2016** 12/12/2016 <https://twitter.com/JeanLucRomero/status/776005126901424128>
- ● **@FrancoisFillon** 254 RT (0.93) : *Pas d'auto-rité de l'Etat sans honnêteté! Le pouvoir doit être* *exemplaire pour être vraiment légitime.* #Fillon2017 - **21/09/2016** 01/02/2017 <https://twitter.com/FrancoisFillon/status/778684351013789696>
- **@benoithamon** 266 RT (0.92) : *Triste pour @Cristiano Ronaldo et le Portugal. Énorme joueur.* #POR-FRA - **10/07/2016** 22/01/2017 <https://twitter.com/benoithamon/status/75222252909924353>
- **@jeanlassalle** 384 RT (0.87) : *Et si la France se remettait à marcher sur l'eau ?* #Presidentielle2017 #Résistons <https://t.co/6oKyVtBLzg> - **30/12/2016** 07/04/2017 <https://twitter.com/jeanlassalle/status/814781999625936896>
- **@ramayade** 207 RT (0.85) : *Pour que les femmes ne soient pas représentées que par les extrêmes (Artaud, Le Pen), @ramayade, une candidature nécessaire* #7novembre16h34 - **07/11/2016** 29/01/2017 <https://twitter.com/ramayade/status/795654465374195713>
- **@VigierPhilippe** 327 RT (0.83) : *#Macron en marche aux frais des contribuables. Il ne sert pas la Fr. , il se sert. Nous allons saisir la Haute Auto...* <https://t.co/auOhcikwIV> - **24/01/2017** 07/03/2017 <https://twitter.com/VigierPhilippe/status/823851682626990080>
- ● **@manuelvalls** 419 RT (0.8) : *La primaire est un formidable moyen pour recréer l'unité.* #CeQuiNousRassemble - **05/12/2016** 29/03/2017 <https://twitter.com/manuelvalls/status/805830310155972608>
- ● **@bayrou** 273 RT (0.78) : *Je ne sais pas - à cette date - qui est E. #Macron, quels sont son projet et son positionnement, avec qui il veut gouverner.* #Bourdindirect - **01/02/2017** 22/02/2017 <https://twitter.com/bayrou/status/826699689038905344>
- ● **@FrancoisFillon** 460 RT (0.75) : *Pour gouverner un pays, ma conviction est qu'il faut être irréprochable. Je pose le principe de l'exemplarité du Président et des ministres.* - **18/09/2016** 04/03/2017 <https://twitter.com/FrancoisFillon/status/777431613923885056>
- ● **@BrunoLeMaire** 625 RT (0.75) : *Que les choses soient claires une bonne fois pour toutes : je ne me rallierai à AUCUN candidat au 2nd tour et je ga...* <https://t.co/yihPCmneMN> - **02/10/2016** 20/11/2016 <https://twitter.com/BrunoLeMaire/status/782567180063813634>
- ● **@bayrou** 532 RT (0.73) : *J'ai lu avec beaucoup de soin le discours d'@EmmanuelMacron mais je n'y*

9. Nous limitons l'étude à ces tweets car ce sont les seuls pour lesquelles nous captions les retweet, voir l'annexe B sur la collecte des données.

10. Attention certaines émoticônes ont dû être retiré du texte.

- ai pas trouvé beaucoup de substance. #LeGrandJury - 05/02/2017 22/02/2017
<https://twitter.com/bayrou/status/828208897261445121>
- @Bernard_Monot 193 RT (0.72) : L'islamisme, ce cancer qui détruit tout. #Fillon, #Valls, #Sarkozy ne l'ont pas arrêté. @MLP_officiel le fera!... <https://t.co/iCvsxdKUKj> - 04/01/2017 21/03/2017
https://twitter.com/Bernard_Monot/status/816686129118375936
 - ● @bayrou 213 RT (0.72) : Je ne suis pas pour que le pouvoir de l'argent prenne le pas sur la politique. #Macron #BourdinDirect @BFMTV @RMCinfo <https://t.co/DittqSBegJ> - 07/09/2016 23/02/2017
<https://twitter.com/bayrou/status/773515676711399425>
 - @Marion_M_Le_Pen 3706 RT (0.68) : En avant vers la victoire! #AssisesMLP <https://t.co/4SYKqNoI01> - 04/02/2017 23/04/2017
https://twitter.com/Marion_M_Le_Pen/status/827991739327713280
 - @EmmanuelMacron 17669 RT (0.65) : I have a message for you guys. #Science-March <https://t.co/ZnkFIksdx> - 10/02/2017 08/05/2017
<https://twitter.com/EmmanuelMacron/status/829849353867128832>
 - @benoithamon 660 RT (0.63) : Oui pour le moral j'écoute Still - Dr Dre [#Punchline #Hamon2017](https://t.co/oTSNCazlkJ) - 23/10/2016 22/01/2017
<https://twitter.com/benoithamon/status/790145245262544896>
 - @MAZUEL_2017 128 RT (0.62) : @KingSalman Sir, may Your Majesty pardon young #AliMohammedAlNimr. Philippe Mazuel, candidate to French Presidency <https://t.co/RP3xShqLiB> - 10/09/2016 26/11/2016
https://twitter.com/MAZUEL_2017/status/774722675361087489
 - ● @lepenjm 496 RT (0.62) : Ni Droite, Ni Gauche : #Macron! Droits d'auteur! Mon pronostic : 3 à 4%. - 16/11/2016 07/05/2017
<https://twitter.com/lepenjm/status/798846004556546049>
 - @GLarrive 112 RT (0.62) : La France n'a besoin, ni de la brutalité extrémiste de Marine Le Pen, ni de l'irréalité bougiste d'Emmanuel Macron... <https://t.co/NlPgZ1YSUB> - 23/02/2017 10/04/2017
<https://twitter.com/GLarrive/status/834761651383566337>
 - @LydiaGuirous 411 RT (0.61) : Retrouvez dans @Valeurs ma tribune face à l'islamisme, 2017 sera "Le quinquennat de la dernière chance" <https://t.co/z1Au8CEkx9> - 09/12/2016 08/03/2017
<https://twitter.com/LydiaGuirous/status/807204636641861632>
 - ● @FrancoisFillon 708 RT (0.61) : Beaucoup de Français ont l'impression de travailler pour ceux qui ne travaillent pas. #PrimaireLeDebat - 17/11/2016 01/02/2017
<https://twitter.com/FrancoisFillon/status/799368039552708609>
 - @Simonnet2 377 RT (0.58) : En vidéo l'arnaque de l'ubérisation, paupérisation organisée par des multinationales hors-la-loi! La vidéo complet... <https://t.co/sEKuPuXxr4> - 23/12/2016 04/03/2017
<https://twitter.com/Simonnet2/status/812410976989564929>
 - @ThomasScuderi 110 RT (0.57) : #Jouer-Collectif pour éviter les extrêmes, tel est le défi #citoyen qui réinventera la #Politique! Merci... <https://t.co/MYeodTN1OH> - 13/11/2016 20/12/2016
<https://twitter.com/ThomasScuderi/status/797771376094564352>
 - @marseille 149 RT (0.57) : Suivez toute l'actualité de Marseille sur Twitter - 20/01/2017 09/02/2017
<https://twitter.com/marseille/status/822396311722807296>
 - @DamienObrador 113 RT (0.57) : La #Patriosphere mobilisée pour la victoire de @MLP_officiel en 2017! #AvecMarine <https://t.co/wmpzFFa35g> - 16/10/2016 16/04/2017
<https://twitter.com/DamienObrador/status/787727421898391553>
 - @OuchikhKarim 436 RT (0.54) : La République est laïque, mais la France est chrétienne. Une vérité qui s'impose à tous, y compris en politique <https://t.co/jHpI0u6Izu> - 04/09/2016 05/11/2016
<https://twitter.com/OuchikhKarim/status/772411000637317121>
 - @JLMelenchon 349 RT (0.54) : @LiemHoangNgoc : Nous allons revaloriser les petites retraites et augmenter le minimum vieillesse. #JLMChiffrage... <https://t.co/HS2gubHopW> - 19/02/2017 16/03/2017
<https://twitter.com/JLMelenchon/status/833320688048099328>
 - ● @FrancoisFillon 372 RT (0.54) : Il faut une justice rapide et ferme. L'impunité zéro doit être la règle! #PalaisDesCongrès <https://t.co/puyJLiyuLj> - 18/11/2016 01/03/2017
<https://twitter.com/FrancoisFillon/status/799690695883657216>
 - @Simonnet2 155 RT (0.53) : Pourquoi le Gvt n'engage pas d'enquête sur Lafarge après les révélations du Monde sur sa collaboration avec #Daech? <https://t.co/q6paaQz4uP> - 06/07/2016 13/08/2016
<https://twitter.com/Simonnet2/status/750711159737741312>
 - ● @FrancoisFillon 133 RT (0.53) : De nous, les Français attendent transparence et intégrité : pour rétablir ordre et confiance, l'exemple doit venir d'en haut. #CNLR - 02/07/2016 01/02/2017
<https://twitter.com/FrancoisFillon/status/749233025184108544>
 - @PhilippePoutou 256 RT (0.52) : Vu sur un mur : "Pour une abstention utile, votez Poutou" <https://t.co/UE3XKkODuh> - 24/11/2016 04/04/2017
<https://twitter.com/PhilippePoutou/status/801892217803722752>
 - @EmmanuelMacron 157 RT (0.52) : C'est un livre où j'explique la vision que j'ai du pays. #Révolution <https://t.co/zwyXTULNaN> - 24/11/2016 27/12/2016
<https://twitter.com/EmmanuelMacron/status/801850774594736128>
 - @Marion_M_Le_Pen 663 RT (0.49) : "#Fillon est pour le droit du sol, le regroupement familial, contre la restauration des frontières, contre la prior... <https://t.co/uYH3mZxDtJ> - 15/12/2016 13/03/2017
https://twitter.com/Marion_M_Le_Pen/status/809324258312781826
 - @PhdeVilliers 151 RT (0.48) : Mathieu Bock-Côté : « Un certain catholicisme se mue en utopie multiculturaliste » <https://t.co/pg5sXONbU7> - 24/01/2017 05/03/2017
<https://twitter.com/PhdeVilliers/status/823894913678974976>
 - ● @manuelvalls 234 RT (0.46) : Le revenu universel que je défends, c'est une grande innovation sociale : garantir un revenu à tous et à tout moment... <https://t.co/zfbsZswUV> - 27/10/2016 24/01/2017
<https://twitter.com/manuelvalls/status/791653817544826881>
 - @lesRepublicains 155 RT (0.46) : Trouvez le bureau de vote le plus proche de chez vous pour voter à la #Primaire2016 des 20 & 27 novembre! <https://t.co/Eb1SFDuU5e> - 13/10/2016

- 09/11/2016
<https://twitter.com/lesRepublicains/status/786494987080990720>
- @OuchikhKarim 277 RT (0.45) : Pour #2017, je confirme ma position au #Scan : «J'appelle à voter pour #MarineLePen». «Mon cœur ne balance pas» ! <https://t.co/4Qv1ufodst> - 14/10/2016 03/11/2016 <https://twitter.com/OuchikhKarim/status/786965898540220416>
 - ● @lepenjm 1749 RT (0.44) : Un éléphant, ça TRUMP, ça TRUMP, un éléphant ça TRUMP énormément... (l'éléphant est l'animal emblème du Parti Républicain) - 20/07/2016 09/11/2016 <https://twitter.com/lepenjm/status/755809512649547776>
 - @alainhoupert 503 RT (0.43) : «La seule activité continue de l'Élysée, c'est l'hostilité à Sarkozy» Jean d'Ormesson <https://t.co/U3HBfMrze7> - 02/10/2016 14/11/2016 <https://twitter.com/alainhoupert/status/782483601225113600>
 - @PhdeVilliers 150 RT (0.41) : Alexandre Devecchio : «Derrière l'affaire Théo, les banlieues en sécession» <https://t.co/mPmz4ZXqre> - 14/02/2017 27/03/2017 <https://twitter.com/PhdeVilliers/status/831510972661641216>
 - @Paris 118 RT (0.41) : .@RElementaires organise une collecte de tampons et serviettes pour la #JournéeInternationaleDesDroitsDesFemmes ... <https://t.co/Mje4mQdWg3> - 08/03/2017 25/03/2017 <https://twitter.com/Paris/status/839509879316443137>
 - @SIEL_Off 146 RT (0.4) : Le #SIEL apporte son soutien à #MarineLePen aux #Présidentielle2017 et aux candidats SIEL aux législatives <https://t.co/ELAHoyhjzY> - 26/02/2017 29/03/2017 https://twitter.com/SIEL_Off/status/835950123855282176
 - ● @FrancoisFillon 303 RT (0.4) : Au sommet de l'Etat, l'intégrité du Président et de ses ministres doit être irréprochable car il n'y a pas d'autorité sans exemplarité. - 18/11/2016 31/01/2017 <https://twitter.com/FrancoisFillon/status/799696415555551232>
 - @Paris 552 RT (0.39) : Parc #RivesDeSeine, venez admirer Paris au fil de l'eau <https://t.co/9gfz5Uy1ln> <https://t.co/8e5jQFSFL1> - 22/03/2017 20/04/2017 <https://twitter.com/Paris/status/844513445823303680>
 - @EmmanuelMacron 1680 RT (0.39) : Non, nous ne devons pas accepter qu'en Tchétchénie des camps pour homosexuels soient réouverts. #MacronPrésident <https://t.co/wriwoHdvoc> - 17/04/2017 02/05/2017 <https://twitter.com/EmmanuelMacron/status/854021146924240896>
 - ● @DLF_Officiel 113 RT (0.39) : .@dupontaignan "Je suis un gaulliste social. Chez DLF, il n'y a pas d'arrière boutique d'extrême droite comme au #FN" #LEntretienPolitique - 15/03/2017 29/04/2017 https://twitter.com/DLF_Officiel/status/842096496417898503
 - @alainjuppe 163 RT (0.39) : Je crois à l'esprit Bordelais. Une ville fidèle à son esprit girondin, à la tempérance, à la modération. Tout ne vient pas d'en haut. - 12/01/2017 06/02/2017 <https://twitter.com/alainjuppe/status/819500637083881473>
 - ● @FrancoisFillon 146 RT (0.37) : Un délit : une sanction ! Voilà la règle qui doit être parfaitement claire pour les délinquants. #PalaisDesCongrès - 18/11/2016 31/01/2017 <https://twitter.com/FrancoisFillon/status/799690849311490048>
 - @OuchikhKarim 206 RT (0.36) : L'#islam est incompatible avec la #République : il doit être mis sous tutelle ! Ma tribune sur @BVoltaire <https://t.co/LvvLYDoClJ> - 16/11/2016 10/12/2016 <https://twitter.com/OuchikhKarim/status/798787380496863232>
 - ● @FrancoisFillon 196 RT (0.36) : Pour être Président de la République, il faut être soi-même irréprochable. #BourdinDirect - 18/11/2016 29/01/2017 <https://twitter.com/FrancoisFillon/status/799519082538209280>
 - @UPR_Asselineau 482 RT (0.35) : Frédéric Robert, maire de Monétier-Allemont : «Non Mr @EmmanuelMacron, La commune, La France ne sont pas à vendre!... <https://t.co/DmQTXqqPDN> - 26/03/2017 04/05/2017 https://twitter.com/UPR_Asselineau/status/846108489479000064
 - ● @FrancoisFillon 249 RT (0.35) : Nous devons entendre les Français. Nous devons entendre leur demande d'honnêteté, de courage et de vérité. #CNlesRepublicains - 14/01/2017 01/02/2017 <https://twitter.com/FrancoisFillon/status/820233545046171648>
 - ● @FrancoisFillon 152 RT (0.35) : C'est le cri de désespoir des Français lassés de travailler pour ceux qui ne travaillent pas qui a guidé la rédaction de mon programme. - 03/09/2016 25/01/2017 <https://twitter.com/FrancoisFillon/status/772024442377670656>
 - @dupontaignan 534 RT (0.35) : Retrouvez mon clip officiel de campagne pour la #présidentielle A découvrir et à partager ! #NDA2017 #DeboutLaFrance <https://t.co/sffjfbkJc> - 28/03/2017 20/04/2017 <https://twitter.com/dupontaignan/status/846759051505651712>
 - @RafikSmati 155 RT (0.33) : Ce que cache vraiment la candidature Macron... <https://t.co/RIwz4VnirU> - 15/01/2017 13/02/2017 <https://twitter.com/RafikSmati/status/820611357456404480>
 - @LydiaGuirous 166 RT (0.33) : Ce n'est pas à la République de s'adapter à l'islam mais à l'islam de s'adapter à la République. #Laïcité ne se négocie pas #Valls ! @leJDD - 31/07/2016 26/08/2016 <https://twitter.com/LydiaGuirous/status/759687863072071680>
 - @LaurencePache 269 RT (0.33) : Merci Mr #Vidalies, pour l'ouverture à la concurrence des #TER, pour l'augmentation des péages autoroutiers, pour l... <https://t.co/orfp5jUWJm> - 19/09/2016 10/01/2017 <https://twitter.com/LaurencePache/status/777993712546811904>
 - @LGRANDGUILLAUME 131 RT (0.32) : #taxis #VTC #loti Je serai vers 18h Gare de Lyon pour commencer la distribution pour la loi <https://t.co/ZOJA9GDXJe> - 23/10/2016 12/11/2016 <https://twitter.com/LGRANDGUILLAUME/status/790191309764038657>
 - @LellouchePierre 109 RT (0.32) : À libre absolu dans @lemondefr, le réquisitoire d'@Anne_Hidalgo contre @EmmanuelMacron. La meilleure attaque à... <https://t.co/eGjmsRsa76> - 12/01/2017 02/05/2017 <https://twitter.com/LellouchePierre/status/819611464168308739>
 - @JeanLucRomero 1248 RT (0.32) : Tweetons tous cette campagne de prévention #VIH que fustigent tous les homophobes et bigots d'un autre temps!... <https://t.co/vho592qDfv> - 18/11/2016 17/01/2017 <https://twitter.com/JeanLucRomero/status/799681071214428160>
 - ● @FrancoisFillon 130 RT (0.3) : Il n'y a pas d'autorité sans prestige, sans une certaine dignité

- et une intégrité certaine. #PalaisDesCongres - 18/11/2016 01/02/2017
<https://twitter.com/FrancoisFillon/status/799684116941676544>
- @benoithamon 372 RT (0.3) : Saisissons la chance de ne pas subir le dérèglement climatique, les mutations de la société, de l'industrie, mais d'e... <https://t.co/omCWDiNrA> - 19/03/2017 04/04/2017
<https://twitter.com/benoithamon/status/843493085770334208>
 - @pierrezeniziot 399 RT (0.29) : Ce moment gênant où tu vas sur le site de #Macron et où il n'y a pas d'onglet #projet... <https://t.co/mHx0WJVxSn> - 21/01/2017 05/02/2017
<https://twitter.com/pierrezeniziot/status/822908822608023552>
 - @PhdeVilliers 776 RT (0.29) : Philippe de Villiers pense à s'afficher aux côtés de Marine Le Pen <https://t.co/xUd3iiVfdF> - 27/03/2017 29/04/2017
<https://twitter.com/PhdeVilliers/status/846416458594832384>
 - @LellouchePierre 125 RT (0.29) : Je le dis avec gravité, la question de l'application de l'art. 68 de la Constitution (destitution pour manquement à ses devoirs) est posée - 19/10/2016 07/11/2016
<https://twitter.com/LellouchePierre/status/788754826712719361>
 - @UPR_Asselineau 158 RT (0.27) : ÉVÉNEMENT : Grand rassemblement de l'UPR le 25 mars à Paris! Réservez votre place sur notre site! #Asselineau2017 <https://t.co/hHAEPnOfyt> - 17/02/2017 05/03/2017
https://twitter.com/UPR_Asselineau/status/832525210729451520
 - @Marion_M_Le_Pen 771 RT (0.27) : #LesRépublicains : "Plutôt que de vous contenter d'un plan B, tournez-vous vers le plan M, le plan Marine" #Fillon <https://t.co/r3rV34j7fB> - 06/02/2017 06/03/2017
https://twitter.com/Marion_M_Le_Pen/status/828571665315532801
 - @Fillon2017_fr 434 RT (0.27) : "Je veux une classe politique exemplaire, avec des ministres qui ne sont pas mis en examen." @FrancoisFillon #Avec-FillonElevonsLeDébat - 24/11/2016 01/03/2017
https://twitter.com/Fillon2017_fr/status/801882003536805890
 - @najatvb 137 RT (0.26) : Avec ce livre, "La vie a plus d'imagination que toi", j'ai voulu parler de la France telle que je la connais, telle... <https://t.co/zLStE8DUCN> - 31/03/2017 30/04/2017
<https://twitter.com/najatvb/status/847764078609260544>
 - @JLMelenchon 919 RT (0.26) : #Juppé et #Fillon ont parlé de l'ISF et des riches. Mais pas un mot sur les 9 millions de pauvres en France... <https://t.co/0pJbZHRpxv> - 24/11/2016 01/04/2017
<https://twitter.com/JLMelenchon/status/801905686296657920>
 - @DidierTauzin 174 RT (0.26) : Exigeons un casier judiciaire vierge à nos hommes politiques! Ils doivent être exemplaires. <https://t.co/VrvvAui83D> - 05/11/2016 11/12/2016
<https://twitter.com/DidierTauzin/status/794954967840854020>
 - @valerieboyer13 163 RT (0.25) : Porter le voile est perçu comme un signe d'allégeance à notre ennemi. Je proposerai un texte pour y mettre un terme, luttons c/ le salafisme - 11/09/2016 22/11/2016
<https://twitter.com/valerieboyer13/status/774908438417670144>
 - @NicolasSarkozy 3203 RT (0.25) : Je souhaite le meilleur pour mon pays et pour celui qui aura à conduire la France que j'aime tant - NS <https://t.co/xNYy61YUVV> - 20/11/2016 30/03/2017
<https://twitter.com/NicolasSarkozy/status/800451750045970432>
 - @MLP_officiel 248 RT (0.25) : "L'immense majorité des homosexuels ne demandent pas le mariage gay. Ils ne réclament qu'un PACS amélioré." #BFM-Politique - 16/10/2016 28/04/2017
https://twitter.com/MLP_officiel/status/787604569593356289
 - @MLP_officiel 648 RT (0.24) : "On doit renvoyer les clandestins chez eux et maîtriser nos frontières nationales, sinon les camps se reconstitueront... <https://t.co/HEZ64kypQn> - 24/01/2017 25/03/2017
https://twitter.com/MLP_officiel/status/823916093064548355
 - @MLP_officiel 611 RT (0.24) : "Nous vivons la fin d'un monde et la naissance d'un autre! C'est le retour des États-Nations!" #Koblenz <https://t.co/IHH2wFSMz4> - 21/01/2017 20/03/2017
https://twitter.com/MLP_officiel/status/822762303925981184
 - @manuelvalls 1483 RT (0.24) : Je suis candidat à la présidence de la République. MV - 05/12/2016 29/01/2017
<https://twitter.com/manuelvalls/status/805827675482951680>
 - @LydiaGuirous 562 RT (0.24) : Message aux "féministes" démissionnaires de gauche qui défendent #burkini voile burqa... au nom d'une fausse Liberté <https://t.co/ku378P1PNf> - 07/09/2016 20/10/2016
<https://twitter.com/LydiaGuirous/status/77358642534222337>
 - @Simonnet2 281 RT (0.23) : Il y a pas écrit La Poste, là! Ne laissons pas les supérettes faire main basse sur la ville, et sur le service pub... <https://t.co/9AJZgoffI4> - 20/11/2016 16/12/2016
<https://twitter.com/Simonnet2/status/800321460078252033>
 - @MLP_officiel 447 RT (0.23) : "Les pompiers et les policiers sont immobilisés chaque jour pour régler les problèmes liés aux jungles de #migrants... <https://t.co/ck0qUBXXhG> - 24/01/2017 26/03/2017
https://twitter.com/MLP_officiel/status/823916442613612544
 - @GilbertCollard 223 RT (0.23) : Et vous demandez comment on peut faire des économies : <https://t.co/5UtaUJDxLV> - 06/01/2017 22/03/2017
<https://twitter.com/GilbertCollard/status/817302224040853504>
 - @ECiotti 1013 RT (0.23) : Je propose à @EliseLucet d'enquêter sur le financement de la campagne de M. #Macron, elle qui a enquêté sur M... <https://t.co/ABahs9ombb> - 06/02/2017 11/03/2017
<https://twitter.com/ECiotti/status/828542977815494656>
 - @valerieboyer13 245 RT (0.21) : L'assassinat de #ZhangChaolin passe sous silence...! <https://t.co/BbGwDjK86w> - 15/08/2016 05/09/2016
<https://twitter.com/valerieboyer13/status/765314139971526657>
 - @valerieboyer13 184 RT (0.21) : Jamais nous ne laisserons prendre en otage ni la démocratie ni le destin de notre pays. Soutien à @Francois-Fillon <https://t.co/VCbOpTfO6Q> - 02/02/2017 01/03/2017
<https://twitter.com/valerieboyer13/status/827165016684130304>
 - @Paris 198 RT (0.21) : #ParisPlages revient du 20/07 au 04/09.

- Le programme complet <https://t.co/hQoJzM430L>
<https://t.co/60ZVJhVPT2> - 19/07/2016
17/08/2016
<https://twitter.com/Paris/status/755342759968530432>
- @OuchikhKarim 308 RT (0.21) : *Qui peut croire un seul instant que #Macron soit capable de garantir aux Français une majorité pour gouverner!* <https://t.co/nr4cDS3HpA> - 09/03/2017 27/03/2017
<https://twitter.com/OuchikhKarim/status/839875214079057922>
 - @MLP_officiel 476 RT (0.21) : *"Il n'y a rien pour nous de plus beau que la France, il n'y a rien pour nous de plus grand que la France!"*... <https://t.co/1jeyJvAeka> - 05/02/2017 21/03/2017
https://twitter.com/MLP_officiel/status/828249040450101250
 - @GilAverous 335 RT (0.21) : *Si on me demande pourquoi je soutiens @NicolasSarkozy, la réponse est là ↓ #ToutPourLaFrance* <https://t.co/u2lSs9ohE9> - 16/09/2016 15/10/2016
<https://twitter.com/GilAverous/status/776765712232747008>
 - @francoisbaroin 163 RT (0.21) : *Rassemblement to-*
 - tal autour de @FrancoisFillon pour la présidentielle.* 2/2 - 27/11/2016 31/01/2017
<https://twitter.com/francoisbaroin/status/802974514191728641>
 - @alainjuppe 668 RT (0.21) : *Je me sens plus proche du Pape François que de Sens Commun et de la #ManifPourTous #JuppéPrésidentCest sur @France2tv #20H* - 21/11/2016 16/04/2017
<https://twitter.com/alainjuppe/status/800781567090130944>
 - @OuchikhKarim 560 RT (0.2) : *#Macron est aussi l'homme qui a vendu #Alstom aux Américains. La France bradée, c'est lui!* <https://t.co/B3H50PsoFH> <https://t.co/6nAy6ipJte> - 05/02/2017 05/04/2017
<https://twitter.com/OuchikhKarim/status/828193741504651264>
 - @Marion_M_Le_Pen 2499 RT (0.2) : *"Tous les musulmans ne sont pas djihadistes, mais tous les djihadistes sont musulmans."* #BBRGrandSud <https://t.co/Kq7TfXzIJW> - 09/07/2016 29/10/2016
https://twitter.com/Marion_M_Le_Pen/status/751776519538745344