



**HAL**  
open science

## Rapport Scientifique projet RSSA

Camille Baulant, Guillaume Sylvestre

► **To cite this version:**

Camille Baulant, Guillaume Sylvestre. Rapport Scientifique projet RSSA. Centre des Hautes Etudes du Ministère de l'Intérieur; Laboratoire de Recherche de l'Ecole Nationale Supérieure de la Police. 2020, pp.94. hal-04879148

**HAL Id: hal-04879148**

**<https://hal.science/hal-04879148v1>**

Submitted on 10 Jan 2025

**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.

# RAPPORT SCIENTIFIQUE PROJET RSSA

*Analyse des Réseaux Sociaux pour améliorer l'Anticipation et l'Adaptation de la réponse d'ordre public*

**Sous la direction de Camille Baulant et Guillaume Sylvestre**

**Financé par du Centre des Hautes Études du Ministère de l'Intérieur et réalisé en lien avec le Centre de Recherche de l'Ecole Nationale Supérieure de Police**

## Table des matières

Résumé du rapport de recherche .....	3
Présentation de l'équipe .....	5
<b>I. Rappel du cadre du projet de recherche .....</b>	<b>6</b>
1. Objectifs du projet et réponse du Granem .....	6
2. Exploitation de l'intelligence artificielle dans le cadre de ce projet.....	6
<b>II. Présentation de la méthodologie .....</b>	<b>8</b>
1. Les enjeux de l'analyse automatique des données issues des réseaux sociaux.....	8
2. L'analyse de graphe social pour interpréter les données de Twitter.....	11
3. Le cadre théorique proposé par l'équipe du projet .....	14
4. Le Social Graph Learning, application de l'IA à la Social Network Analysis .....	21
<b>III. Enseignement des analyses des corpus de données .....</b>	<b>22</b>
1. Les sujets de controverse publique : Anti-spécisme, LaPurge .....	22
2. Les sujets de militantisme sociétal : Pacte de Marrakech, Extinction Rebellion à Italie 2 .....	25
3. Les sujets portés par des acteurs locaux : Europacity, Extinction Rebellion Lafarge, ZAD Dune.....	30
<b>IV. Résultats de l'application de l'IA.....</b>	<b>34</b>
1. Une méthodologie rigoureuse pour s'assurer de la qualité des données avant l'analyse .....	34
2. Procédure d'analyse des données avec l'outil DataRobot .....	35
3. Mise en œuvre des modèles de machine learning.....	37
4. Nos pistes de travail pour perfectionner le modèle IA.....	39
Conclusion .....	40
V. Bibliographie .....	41
VI. Annexes .....	44
1. Analyse des échanges Twitter sur les manifestations anti-spécistes en septembre 2018 .....	45
2. Analyse des échanges Twitter sur #Lapurge en octobre 2018.....	54
3. Analyse des échanges Twitter sur le Pacte de Marrakech en novembre – décembre 2018 .....	58
4. Analyse des échanges Twitter sur Extinction Rebellion à Italie 2 en octobre 2019 .....	64
5. Analyse des échanges Twitter sur Notre-Dame-des-Landes en janvier 2016 .....	70
6. Analyse des échanges Twitter sur Europacity en octobre 2019.....	74
7. Analyse des échanges Twitter sur la ZAD Dune en janvier 2020.....	84
8. Analyse des échanges Twitter sur Extinction Rebellion et Lafarge en février 2020 .....	88

## Résumé du rapport de recherche

L'analyse des échanges sur les réseaux sociaux est un sujet complexe. Si de nombreux travaux de recherche s'y intéressent, ceux-ci vont privilégier des méthodologies d'analyse sémantique, en s'appuyant sur des corpus de données limités par les capacités de collecte des outils. C'est un double écueil : l'analyse sémantique n'est efficace que pour interpréter des **éléments connus**, et elle nécessitera un gros travail en amont pour préparer les données.

Le champ de la social network analysis est une première piste pour mieux analyser les échanges sur les réseaux sociaux. En s'intéressant aux liens entre les individus représentés par leurs interactions sociales, qui se retrouvent via les mentions et partages sur les réseaux sociaux, **on peut dépasser les limites de l'analyse sémantique**. La social network analysis permet notamment d'identifier précisément les entités d'un corpus qui ont le plus contribué à diffuser l'information, et de les caractériser, via différents algorithmes, avec l'outil Gephi. **Gephi permet de comprendre les raisons de la forte diffusion d'une information spécifique**, en dehors d'une actualité générant une forte réaction émotionnelle : légitimité d'un expert, opération d'influence, action militante ... en vérifiant les données brutes via une plateforme collectant les données.

En choisissant le réseau social Twitter, avec des données obtenues via un outil professionnel, il est possible de limiter au maximum les biais liés à la collecte de données. Twitter permet d'accéder à toutes les informations ; un outil spécifique à Twitter permet de collecter les données avec des règles complexes pour limiter les risques de perte d'informations ; si cet outil est bien utilisé, les erreurs d'exports de données seront évitées. Afin d'obtenir des exports de données exploitables pour notre projet et de limiter les risques d'erreurs dans la sélection de l'information, **nous avons utilisé la plateforme Visibrain**.

**Le choix de Twitter est rendu essentiel par l'objet de l'étude : identifier via l'intelligence artificielle des indicateurs d'une viralité potentielle de la diffusion de l'information**. Si l'on veut entraîner efficacement une IA sur plusieurs corpus de données, **il faut que ces corpus soient construits sur des bases comparables**. Les données limitées de Facebook, Instagram ou Snapchat ne sont pas exploitables pour cet objet. Néanmoins, la mise en place d'un système d'alerte via Twitter permettra ensuite de compléter / vérifier les signaux obtenus via des recherches complémentaires sur ces réseaux sociaux.

Nous appliquons dans ce projet une démarche innovante de **Social Graph Learning**, qui consiste à entraîner des modèles de machine learning à partir des calculs statistiques obtenus via la social network analysis. Afin de qualifier précisément les corpus de données Twitter qui seront exploités, ceux-ci ont été pré sélectionnés et étudiés via une méthodologie rigoureuse, définie dans un cadre universitaire avec des étudiants du master 2 Intelligence Economique et Stratégies Compétitives à l'International d'Angers.

Les corpus de données sont ensuite analysés via un outil de social network analysis pour identifier précisément les diffuseurs de l'information, et contextualiser le sujet étudié. **Cette analyse est complétée par les travaux de recherche du laboratoire Granem sur les différents types de leaders (leader hiérarchique, leader économique, leader go-between), qui se révèlent cruciaux dans les réseaux économique humains**.

Cette pré analyse permettra, dans une seconde étape, d'identifier les comptes Twitter qui pour chaque corpus ont contribué à diffuser l'information avant que celle-ci devienne virale. Ces comptes Twitter seront les cibles de dizaines de modèles de machines learning exploités via un outil d'IA qui permet de comparer précisément leur efficacité, DataRobot.

Les modèles de machine learning les plus efficaces ont ainsi été capables d'identifier les caractéristiques permettant de retrouver les comptes Twitter responsables de la viralité d'un sujet donné sans exploiter de données personnelles, à partir de la dynamique des échanges Twitter. Il est donc possible d'identifier, en ciblant un corpus de données précis, si un sujet deviendra viral ou non.

Pour être exploitable en conditions réelles, **il sera nécessaire d'affiner le modèle le plus performant via des tests sur d'autres corpus**. Les tests effectués ont démontré que les sujets analysés avaient un faible impact sur les résultats, ce qui facilitera la constitution de corpus tests.

## Présentation de l'équipe

Madame Camille BAULANT, fondatrice et directrice du Master Intelligence Economique et Stratégies Compétitives à l'Internationale (Université d'Angers), est Professeur agrégé des Facultés de Sciences Economiques depuis 1994.

Rattachée au laboratoire Granem, ses champs de recherche combinent l'économie internationale (compétitivité, IDE, risque pays, relations méso-économiques, ...) et l'économie de la connaissance (leadership transformationnel, coopération, réseaux, clusters ...). Elle publie et participe régulièrement à des colloques de niveau international.

Monsieur Guillaume SYLVESTRE, expert en Social Network Analysis, intervenant dans le master Intelligence économique et stratégies compétitives à l'international. Guillaume Sylvestre a plus de dix ans d'expérience dans l'analyse des données issues du web en général, et des réseaux sociaux en particulier. Directeur de l'innovation numérique à l'ADIT, leader européen de l'intelligence économique, il développe des méthodologies innovantes d'analyse de l'information combinant intelligence humaine et outils de collecte et d'analyse des données.

Son expertise couvre notamment l'impact stratégique de l'innovation sur les secteurs économiques et l'identification des signaux faibles via les réseaux sociaux. Il donne des cours à l'Ecole de Guerre Economique et l'IHEDN sur l'analyse des données et la datavisualisation, et intervient à l'Université d'Angers sur des projets de recherche liant Big Data, Réseaux sociaux et IA.

Avec la participation des étudiants de la promotion 2019-2020 master 2 Intelligence Economique et Stratégie compétitives de l'université d'Angers : André DAGO, Antonin ROHARD, Han Hoan NGUYEN (M1), Ghita MARRAKCHI, Hajar NAQACHI, Marvin EMERY, Sylman TEULIERES, Ruxanda BOTEA, Denisa APOSTOL, Céline METAIS, Clarisse BOUET, Emeline Launay, Mengyao WANG, Zakaria TISSAFI IDRISSE.

Et de Madame Mouna EL HADDANI, docteure en intelligence économique du Granem, spécialiste de la veille web et de la veille terrain, qualifiée en section information et communication en mars 2020, ATER, intervenante en logiciel de veille (Sindup, Ami Software) dans le master Intelligence économique et stratégies compétitives à l'international.

Camille Baulant et Guillaume Sylvestre encadrent la thèse CIFRE (financée par GFI Informatique) de Julian Rioche sur Le changement de paradigme lié à l'Internet of Things, étude des enjeux, des conséquences et des externalités générées dans le domaine de la e-santé.

Le GRANEM (Groupe de Recherche ANgevin en Économie et Management) est une équipe de recherche pluridisciplinaire (économie et gestion) de l'Université d'Angers, rattachée à la Fédération CNRS n°3435 TEPP (Travail, Emploi et Politiques Publiques). Cette équipe d'accueil n° 7456 regroupe 51 enseignants-chercheurs, 43 doctorants, et 4 BIATSS.

# I. Rappel du cadre du projet de recherche

## 1. Objectifs du projet et réponse du Granem

L'intitulé de l'appel d'offre du Centre des Hautes Études du Ministère de l'Intérieur était « *Les nouveaux usages des réseaux sociaux dans les nouvelles dynamiques de contestation (nuit debout, gilets jaunes, extrémisme) afin d'améliorer l'anticipation et l'adaptation de la réponse d'ordre public à travers l'utilisation de l'intelligence artificielle.* »

Nous avons proposé d'adresser les éléments suivant via notre méthodologie de recherche :

- Améliorer l'analyse des réseaux sociaux dans les nouvelles dynamiques de contestation (nuit debout, gilets jaunes, extrémisme) afin d'améliorer l'anticipation et l'adaptation de la réponse d'ordre public ;
- Identifier l'apparition de nouvelles thématiques / signaux faibles portés par des groupes spécifiques dans le champ de la contestation sociale ;
- Caractériser la typologie des échanges liés à ces signaux faibles pour comprendre la logique de structuration du mouvement : centralisée, éclatée, indépendante ;
- Souligner la présence ou non d'éventuels relais au sein des communautés de l'information vérifiée / non déformée.

Pour ce faire, tous les corpus de données sélectionnés pour entraîner l'IA ont été analysés par l'équipe de recherche et le groupe étudiant. Cela nous a permis d'en définir une typologie et des caractéristiques spécifiques.

## 2. Exploitation de l'intelligence artificielle dans le cadre de ce projet

Nous travaillerons sur les données issues de Twitter, seul réseau social qui donne accès à l'ensemble de ses données, ce qui implique certains biais qui seront pris en compte dans l'analyse. **Nous nous intéresserons aux échanges qui sur une période allant d'une semaine à un mois donnent lieu à une très forte viralité, en lien avec des manifestations publiques.** Notre premier objectif sera d'automatiser la détection de cette viralité sur des corpus de données spécifiques.

La majorité des travaux d'analyse de données des réseaux sociaux cherchent à définir un modèle global qui s'appliquerait à tous les cas : langues, thématiques, temporalité ... pour des objets larges : Détection d'un bad buzz sur une marque, identification de fausses informations, anticipation de campagnes d'influence ... **Pour autant, chaque modèle développé va fonctionner sur quelques cas spécifiques liés à la préparation des données.**

L'essence de notre travail ne consiste pas dans la mise au point d'une intelligence artificielle mais dans **la définition d'une méthodologie de qualification et de pré analyse des données afin d'identifier un modèle de machine learning efficace pour détecter la viralité**. Sans données pertinentes, aucune IA ne donne de résultats exploitables.

L'intelligence artificielle est une notion très populaire aujourd'hui, pas toujours utilisée à bon escient. La plupart des start-up labellisées « French Tech » indiquant faire de l'intelligence artificielle n'exploitent pas d'IA. Une rapide définition des termes que nous développerons dans notre travail de recherche s'avère donc nécessaire.

**Social Network Analysis / Analyse du graphe social** : il s'agit de l'analyse des structures sociales (connexions humaines notamment) par la cartographie des relations entre les entités. Des logiciels permettent de visualiser ces échanges via un import de données, par exemple issues du réseau social Twitter, et de les analyser via des **algorithmes** et des **calculs statistiques**. Nous utiliserons le logiciel libre et gratuit **Gephi**.

**Algorithme** : Nous exploiterons les **algorithmes** de spatialisation de **Gephi**, qui permettent de positionner les entités d'un réseau sur un graphe en fonction de leurs relations. Le positionnement est lié à la dynamique d'un corpus de données. Cela nous permet de mieux analyser les corpus Twitter qui seront exploités dans ce projet.

**Calcul Statistique** : **Gephi** propose plusieurs dizaines de **calculs statistiques** liés aux propriétés d'un graphe social, qui s'appliquent ensuite à chaque entité. Ils correspondent à des travaux de recherche en sociologie. Les **calculs statistiques** effectués via **Gephi** sur un corpus peuvent être exportés pour y indiquer les **cibles** que l'on cherche à retrouver via l'**IA**.

**Cible** : Dans un projet d'**IA**, la **cible** correspond aux caractéristiques que l'on souhaite retrouver en appliquant un modèle de **machine learning**. Pour ce projet, la cible correspondra aux comptes Twitter responsables de la viralité d'un sujet, que nous aurons identifié via la **social network analysis**.

**Modèle de machine learning** : Le machine learning est un des modes d'exploitation de l'**IA**, qui fonctionne en « **entraînant** » un modèle avec un ensemble de données d'apprentissage indiquant la **cible** à prédire. Il est donc crucial d'utiliser plusieurs corpus de données adaptés aux objectifs et d'en retirer les éléments pouvant biaiser la prédiction. Après plusieurs phases d'entraînement, l'**IA** pourra reconnaître des cibles sur de nouvelles données. Il existe plusieurs modèles de **machine learning** accessibles en open-source, qui fonctionneront sur certains types de données et pas d'autres. Pour un projet innovant, il n'existe pas de modèle adapté, il est donc préférable d'en tester plusieurs.

**Biais** : Dans un contexte d'utilisation de l'**IA**, les biais correspondent à une mauvaise sélection des corpus de données, une méthodologie erronée, une interprétation hasardeuse des résultats du machine learning ... qui amènent des résultats faussés par rapport aux objectifs d'application de l'**IA**. **C'est l'expertise humaine et la méthodologie d'analyse qui permettront d'éviter les biais ou à minima de les prendre en compte dans l'analyse**. Un des biais évidents de l'analyse de données Twitter, ou web au sens large, sera leur manque de représentativité, qui variera selon les sujets étudiés.

**Faux positif** : La sensibilité du modèle de **machine learning** induit que pour obtenir 100 % de détection des **cibles**, il y aura nécessairement des erreurs statistiques. Ce seront les **faux positifs**, c'est-à-dire pour notre projet que des comptes Twitter non responsables de la viralité seront indiqués comme des cibles. Il est important de limiter le nombre de **faux positifs** pour obtenir un modèle efficace.

## II. Présentation de la méthodologie

### 1. Les enjeux de l'analyse automatique des données issues des réseaux sociaux

Les traces numériques en ligne laissées par les internautes, qui permettent de nous renseigner sur l'activité sociale réelle (Rogers, 2015), sont extrêmement nombreuses et riches dans les réseaux sociaux Facebook, Twitter ou LinkedIn qui disposent de centaines de millions d'utilisateurs journaliers dans le monde, plusieurs dizaines de millions en France. Ces services stockent les données des messages postés par leurs utilisateurs et des interactions liées, et se prêtent à des analyses quantitatives et automatisées sur des grands ensembles de données (Cha, Haddadi, Benevenuto, Gummadi, 2010). Les échanges sur les réseaux sociaux sont une matière riche en information, dont le volume ne cesse de croître ... et sont de plus en plus complexe à interpréter.

Les volumes d'informations issues des réseaux sociaux croient de façon exponentielle, rendant très complexe la détection de tendances ou de signaux faibles. Si les réseaux sociaux eux-mêmes proposent des outils dédiés, ceux-ci sont souvent biaisés et de facto des boîtes noires :

- Facebook personnalise les informations remontées sur chaque compte utilisateur, qu'il s'agisse du flux global ou de résultats de recherche, en fonction de l'activité du compte, de celle de son réseau ...
- Twitter propose par zone géographique des « trending topics » qui ont tendance à valoriser des opérations commerciales, et ne permettent pas d'identifier des variations liées à des sujets très commentés habituellement.

Cela signifie que l'information n'est pas diffusée également au sein des réseaux sociaux à leurs utilisateurs. Les algorithmes qui définissent la circulation d'une publication vont la diffuser aux abonnés d'un compte Twitter et Facebook selon les paramètres cités, et en fonction de critères évolutifs (Jammet, 2018). **Il en résulte l'impossibilité de prévoir la viralité potentielle d'un sujet, en dehors d'une information générant une forte réaction émotionnelle** (Alloing, Pierre, 2017). Un seul élément reste cohérent pour interpréter la diffusion d'une publication : l'engagement que celle-ci suscite, qui est recherché par les réseaux sociaux. Les informations suscitant des réactions nombreuses favoriseront la diffusion de publicités aux utilisateurs, première ressource financière des réseaux sociaux. Elles sont donc logiquement mises en avant par les algorithmes des réseaux sociaux.

Les évolutions des réseaux sociaux montrent également la nécessité **de prendre du recul** sur leurs usages. Facebook, après avoir connu de nombreux scandales liés à l'exploitation de l'information de ses utilisateurs et à la manipulation électorale via la propagation d'informations non vérifiées, s'est recentré sur la valorisation des interactions entre les comptes personnels. Les contenus proches géographiquement partagés via des groupes locaux sont également privilégiés (le réseau social incitant par ailleurs les PME à utiliser ses publicités pour cibler leurs zones géographiques), ce qui augmente les risques de manquer des informations sans outils / méthodologie spécifique.

Les algorithmes des réseaux sociaux, en particulier Twitter, favorisent les réactions émotionnelles à l'actualité, ou la controverse entre deux camps irréconciliables. Les "trending topics" correspondent souvent à des réactions à chaud sans lendemain, qui sont difficilement exploitables. Les organisations se retrouvent souvent, selon leurs images, mêlées à ces débats, que cela soit cohérent ou non, comme l'industrie pharmaceutique sur la chloroquine dans le contexte du Covid-19.

**S'il est nécessaire de suivre ces réactions très reprises ponctuellement, il est important pour l'analyse stratégique de pouvoir anticiper des évolutions.** Un bon dispositif de veille de "signaux faibles" liés à des risques doit être capable de repérer des sujets moins porteurs, moins d'actualité, mais qui peuvent préfigurer le buzz de demain. Aujourd'hui, **les outils de veille et d'analyse de l'information des réseaux sociaux sont adaptés à l'identification en temps réel des réactions émotionnelles / des controverses,** mais pas à la détection de nouvelles tendances. Ces plateformes sont très efficaces pour suivre des termes / comptes précis, voir identifier des variations de volumétrie (Guilain, 2015). La détection de l'inconnu, de la nouveauté, ne fonctionnera pas ou seulement dans des cas bien délimités.

L'analyse des informations issues des réseaux sociaux est complexe et problématique à plusieurs niveaux. Si le suivi de la réputation des entreprises et des organisations a permis le développement de nombreux outils et plateformes de veille dédiées, elles montrent très vite leurs limites :

- Faible exploitation des méta-données, qui peuvent difficilement être croisées avec des filtres mots-clés complexes ;
- Ajouts de modules de cartographies / détection d'influenceurs souvent biaisés pour valoriser le client ;
- Débuts d'exploitation de l'intelligence artificielle pour détecter des informations nouvelles mais qui ne seront pas adaptées aux corpus étudiés.

L'analyse de sentiments, mise en place depuis une dizaine d'années, n'a jamais été fiable malgré l'évolution des outils de data mining. La raison en est simple : la compréhension du langage humain écrit et de son contexte n'est pas adaptée aux algorithmes d'analyse automatique de l'information, dès lors que l'interprétation peut varier d'une personne à une autre. Plusieurs travaux soulignent ainsi la difficulté d'analyser la polarité des messages, notamment d'identifier le sarcasme, l'ironie ... et plus simplement d'en comprendre le contexte : les mentions d'un homme politique ne sont pas forcément en lien avec son activité politique en tant que tel (Boyadjian et Velcin, 2017).

De plus, ce type de technologie basé sur des corpus d'informations déjà qualifiées n'arrivera pas à identifier de nouvelles tendances. A contrario, il sera très efficace pour détecter et analyser des éléments connus : problématiques de service client, identification de produits, de personnalités ... l'intelligence artificielle permettra d'effectuer une pré-analyse de ces données.

Certaines plateformes intègrent des modules « d'intelligence artificielle » pour identifier de nouvelles informations, en comparant en temps réel les informations partagées avec le bruit ambiant. Les limites sont là aussi nombreuses, notamment le problème de la langue, puisque les outils sont souvent paramétrés avec des exemples en anglais.

## Tableau comparatif de différents cas d'usage de l'IA dans les plateformes de veille

CAS D'USAGE DE L'IA	VALEUR AJOUTEE	LIMITES
Analyse sémantique de l'historique de la veille pour automatiser la surveillance	Identifier automatiquement des documents similaires à ceux déjà validés	1. Imprécis si pas de possibilité d'adapter l'IA 2. Ne permet pas d'identifier de nouveaux sujets
Détection de « sentiments » sur les réseaux sociaux	Visualiser les éléments négatifs et positifs	Ne fonctionne qu'avec un entraînement sur des données très qualifiées
Qualification de l'information recherchée selon des catégories précises	Mettre en avant les partenariats, les fournisseurs, etc ... autour d'une société	Fonctionnera uniquement dans des contextes prédéterminés
Identification de nouveaux termes viraux	Alerter sur l'apparition de signaux non anticipés	1. Beaucoup de bruit 2. Précision des thématiques
Sélectionner des « influenceurs » sur un corpus	Pointer les utilisateurs les plus influents sur un sujet donné	1. Risque de biais via des robots / opérations d'influence 2. Critères en général non transparents

Source : Camille Baulant et Guillaume Sylvestre, 2020

Il est donc nécessaire d'enrichir les données issues des réseaux sociaux d'éléments qui seront plus interprétables pour une intelligence artificielle, et d'établir des méthodologies d'exploitation de l'intelligence artificielle adaptées au besoin de comprendre la structuration d'un corpus. Il s'agit donc de combiner l'expertise humaine d'analyse des données des réseaux sociaux au travail de data scientistes, qui permet ainsi d'obtenir des résultats plus pertinents (Boyardjian et Velcin, 2017).

**Le réseau social Twitter permet un accès payant à l'intégralité des données publiées de façon publique par ses utilisateurs.** C'est une différence fondamentale avec Facebook, qui donne accès à très peu de contenu, tout comme Instagram, ou encore LinkedIn qui permet des recherches sur le contenu publié mais uniquement de façon limitée. Les données des tweets ne peuvent être obtenues que via des fournisseurs de service partenaires de Twitter, qui devront respecter des conditions d'utilisation strictes sur la réutilisation des données et la protection des informations des utilisateurs. Outre les messages, URL, photos, vidéos et émoticônes partagées, l'accès à ces plateformes permet d'exploiter les métadonnées des tweets : localisation lors de la publication, localisation indiquée par le compte, nombre d'abonnés et d'abonnements, etc ...

Nous avons décidé d'utiliser **la plateforme de veille Visibrain**, qui permet une collecte fine des données, fourniture des exports propres et exploitables notamment pour une analyse par l'IA, et intègre une analyse précise des résultats. Visibrain permet ainsi :

1. L'accès à toutes les données Twitter sans limite dans le temps, et à des dizaines d'opérateurs pour une collecte fine et exhaustive de l'information ;
2. Deux exports de données réexploitables pour l'analyse big data via des outils de business intelligence ou de social network analysis ;
3. Différents modules d'analyse pour faciliter l'interprétation des échanges sur Twitter.

Twitter ne reflète pas l'opinion publique sur un sujet donné, et dispose de moins d'utilisateurs que Facebook en France (environ 10 millions contre 33 millions pour Facebook, la majorité des utilisateurs de Twitter étant par ailleurs inactifs). Pour autant, c'est un outil exploité par tous les politiques et les journalistes (Rieder et Smyrnaio, 2012), et qui permet une très forte exposition publique pour un coût dérisoire. **Il est donc privilégié par les activistes et tout groupe plus ou moins constitué voulant influencer l'opinion.** Ce sont les minorités actives qui utilisent Twitter le plus efficacement. De plus, selon les groupes et les sujets, Twitter sera plus ou moins pertinent pour l'analyse.

Les activistes environnementaux sont historiquement très présents sur Twitter, et ont par réaction amené les groupes agroalimentaires puis les agriculteurs eux-mêmes à investir le réseau social pour influencer le débat public en 2019. A contrario, dans la santé, si les anti-vaccins sont historiquement très actifs, la plupart des patients victimes d'effets secondaires de médicaments en France échangent via des groupes Facebook privés et sont donc très peu visibles.

Il faudra donc adapter l'interprétation des données Twitter à chaque thématique, sans oublier de suivre les autres réseaux sociaux : les gilets jaunes se sont structurés et développés via Facebook (Sebbah B., Marchand P., Souillard N., Loubère L., Smyrnaio N., Renard J., 2019).

## 2. L'analyse de graphe social pour interpréter les données de Twitter

La très grande quantité de données issues des réseaux sociaux en général, et de Twitter en particulier, permet de disposer de modèles d'analyse ne se limitant à l'exploitation du texte ou l'utilisation des méta-données. Il est également possible de détecter des communautés et d'identifier les leaders au sein de ces communautés.

La recherche d'influenceurs est souvent limitée à des informations quantitatives, et donc facilement biaisées. Plusieurs études ont ainsi montré que **le nombre d'abonnés / de like d'un compte sur les réseaux sociaux ne préjuge pas de son éventuelle capacité à influencer** (Cha, Haddadi, Benevenuto, Gummadi, 2010 ; Beauvisage, Beuscart, Couronné, Mellet, 2011). A contrario, les retweets ou mentions montrent une proximité **et des interactions** entre des acteurs (Boyd, Golder & Lotan, 2010).

Les interactions au sein des réseaux sociaux peuvent être représentées et analysées via l'analyse du graphe social, c'est-à-dire l'étude des interactions entre les membres d'un réseau. D'abord appliquée aux réseaux humains dans le cadre de la théorie des graphes, l'analyse du graphe social est aujourd'hui appliquée sur les données Twitter via la Social Network Analysis (Grandjean, 2016).

Le champ de recherche de la social network analysis exploite un outil d'analyse des communautés par la datavisualisation : Gephi. Cet outil, libre et gratuit, développé à l'origine par des étudiants de l'université de Compiègne, est de plus en plus utilisé par les professionnels de l'analyse des réseaux sociaux, car il permet d'interpréter de façon très efficace et précise des grands corpus de données Twitter.

En proposant une cartographie des échanges Twitter à partir des tweets et mentions des comptes, Gephi permet de visualiser le réseau des acteurs (Grandjean, Jacomy et Girard, 2016). La position des comptes Twitter est déterminée par l'algorithme de spatialisation Force Atlas 2 (Jacomy, Venturini, Heymann, Bastian, 2014), qui met en évidence les communautés correspondant aux interactions. Cela permet de constater visuellement l'effet « bulle de filtre » des échanges sur les réseaux sociaux : il est très difficile pour des publications issues d'une communauté donnée d'être partagées par d'autres communautés.

Gephi permet en outre d'utiliser plusieurs calculs statistiques pour analyser les interactions entre les tweets étudiés :

- **Modularity Class** : identifier des communautés au sein des échanges sur Twitter (Blondel, 2008). C'est un calcul statistique qui décompose le réseau en sous-structures selon un score affecté à chaque entité.

Chaque compte Twitter d'un corpus étudié sera ainsi affecté à une communauté selon les comptes avec lesquels il échange. **Sur des corpus de plusieurs dizaines de milliers de tweets, Gephi identifiera de 500 à 15000 communautés différentes, mais entre 50 et 75 % des échanges vont correspondre aux 10 premières communautés.** On peut ainsi avoir un premier niveau d'analyse si ce score moyen évolue dans le corpus étudié.

- **Eigenvector Centrality** : prendre en compte l'influence des comptes Twitter au sein de leur propre communauté. Ce calcul croise le nombre de liens (mentions / retweets) d'un compte Twitter avec la qualité de ses liens dans le réseau (le fait d'être relié à des entités ayant elles-mêmes beaucoup de liens).

Cela permet de mettre en évidence les comptes Twitter très mentionnés (notamment par d'autres comptes eux-mêmes très mentionnés) tout en diminuant l'importance des comptes très actifs (donc générant des mentions) mais peu repris par d'autres utilisateurs.

- **Betweenness Centrality** : identifier les acteurs au centre des échanges entre les différentes communautés d'un corpus de tweets.

Ce calcul permet de valoriser les acteurs au centre des échanges entre les communautés, sans être pour autant très mentionnés. **Il est important pour détecter des comptes Twitter peu repris mais influents en dehors de leur réseau, mais il peut générer des faux positifs** (un compte Twitter membre d'une petite communauté ayant deux ou trois mentions d'autres communautés aura un score très important). Il faut donc vérifier les résultats proposés par Gephi avec les données sources.

La combinaison de ces calculs statistiques à partir de corpus Twitter permet de mettre en évidence notamment :

- Des campagnes / stratégies d'influence, ponctuelles ou structurées sur le long terme ;
- L'emploi de robots / comptes inactifs pour pousser une thématique ;
- L'apparition de sous-communautés liées à un sujet donné et donc de signaux faibles concernant la montée en puissance d'un nouveau sujet / d'une thématique spécifique.

Exploiter des outils d'intelligence artificielle à partir de corpus Twitter enrichis d'analyses statistiques liées à l'analyse de graphe social permettra d'entraîner des algorithmes qui auront vocation à générer des alertes à destination d'experts : apparition de nouvelles thématiques, constitution d'une sous-communauté, identification de nouveaux influenceurs ... A noter que **l'outil Gephi, s'appuyant sur la dynamique des échanges au sein d'un corpus Twitter, n'exploite pas les métadonnées personnelles des comptes Twitter** : bio, abonnés, abonnements, localisation, date et heure des tweets ...

Il est nécessaire de disposer d'une expertise avancée de Gephi pour exploiter au mieux ses algorithmes, notamment en utilisant les filtres pour combiner certains éléments. Il est ainsi possible de définir des méthodologies pour anticiper des tendances à travers l'évolution des échanges sur des corpus de données précis, qu'il s'agisse d'identifier de nouveaux acteurs militants (Sylvestre, 2016) ou de vérifier l'influence et la qualité du réseau d'une startup (Sylvestre, 2020).

Pour autant, Gephi n'est pas exploitable pour mettre en place des alertes sur la détection de signaux : sans expertise humaine pour analyser le corpus de données, a posteriori d'un évènement, on ne pourra pas donner du sens aux algorithmes de Gephi. Aucune recherche n'a appliqué l'IA à Gephi pour atteindre cet objectif. Une start-up rachetée en février 2020 par Twitter, Fabulous AI, a développé une méthodologie pour identifier les « fake news » à partir de l'analyse des interactions sur les Twitter, mais sans faire appel à Gephi (Federico Monti et Al, 2019).

### 3. Le cadre théorique proposé par l'équipe du projet

Afin de répondre aux objectifs du projet, nous avons choisi d'analyser des échanges Twitter correspondant à différents mouvements sociaux, sur des sujets variés (immigration, droit des animaux, opposition aux grands projets, développement durable) et avec des ressorts de mobilisations spécifiques (manipulation de l'information, acteurs locaux, messages à forte charge émotionnelle ...).

Nous effectuerons une analyse avec Gephi en appliquant notre méthodologie à l'analyse des échanges au sein de ces corpus après qu'ils soient devenus viraux, pour identifier les principales communautés et leurs leaders. Une seconde analyse sera ensuite réalisée en arrêtant la collecte des données 24h avant que chaque sujet devienne viral, qui permettra ensuite de qualifier les comptes Twitter en fonction de leur importance pour la viralité du sujet. Le fichier qualifié obtenu sera exploité pour entraîner l'IA.

Les analyses des corpus Twitter combineront notre expertise en analyse des réseaux sociaux (1) et nos travaux de recherche sur les leaders d'opinions (2) appliqués aux réseaux sociaux afin de qualifier précisément les corpus de données qui seront analysés.

Après analyse et sélection des corpus pertinents (3), nous appliquerons le **social graph learning**, méthodologie mise en place pour ce projet (4).

#### 1. Adapter l'expertise en analyse des réseaux sociaux au contexte du projet

La littérature scientifique sur l'utilisation des réseaux sociaux pour la prédiction de tendances (Phillips, Dowling, Shaffer, Hodas, Volkova, 2017) indique l'importance de filtrer le bruit, de prendre en compte les biais propres aux réseaux sociaux, et d'adapter le modèle d'analyse à la spécificité des sujets traités :

- **Quelle que soit la thématique étudiée, l'essentiel des données Twitter correspond à du "bruit"**, c'est-à-dire des éléments peu ou pas intéressants par rapport à ce que l'on recherchera. Ce sera d'autant plus le cas sur des sujets innovants, où les effets d'annonces et les campagnes marketing sont démultipliés. **Il faut arriver à supprimer ce bruit ;**

- **Les algorithmes de Twitter, comme ceux des autres réseaux sociaux, amplifient la viralité de certains échanges dès lors qu'ils suscitent de nombreuses réactions.** Les contenus à fort potentiel de réaction émotionnelle ou proposés par des leaders dans leur secteur obtiendront donc énormément de visibilité, favorisant les experts en communication qui sont influents dans leur communauté, mais rarement en dehors ;

- **Aucun modèle d'analyse "clé en main" ne fonctionne pour détecter les signaux faibles sur Twitter** : chaque problématique et chaque thématique nécessiteront d'affiner une méthodologie rigoureuse pour obtenir des résultats pertinents.

Il faudra donc s'assurer, outre la qualité des données récoltées, de bien analyser au préalable le contexte des corpus de données sélectionnés afin de bien interpréter les résultats. **Nous formulons comme hypothèse qu'il y aura une variation de la précision de l'algorithme d'IA pour détecter les comptes Twitter influents selon les sujets étudiés.**

Les combinaisons via les filtres de Gephi de certains algorithmes permettent d'obtenir des éléments d'analyse plus fins pour l'identification des types d'influenceurs qui contribuent à la propagation de l'information.

Si on s'intéresse aux comptes Twitter capables de diffuser l'information auprès d'autres communautés, et donc de contribuer à la viralité d'un sujet, on recherchera en premier lieu ceux qui sont à la fois influents au sein de leur communauté et repris dans d'autres communautés.

Pour autant, les comptes peu influents dans leur propre communauté mais repris en dehors sont également essentiels à la diffusion de l'information. Ce sont ces deux types d'influenceurs que nous chercherons à identifier automatiquement via l'IA dans ce projet.

### Identification des types d'influenceurs sur Twitter via Gephi

	Relais auprès d'autres communautés	Relayé uniquement par sa communauté
Influence importante dans sa communauté	Compte Twitter très influent sur le sujet étudié (Information virale, transpartisan ...).	Compte Twitter influent pour ses abonnés mais qui ne dépasse pas son réseau (leader politique ...).
Pas influent dans sa communauté	Compte Twitter non influent dans son réseau mais capable d'être relayé ponctuellement (expertise, information virale ... ou faux positif de Gephi).	Utilisateur classique de Twitter, à surveiller s'il change de propriétés.

Source : Camille Baulant et Guillaume Sylvestre, 2020

Nous avons également repris et adapté la typologie des échanges sur les réseaux sociaux définie par le Pew Research Center (Smith, Rainie, Shneiderman, Himelboim, 2014). Cette typologie permet d'affiner la compréhension de la dynamique des échanges sur Twitter via Gephi.

## Typologie des communautés Twitter adaptées du Pew Research Center

TYPE DE RESEAUX DEFINIS PAR LE PEW RESEARCH CENTER	ORGANISATION DES COMMUNAUTES	NIVEAU D'ECHANGES ENTRE LES COMMUNAUTES	EXEMPLES
<b>FOULE POLARISEE</b>	2 communautés principales rassemblant environ 50 % des comptes Twitter	Echanges limités, car chaque communauté est très polarisée ; Présence de très nombreuses micro-communautés isolées	Controverse avec deux camps identifiés, les POUR et les CONTRE : opposition à un projet local ou national, soutien à une personnalité controversée
<b>FOULE ETROITE</b>	Une dizaine de communautés regroupent la quasi-totalité des comptes Twitter	Echanges nombreux entre les différentes communautés	Sujet peu évoqué, consensus entre différents acteurs : conférence / colloque scientifique ou professionnel, groupes militants peu influents ...
<b>GROUPE DE MARQUES</b>	Une dizaine de communautés regroupent entre 50 et 70 % des comptes Twitter	Peu d'échanges entre les communautés, nombreuses micro-communautés isolées	Discussions autour d'un sujet, d'une marque évoqué par des acteurs ayant des points de vue différents
<b>GROUPE DE COMMUNAUTES</b>	Une dizaine de communautés regroupent la quasi-totalité des comptes Twitter	Echanges entre les communautés via des influenceurs	Sujet très médiatisé qui suscite des échanges entre des acteurs qui se connaissent : Conférence sur le Climat, Consumer Electronic Show
<b>RESEAUX DE DIFFUSION</b>	Une communauté principale qui domine les échanges	Tout part de la communauté principale	Sujet / information à fort potentiel de réaction émotionnelle : images / vidéos choquantes, dérapage d'une personnalité, révélations politiques
<b>RESEAUX DE SOUTIEN</b>	Une communauté principale qui domine les échanges	Tout va vers la communauté principale	Soutien à une cause / personnalité sans réelle controverse (un seul camp domine les échanges)

Source : Camille Baulant et Guillaume Sylvestre, 2020

La collecte des tweets s'effectue en reprenant les mots-clés et hashtags pertinents pour chaque sujet validé, sur une période étendue afin de bien cibler l'analyse des communautés actives. Nous avons en parallèle vérifié que des erreurs d'orthographe n'avaient pas empêché la collecte de tweets importants pour l'analyse. En effet, la majorité des travaux de recherche sur Twitter ciblent des hashtags ou des termes définis dans un cadre universitaire, qui ne sont pas forcément adaptés aux pratiques des utilisateurs. Ce point est critique : certains échanges entre des acteurs clés ne sont pas identifiés, l'entraînement de l'IA et donc les résultats obtenus seront faussés.

## 2. Identifier les nouveaux leaders d'opinion des réseaux sociaux

Les échanges sur les réseaux sociaux en général, et Twitter en particulier, correspondent principalement à des interactions entre des individus. Ces interactions sont généralement dominées par des « **influenceurs** », **véritables leaders d'opinion qui seront repris par leur communauté, voir d'autres communautés**. C'est dans ce dernier cas qu'ils auront vraiment de l'influence, **car ils deviendront alors des « diffuseurs » d'information et permettront à un sujet de devenir viral**. **Nous nous intéresserons ici à deux principaux types de leaders d'opinion que l'on retrouvera sur Twitter** : les « **leaders économiques** » (Muller, Pénin, 2008 ; Baulant 2015) et les « **leader go-between** » (Amisse, Baulant, Muller, Vargas, 2008 ; Baulant, 2017).

Les « leaders économiques » se distinguent des « leaders hiérarchiques » de la société fordiste car ils peuvent se définir par deux grandes capacités inimitables : « voir loin » et « s'assurer que tous les membres du réseau auront une retombée réelle de leurs actions ». **Les leaders économique travaillent essentiellement sur des concepts assez généraux et mobilisent un réseau assez dense**. Une attention particulière est donnée aux « liens faibles », acteurs difficiles à identifier au préalable car souvent en retrait mais dont l'action a été mise en évidence dans les travaux de Granovetter (1973) et de Crozier (1977). L'objectif stratégique est de réussir à « amalgamer » plusieurs causes (par exemple environnementale, sociale, politique, économique, locale...) de façon à pouvoir avoir une résonance multiple et à tous les niveaux (local, national et mondial). **Ces « leaders économiques » correspondront sur les corpus étudiés à des acteurs publics / des entités très connues et visibles**, intervenant sur de multiples sujets et repris à minima par leur propre communauté. Faciles à identifier via leur nombre d'abonnés et de mentions Twitter, ils n'expliquent pas à eux seuls la viralité des échanges.

Le second type de leader travaille plus à court terme. **Ils sont souvent plus difficiles à identifier car leurs relations sont essentiellement personnelles, humaines, non linéaires** (Rallet, Torre, 2004 ; Kotter, 2007). Un mot « banal » pourra, grâce à une vidéo ou un contexte particulier, lancer avec succès une manifestation ponctuelle qui aura des conséquences pouvant dépasser largement l'évènement initial. En économie, on les appelle des « **leaders go-between** » (Amisse, Baulant, Muller, Vargas, 2008 ; Baulant, 2017) car ils ont des compétences interpersonnelles bien précises (Goleman, 2011) qui ont le pouvoir « d'unir les acteurs dans l'action », même si ces derniers peuvent diverger sur les stratégies de long terme. **Ils travaillent, même quand ils agissent sur les réseaux sociaux, essentiellement pour le niveau « local » de façon à mobiliser tout un tissu d'acteurs** (élus, citoyens, ONG, entreprises...) qui pourra avoir une influence non négligeable sur des niveaux supérieurs ou des intérêts financiers puissants. Ces leaders go-between sont cruciaux pour expliquer l'apparition de nouvelles thématiques, **du fait de leur capacité à mobiliser différentes communautés via leurs connexions**. Ils seront influents sur des sujets précis, et pas sur d'autres, à court et à moyen terme.

**Ce sont donc les leaders go-between que nous essaierons de détecter avec l'IA, en l'entraînant avec des corpus de données correspondant à différents contextes que nous allons décrire.**

### 3. Choix des corpus analysés

Trois thématiques différentes ont été retenues initialement pour ce projet : les mouvements sociaux, les actions de défense de l'environnement, et les mouvements d'extrême droite / gauche autour des Fake News.

Le choix des cas d'études a été effectué avec les étudiants du Master IESCI, afin d'éviter les biais de sélection de l'information, selon un cahier des charges précis :

- Le sujet devait correspondre à une des trois thématiques retenues ;
- Le sujet devait avoir provoqué au moins une manifestation ;
- La période était comprise entre juin 2018 et décembre 2019 ;
- Le sujet devait avoir été poussé par des acteurs français.

Les cas d'études soumis ont ensuite été analysés avec le concours des étudiants du master 2 IESCI de l'université d'Angers. Nous avons effectué une sélection des cas d'études pertinents pour notre méthodologie :

#### **Analyse des cas d'études proposés par les étudiants du M2 IESCI**

Sujet	Thématique	Période(s)	Validation
Pacte de Marrakech	Extrême droite / Immigration	Novembre / Décembre 2018	Cas d'étude retenu
Fake news sur la cession de l'Alsace Lorraine	Extrême droite / Souveraineté nationale	Janvier 2019	Non retenu : forte viralité mais pas d'impact en dehors des militants
Scandale sanitaire du Chlordécone	Santé / Environnement	Juin 2018 / février 2019	Non retenu : peu de viralité
Anti spécisme	Environnement / cause animale	Septembre 2018 / mai 2019	Cas d'étude retenu pour la période de septembre 2018
Fake News étudiant mort à Tolbiac	Mouvement social / violences policières	Du 20 au 24 avril 2018	Non retenu : réaction émotionnelle
Scandales alimentaires	Alimentation et santé	2018 - 2019	Cas d'étude non retenu, viralité trop étalée mais lien effectué avec l'antispécisme

Source : Master 2 IESCI, 2019

Il apparaît ainsi que certains sujets ayant suscité une couverture médiatique importante, voir des réactions localisées marquées, ont par la suite peu mobilisé aussi bien sur les réseaux sociaux que sur le terrain. Cela même si les mouvements en question n’ont pas obtenu satisfaction. Les analyses des étudiants ont permis par ailleurs de comparer les ressorts de la viralité de ces sujets :

### Comparaison des réactions selon l’origine du mouvement

<b>SCANDALES ET MOUVEMENTS</b>	<b>FAKE NEWS</b>
<ul style="list-style-type: none"> <li>- Antispécisme</li> <li>- Faux Steak</li> <li>- Chlordécone</li> </ul>	<ul style="list-style-type: none"> <li>- Pacte de Marrakech</li> <li>- Alsace Lorraine</li> <li>- Mort de Tolbiac</li> </ul>
<ul style="list-style-type: none"> <li>- <b>Effet de réseau</b></li> <li>- <b>Facilité d’accès à l’information accrue</b></li> <li>- <b>Action/Réaction</b></li> <li>- Création de groupes et rassemblement</li> <li>- <b>Perte de confiance</b></li> <li>- <b>Partis extrémistes</b></li> </ul>	<ul style="list-style-type: none"> <li>- <b>Partis extrémistes</b></li> <li>- <b>Désinformation</b></li> <li>- <b>Sentiment de peur</b></li> <li>- Gilets jaunes</li> <li>- <b>Perte de confiance Etat/Force de l’ordre</b></li> <li>- <b>Montée du nationalisme</b></li> </ul>

Source : Master 2 IESCI, 2020

Pour les scandales sanitaires et l’antispécisme, ce qui semble dominer sont les effets réseaux qui provoquent des actions et réactions en chaîne (grands rassemblements) lancées par les leaders d’opinion, ce qui crée une mobilisation citoyenne et ponctuellement des polémiques avec des camps opposés.

Pour les fake news, les leaders utilisent plutôt la désinformation de façon à favoriser la peur au sein de la population et la perte de confiance dans l’Etat accompagnée de la montée des nationalismes. Le Pacte de Marrakech ou le faux mort de Tolbiac sont devenus viraux parce que ces informations confirmaient des biais propres à l’extrême droite et à l’extrême gauche. Si le cas de Tolbiac a très vite été invalidé, pour le Pacte de Marrakech il est plus complexe de démontrer qu’il s’agit d’une Fake News puisqu’elle s’appuie sur l’interprétation d’un texte officiel.

En complément, trois sujets ont été ajoutés par l'équipe de recherche, et deux par l'équipe de suivi du CHEMI (en gras) :

Sujet	Thématique	Période(s)
Europacity	Environnement / opposition à un projet d'envergure nationale	Septembre – octobre 2019
Extinction Rebellion à Italie 2	Environnement	Août – octobre 2019
Extinction Rebellion et Lafarge	Environnement / opposition à un projet local	Janvier – mars 2020
<b>La ZAD Dune</b>	<b>Environnement / opposition à un projet local</b>	<b>Janvier – février 2020</b>
<b>#LaPurge</b>	<b>Extrémisme / défiance envers la police</b>	<b>Octobre 2018</b>

Chaque analyse correspond à une mobilisation effective par la suite, sauf pour #LaPurge. Certaines mobilisations se sont inscrites sur le temps long, comme pour Europacity, d'autres se sont construites très vite comme Extinction Rebellion.

Enfin, nous avons ajouté une analyse des échanges sur Notre-Dame-Des-Landes en janvier 2016, pour comparer ce corpus ancien avec les données plus récentes, et vérifier si la dynamique des échanges Twitter a évolué fortement depuis.

Il faudra bien entendu enrichir après l'étude les corpus d'études, mais il apparaît clairement qu'il est impossible de prédire le succès d'une mobilisation / d'un mouvement avec les seules caractéristiques des communautés qui en sont à l'origine.

#### 4. Le Social Graph Learning, application de l'IA à la Social Network Analysis

Nous utilisons le terme de « **Social Graph Learning** » pour la méthodologie que nous proposons visant à appliquer l'IA aux données issues de Gephi, qualifiées par une expertise humaine. **C'est donc la qualité de la préparation de nos données, à partir de l'analyse via Gephi des corpus Twitter**, qui permettra d'exploiter au mieux l'IA dans ce projet.

Le logiciel Gephi sera utilisé sur chaque corpus pour **analyser la propagation de l'information entre les communautés et les comptes influents** sur la période de viralité.

Une seconde analyse sera effectuée en **exploitant le corpus des premiers échanges sur le sujet** (au maximum un mois en arrière) **jusqu'à la veille de l'accélération des échanges**. Toutes les statistiques de Gephi seront exploitées, les données exportées et qualifiées comme suit :

- Identification de tous les comptes ayant participé à la diffusion entre les communautés et élimination des faux positifs par vérification du corpus ;
- Mise à l'écart des comptes les plus influents (personnalités politiques / médiatiques) pour entraîner l'algorithme sur les acteurs ponctuels de la diffusion de l'information ;
- Suppression des données personnelles du corpus, ajout pour chaque cible d'informations sur la typologie du compte et sa capacité de viralité.

Ces critères correspondent à notre expertise sur l'analyse des réseaux sociaux. **Ils permettront pour chaque corpus comprenant entre 1000 et 5000 comptes d'identifier entre 35 et 85 cibles environ.**

L'identification de ces comptes considérés comme influents et moteurs de la visibilité de la thématique étudiée sera la cible de l'outil IA, qui entraînera ses algorithmes pour retrouver des comptes ayant les mêmes caractéristiques via les calculs de Gephi.

Les exports qualifiés seront anonymisés en retirant toutes les données personnelles des comptes Twitter : Nom, bio, localisation ... Seul le nombre d'abonnés et d'abonnements sera encore présent, sachant qu'ils ne permettent pas d'identifier précisément un utilisateur puisque ces valeurs sont dynamiques.

Nous avons choisi pour appliquer l'IA aux données qualifiées l'outil DataRobot, qui permet de tester plusieurs dizaines de modèles de machine learning. Nous aurons un accès gratuit à cet outil, avec l'accompagnement d'un data scientist expert pour interpréter les résultats. DataRobot permet notamment de :

- Identifier les éléments des corpus de données qui expliquent le mieux les résultats obtenus, pour vérifier l'existence de biais statistiques ;
- Comparer deux fichiers cibles pour évaluer l'efficacité des algorithmes de machine learning selon le sujet analysé ;
- Vérifier si certains calculs effectués par Gephi provoquent des erreurs (afin de les retirer du corpus).

### III. Enseignement des analyses des corpus de données

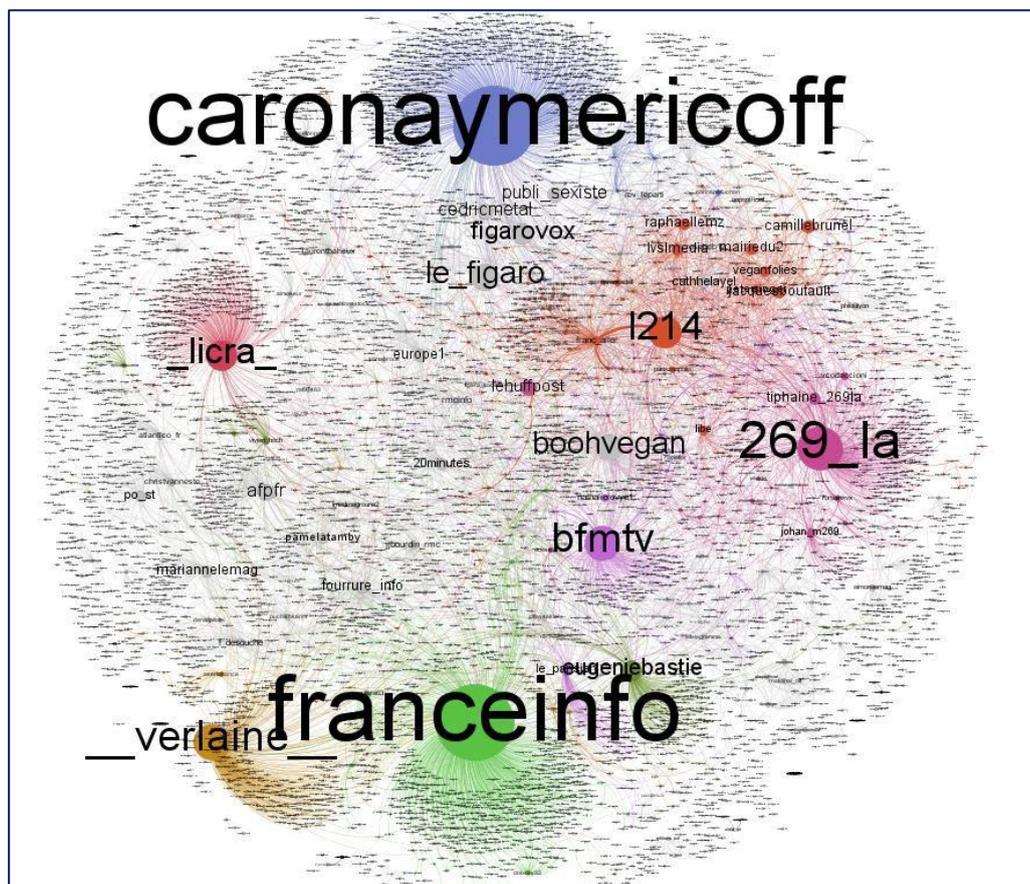
Nous avons effectué une analyse détaillée de chaque corpus de données, qui sont intégrées en annexe. Nous pouvons en tirer plusieurs enseignements sur la typologie des échanges.

#### 1. Les sujets de controverse publique : Anti-spécisme, LaPurge

Ces thématiques de mobilisation renvoient à une controverse entre deux camps identifiés, dont les idées sont irréconciliables. Pour #LaPurge, on aura d'un côté des personnes légitimant des violences contre les forces de l'ordre, et de l'autre les défenseurs de l'ordre républicain. Le sujet a obtenu un très faible écho, avec 3458 mentions Twitter du 28 octobre 2018 au 12 novembre 2015, dont 88 % de retweets, par 2679 comptes. Les mentions en soutien aux forces de l'ordre sont très majoritaires, et leurs contradicteurs peu nombreux. L'analyse de ce corpus servira donc principalement à vérifier la capacité de l'algorithme à retrouver des cibles sur un corpus restreint.

Le cas des manifestations antisépécistes est très révélateur des stratégies d'influence actuelles, qui opposent les militants, ciblant la société civile, et les acteurs industriels, ciblant les politiques, les deux cherchant également à attirer l'attention des médias. Nous avons étudié les échanges avant et après une action organisée contre des bouchers le 22 septembre 2018. Du 31 août au 1<sup>er</sup> octobre 2018, 8307 tweets ont été publiés par 5159 utilisateurs, la mobilisation ayant eu lieu le 22 septembre.

#### Cartographie via Gephi des comptes les plus mentionnés par leurs communautés sur l'antisépécisme du 31 août au 1er octobre 2018





Les sujets de controverse publique qui entraînent des mobilisations tendront à avoir les caractéristiques suivantes sur Twitter :

- Présence de contradicteurs peu influents mais très actifs, qui seront soutenus après une action par des influenceurs liés au secteur attaqué ;

- Communautés militantes structurées autour d'un leader go-between influent pour mobiliser en dehors de ses communautés, qui peuvent avoir des avis divergents (ici visibilité médiatique et autour des publics vegans) ;

- Faible visibilité des acteurs politiques : le soutien du Ministre de l'Agriculture aux bouchers après l'action d'un « commando vegan » contre une boucherie le 5 mai a ainsi été peu repris.

Les tensions liées aux critiques du secteur agricole, qui sont multiples et très relayées notamment dans les médias (violences animales, scandales alimentaires, risques liés aux pesticides), **ont ainsi amené les agriculteurs et leurs soutiens à utiliser fortement Twitter en 2019**, notamment via le terme #Agribashing, promu initialement par la FNSEA.

Une même évolution pourrait voir le jour sur les questions de santé, traditionnellement peu mobilisatrices en dehors des associations de patients (Dépakine, Mediator, Chlordécone) et du soutien de certains politiques, suite au Covid-19. **L'opposition entre les soutiens de Didier Raoult et les défenseurs des protocoles médicaux, regroupés sous le hashtags #NoFakeScience** semble le préfigurer.

## 2. Les sujets de militantisme sociétal : Pacte de Marrakech, Extinction Rebellion à Italie 2

Les grands sujets de crispation sociétale, qui dépassent le cadre militant, sont ceux susceptibles de mobiliser très vite, à travers notamment les réseaux sociaux et Twitter, mais pas uniquement. On retrouvera deux cas de figure :

- L'utilisation de Twitter pour « alerter » sur un sujet pas ou peu traité par les médias, contre une « manipulation » politique, via une information fautive ou détournée. **Ce sont principalement les acteurs d'extrême gauche ou d'extrême droite qui en seront à l'origine, leurs militants reprenant automatiquement les informations partagées ;**

- La construction d'une action militante en dehors de Twitter ou de tout réseau social accessible publiquement, **mais qui en parallèle générera des conversations non pas sur l'action mais sur le réseau militant en question.** Ici ce sont surtout pour les thématiques environnementales que ce mode d'action se retrouvera.

Extinction Rebellion, mouvement né au Royaume-Uni en automne 2018 et qui est présent dans 55 pays, prône la désobéissance civile et des actions coup de poing pour obtenir des actions concrètes contre le changement climatique, tout en restant non-violent. Extinction Rebellion France a été lancé le 24 mars 2019 et recrute très vite près de 10 000 militants prêts à s'engager pour ses actions. Ceux-ci n'ont souvent jamais milité, et se retrouvent dans une structure très horizontale, sans réelle hiérarchie. A partir de 3 militants, des actions peuvent être organisées. Cela en fait un mouvement capable de monter très vite en puissance et dont il est difficile d'anticiper la portée.

Ainsi, l'action « Notre cœur brûle » sur les incendies de forêt en Amazonie, organisée à Paris le 31 août, a réuni 250 militants qui se sont coordonnés en moins d'une semaine. La militante qui a proposé l'action avait rejoint le mouvement un mois plus tôt. Le 7 octobre 2019, les groupes XR de plusieurs pays ont lancé des actions de blocage dans le cadre de la « Rebellion Internationale d'Octobre ». A Paris, cette action a commencé le 5 octobre avec l'occupation du centre commercial Italie 2 par 2000 militants. La préparation de cet événement a eu lieu sur plusieurs mois.

Sur un total de 22 353 tweets publiés entre le 01 septembre et le 6 octobre 2019, près de 18000 correspondent à l'occupation du centre commercial Italie 2 initiée le 5 octobre. Il y a 87 % de retweets par 6975 utilisateurs. Les médias ont tous largement repris cette manifestation comme le montre la cartographie des comptes Twitter les plus influents au sein de leur communauté :





**échanges sur Extinction Rebellion se sont accélérés avant l'action d'Italie 2.** Les interactions entre ces deux types de leader, largement complémentaires, ont permis de structurer des groupes d'horizons variés et de créer une triple proximité « organisée », « thématique » et « géographique » (Massé, Thibaut, 2000 ; Torre, Rallet, 2004 ; Baulant , 2017) capable de dépasser les intérêts de chaque groupe mobilisé.

Les caractéristiques de ces sujets sociétaux seront :

- L'action d'un ou plusieurs leaders économiques identifiés (même ponctuellement) par les militants (personne ou organisation) qui seront initiateurs d'une action ou propagateur auprès des différentes communautés ;

- La participation de plusieurs leaders go-between au relais de l'information pour toucher des communautés multiples dans leur réseau ;

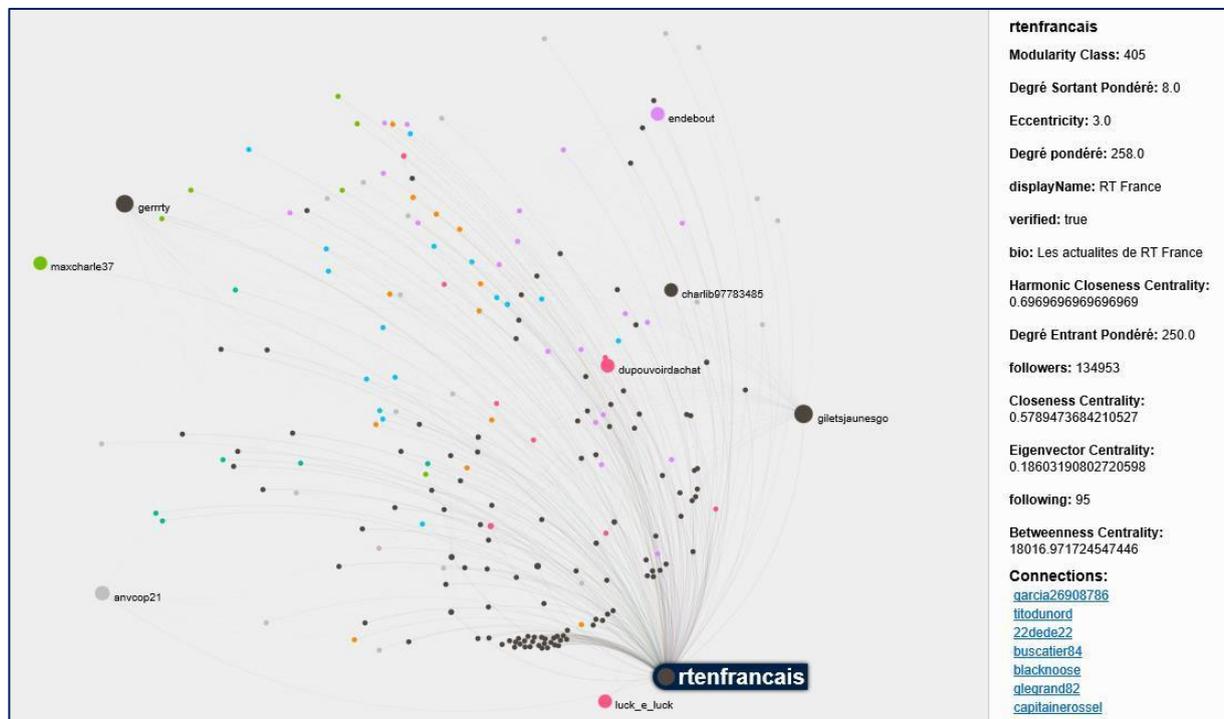
- L'utilisation ponctuelle de fausses informations qui correspondent aux biais de confirmation des communautés ciblées qui fait que les acteurs n'ont, a posteriori, aucune difficulté à trouver des explications raisonnables à un évènement hautement improbable (Klein, 2011).

Dans ce cadre, on peut penser que l'effet des mobilisations autour du COVID-19 sera extrêmement fort car la parole publique est aujourd'hui totalement décrédibilisée. C'est donc un boulevard pour la contestation en général et celle impliquant des fausses informations en particulier. La parole publique ne pourra se rétablir que si des efforts sincères des dirigeants politiques sont faits en ce sens, en s'appuyant sur des connaissances et des expériences historiques aussi bien que récentes. A long terme, plus qu'une stratégie de communication de crise, **le gouvernement a besoin de vrais « leaders économiques » capables de combiner histoire, culture, expérience et ouverture pour proposer une vision nouvelle du problème et des solutions originales et adaptées** (Duggan, 2012, Baulant, Blondel, 2014). Il est essentiel en complément de s'appuyer sur les leaders go-between reconnus par les différentes communautés pour s'assurer de la propagation des messages et de leur forte diffusion.

Une accélération de ces actions de défiance envers le pouvoir public est à prévoir, d'autant plus qu'outre les médias militants, **elles seront très relayées par des agents d'influence tels que Russia Today**. Ce média, qui privilégie la couverture des mouvements sociaux, **avec un focus sur les violences policières, est devenu de fait une source d'informations fiable pour les personnes ne soutenant pas le gouvernement**. Les gilets jaunes notamment se sont beaucoup informés via Russia Today et ses vidéos des manifestations, louant notamment l'absence de commentaires éditoriaux comme sur les chaînes d'infos en continu françaises.

Outre le Pacte de Marrakech, Russia Today s'est intéressé tôt à Extinction Rebellion, comme on le voit sur cette cartographie des mentions du compte Twitter du média.

### Cartographie dynamique via Sigma - Gephi des échanges autour du compte @RTenfrançais sur Extinction Rebellion du 1er septembre au 6 octobre 2019



Plus récemment, le média Twitter Conflits, créé en janvier 2020, et qui a connu une forte progression de son audience pendant la crise Covid, commence à diffuser des informations non vérifiées, avec un fort impact. Il est intéressant de noter que Conflits a qualifié le 11 janvier 2020 Extinction Rebellion au Royaume-Uni de « groupuscule climatique », et après le confinement a couvert plus positivement des actions militantes effectuées en France, notamment le blocage des postes d'Orly le 26 juin 2020.

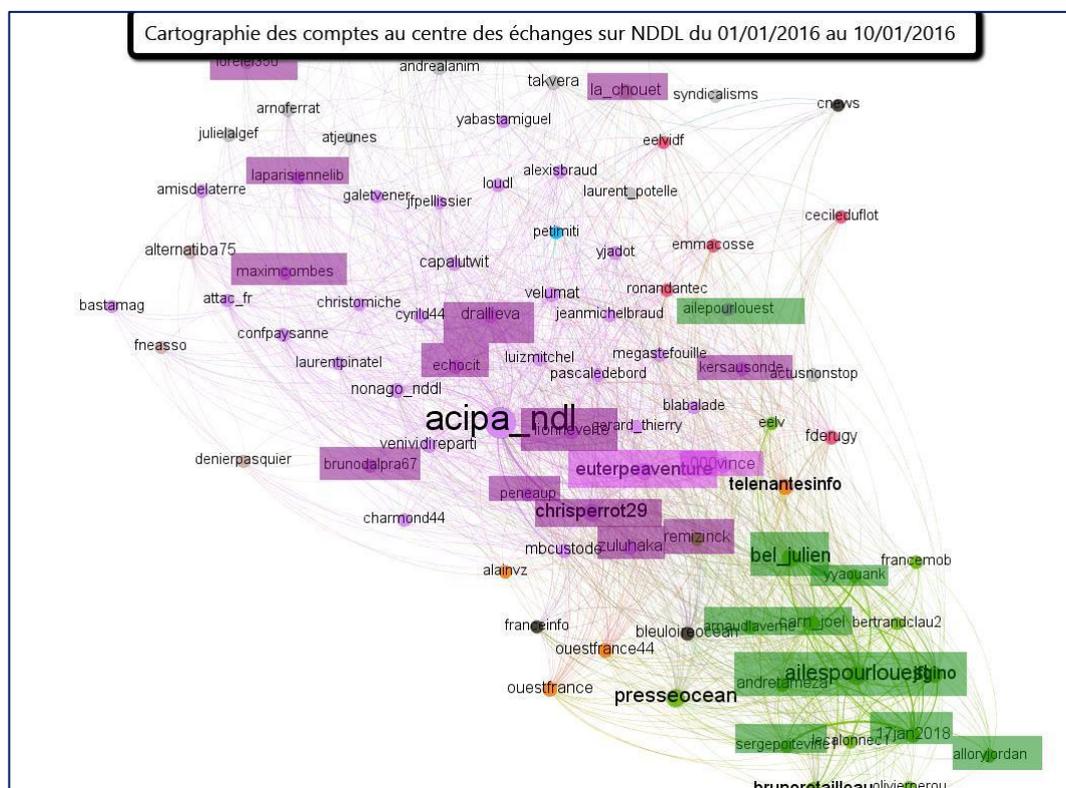
### 3. Les sujets portés par des acteurs locaux : Europacity, Extinction Rebellion Lafarge, ZAD Dune

Le premier sujet porté par des acteurs locaux et ensuite très relayé sur Twitter est l'opposition à la construction de l'aéroport de Notre-Dame-Des-Landes. Le mouvement était très suivi par la presse locale et nationale, notamment avec l'essor de nouvelles « ZAD », **mais a suscité peu d'échanges sur Twitter avant l'annonce en octobre 2015 de la reprise des travaux**, suspendus entre 2012 et 2015. Sur les mois précédents de l'année 2015, il y a en moyenne 6000 tweets sur le sujet, avec quelques pics lors des nouvelles importantes (18 000 tweets en juillet 2015 suite au rejet des recours des opposants, 14 000 tweets en décembre 2014 avec plusieurs articles sur les « zadistes ») et un énorme pic de 70 000 tweets à l'occasion de la manifestation du 22 février 2014 contre le projet à Nantes.

Ce qui va changer en 2016, c'est une relance des manifestations en prévision de la reprise des travaux, qui va accélérer la dynamique Twitter jusqu'au référendum en juin 2016, avec une moyenne de 70 000 tweets par mois sur le premier semestre 2016, dont 130 000 entre la première manifestation de l'année le 9 janvier et l'annonce du référendum le 12 février.

**La cartographie des comptes au centre des échanges sur le début du mois de janvier 2016 montre une opposition entre deux camps équilibrés, dont les principaux influenceurs sont des figures nationales :**

#### **Cartographie via Gephi des comptes au centre des échanges sur Notre Dame des Landes entre le 1er et le 9 janvier 2016 – application d'un filtre pour garder les comptes ayant un score élevé**



Les opposants à l'aéroport sont dans la communauté en violet, les soutiens dans la communauté verte. On notera que les leaders d'EELV sont à l'écart des deux communautés, préfigurant l'absence totale d'influence de ce parti politique sur les mouvements sociaux récents tel qu'Extinction Rebellion.





Une des grandes caractéristiques du leader go-between est d'être capable de relier deux communautés isolées l'une de l'autre sans proximité thématique (Karaszewski, Lis, 2013 ; Baulant, 2017). Ainsi, **par l'intermédiaire du leader go-between, les messages des acteurs locaux** (élus, citoyens activistes écologiques) **peuvent avoir une retombée nationale ou, à l'inverse, un leader national pourra influencer le débat local**. Dans les deux cas, le leader go-between se trouve dans une position stratégique puisqu'il peut soit diffuser, soit bloquer l'information de l'une des communautés.

L'importance de ces leaders go-between locaux, moins étudiés car moins médiatisés que les leaders économiques, s'est retrouvée dans la mobilisation d'Extinction Rebellion contre Lafarge, qui a impliqué des militants et élus locaux très actifs, et dans la ZAD Dune. **C'est l'élément principal expliquant la capacité de mobilisation pour ces deux actions**. Leur succès dans la durée dépendra de la capacité des leaders go-between de développer leur influence en fédérant des communautés autour d'eux et en s'appuyant sur leur connaissance fine de terrain des acteurs locaux.

## IV. Résultats de l'application de l'IA

### 1. Une méthodologie rigoureuse pour s'assurer de la qualité des données avant l'analyse

Nous avons exploité l'IA dans le cadre d'un apprentissage supervisé, c'est-à-dire **en indiquant aux modèles de machine learning de l'outil DataRobot les cibles que nous cherchons à atteindre**. Les « cibles » sont les comptes correspondant aux leaders go-between identifiés sur chaque corpus analysé, avant que la diffusion des tweets ne devienne virale.

Les analyses des corpus ont permis de définir précisément les cibles de l'IA. Celles-ci ont été effectuées par les étudiants et par l'équipe de recherche afin de s'assurer qu'il n'y avait pas d'erreur dans l'interprétation des données. Pour chaque corpus, nous avons identifié le moment où le sujet devient viral, pour exporter les échanges antérieurs à cette période spécifique. **Celle-ci doit être adaptée à chaque corpus de données, elle peut correspondre à quelques jours ou plusieurs semaines. Ces analyses croisées ont également permis d'identifier et de corriger des bugs d'exports des corpus Twitter depuis Gephi, qui provoquaient la disparition de certains comptes Twitter.**

Toutes les données personnelles des comptes Twitter ont été supprimées pour l'analyse, chaque compte ayant obtenu un identifiant aléatoire. Cela nous permettra d'exploiter l'IA en respectant le RGPD, puisque seules les variables liées à la dynamique des échanges sur un sujet sont pertinentes.

Les choix suivants ont été effectués pour limiter les risques de biais liés aux corpus de données et aux caractéristiques d'un sujet :

- Les comptes Twitter de personnalités / organisations d'envergure nationale et donc influents sur de multiples sujets, **correspondant à des leaders économiques (« voir loin » et « s'assurer d'un gain réel pour tout son réseau ») ou à des leaders hiérarchiques (pouvoir politique ou financier)**, ne sont pas sélectionnés comme cible. Ils restent néanmoins indiqués dans le corpus pour vérifier leur pertinence éventuelle, sachant que la dynamique de propagation sur Twitter ne s'explique pas par la seule intervention de ces acteurs. **Cela permet de s'assurer que les algorithmes de machine learning chercheront à identifier les propriétés des leaders go-between** avant tout ;

- La qualification des comptes Twitter préalable à un évènement se fera sur une période d'au moins une semaine avant l'évènement, jusqu'à un mois avant celui-ci selon la dynamique des échanges. Le corpus s'arrêtera la veille de l'évènement, pour augmenter la possibilité d'identifier des caractéristiques pertinentes via l'IA. Cela nous permet de disposer de conditions idéales pour vérifier la capacité de l'IA à retrouver des leaders go-between sur les différents corpus ;

- La dynamique de propagation des tweets propre à chaque compte identifié est vérifiée en utilisant un export dynamique de la cartographie Gephi, permettant de visualiser le réseau propre à chaque compte Twitter pour vérifier s'il s'agit bien de leaders go-between, capables de relayer l'information à plusieurs communautés.

Pour l'ensemble des corpus de données, nous avons obtenu entre 50 et 100 cibles pour l'IA, sur un total de 1000 à 4000 valeurs. C'est donc un ensemble de données très limité qui est utilisé, afin de définir des seuils d'alertes les plus sensibles possibles.

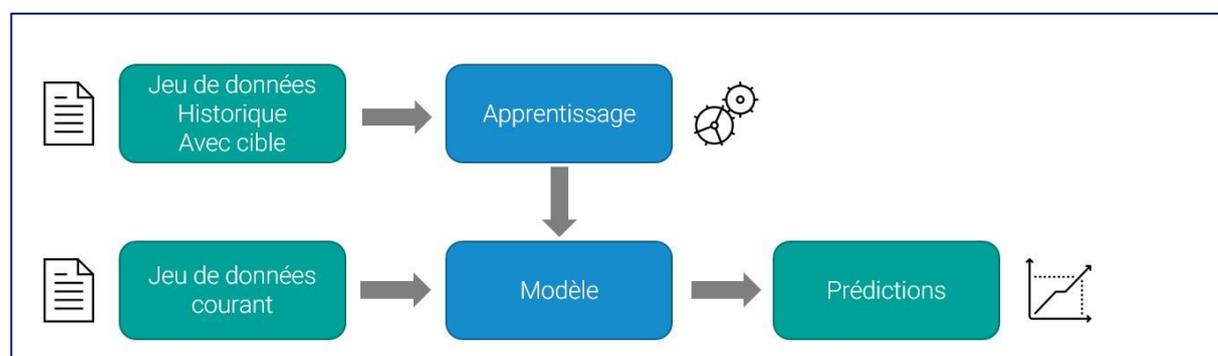
Ce travail de préparation et de qualification des données, après analyse humaine et avant analyse par l'outil d'IA correspond à la moitié du temps passé sur le projet. C'est une étape indispensable au succès d'un projet d'IA (Sadin, 2018), qui ne peut s'appuyer sur la seule analyse de données non préparées pour un objectif précis.

## 2. Procédure d'analyse des données avec l'outil DataRobot

7 corpus de données Twitter ont été exploitées pour créer un modèle de prédiction nous permettant d'identifier les leaders go-between dans un réseau afin d'anticiper la viralité d'une information.

Le modèle d'analyse a été construit avec un data scientist de la société Mydral, expert du logiciel DataRobot, qui permet d'utiliser des dizaines de modèles de machine learning issus de bibliothèques open source ou propriétaires.

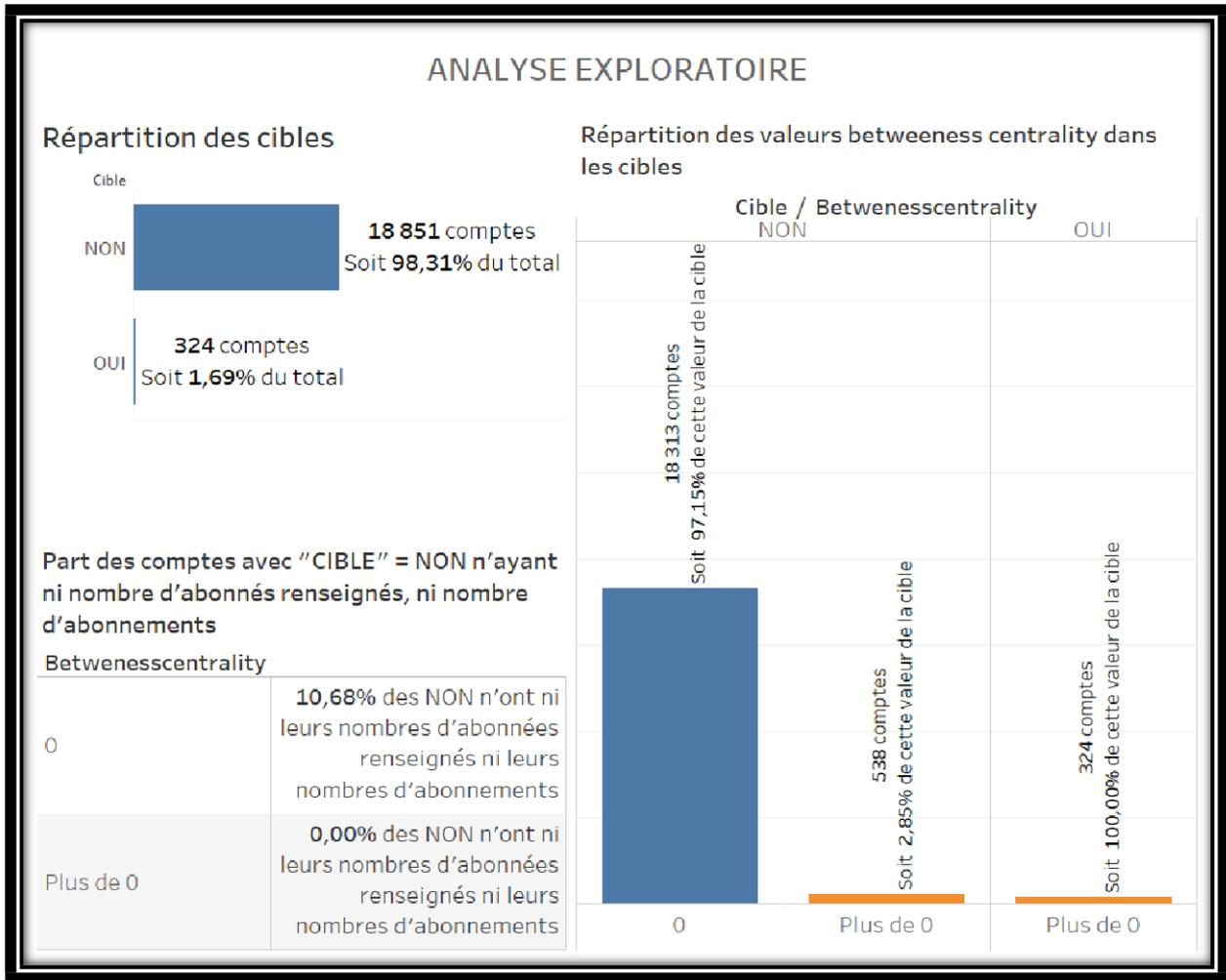
### **Fonctionnement des modèles de machine learning exploités par DataRobot**



*Un modèle de Machine Learning comme combinaison de pré-traitement des données et d'un algorithme reçoit un ensemble de données d'apprentissage indiquant la cible à prédire. Sa précision sera validée sur un autre ensemble en comparant la prédiction aux données réelles. Une fois validé, ce modèle est utilisé pour prédire la cible sur de nouvelles données (source : Mydral & DataRobot).*

6 étapes de traitements et d'itérations ont été réalisées dans le cadre de la constitution de notre modèle de prédiction. Ces étapes comprennent une phase d'analyse descriptive des données Twitter permettant d'orienter le traitement de ces données avant l'application de modèles de prédiction.

**Analyse préparatoire pour déterminer les caractéristiques des données via Tableau Desktop**

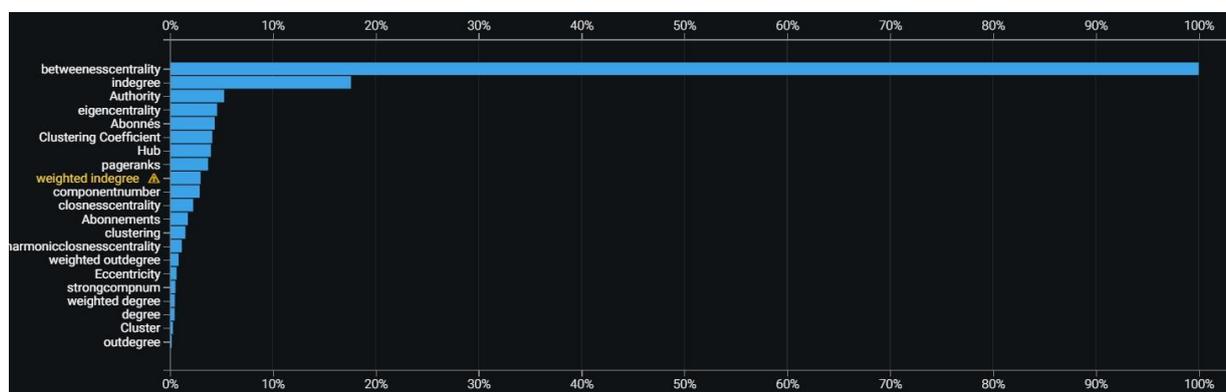


### 3. Mise en œuvre des modèles de machine learning

Dans un premier temps, 4 sets de données ont été utilisés pour entraîner les modèles, et 2 pour vérifier la qualité de prédiction de l'IA. Le nombre de cibles négatives a été sous échantillonné pour éviter que certains modèles n'arrivent pas à faire la différence entre les cibles positives et négatives. Dans ces conditions, 82,5 % de cibles positives ont été détectées.

DataRobot permet d'identifier les variables qui impactent le plus les résultats obtenus par l'IA. Ici, c'est le score de betweenness centrality, qui a un poids prépondérant sur les cibles négatives, tout en ne permettant pas de définir les cibles positives :

#### Identification par DataRobot des variables impactant le plus l'IA pour déterminer les cibles après la première analyse



Les étapes suivantes ont donc consisté à jouer sur les valeurs de betweenness centrality et d'autres variables identifiées comme problématiques, afin d'améliorer le modèle de prédiction, pour approcher les 100 % de détection des cibles positives. Le taux de faux négatif était par contre élevé, proche de 50 %.

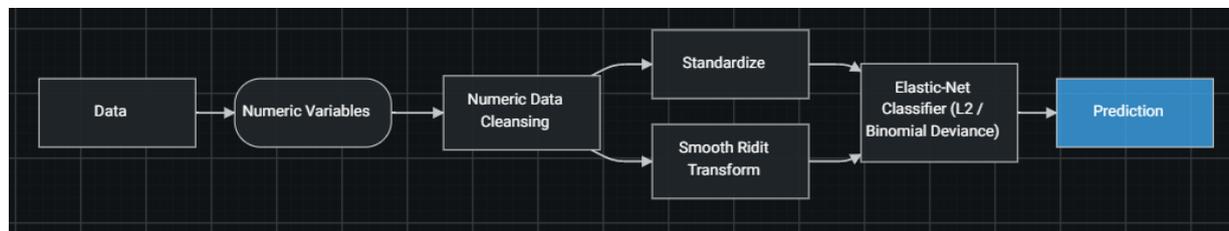
A noter qu'il n'y a qu'une faible différence dans les résultats selon les sets de données sélectionnés pour l'entraînement ou la validation : **quel que soit le sujet / la thématique abordée, la détection des leaders go-between s'effectue de la même façon pour l'IA, dès lors que la méthodologie de préparation des sets de données est la même.** Certaines variables problématiques ont été retirées pour améliorer les résultats. **Il apparaît donc qu'un facteur essentiel pour un projet d'IA n'est pas d'utiliser le plus grand nombre de variables possibles, mais plutôt de sélectionner celles qui seront les plus pertinentes pour le projet.** On parle de problème de sur-apprentissage des modèles qui modélisent du bruit et en conséquence ne sont pas aussi performant sur les jeux de données courants. C'est l'intérêt de disposer d'un jeu de données dit « holdout » pour le vérifier.

DataRobot met visuellement en avant les variables les plus importantes pour chacun des modèles. On peut réduire le nombre de variables exploitées par l'IA et conserver uniquement les plus importantes pour améliorer la performance du modèle. Par ailleurs DataRobot met ce qu'on appelle la dépendance partielle à disposition de l'utilisateur pour chacun des modèles. On peut ainsi comprendre comment le modèle se comporte selon les valeurs des variables les plus importantes. C'est grâce à l'interprétabilité des modèles et la dépendance partielle que Data Scientistes et experts métiers peuvent valider ensemble le bon fonctionnement d'un modèle pour établir les prédictions.

En appliquant un sous-échantillonnage des valeurs cibles négatives, sur-représentées sur les sets de données tests, **nous avons mis au point un modèle permettant d'identifier 100% des vrais leaders go-between avec un taux de faux positifs entre 22 % et 27 %.**

L'algorithme qui nous a permis d'obtenir les meilleurs résultats correspond à un modèle à directions de descente, destiné à minimiser une fonction réelle différentiable définie sur un espace euclidien ou, plus généralement, sur un espace hilbertien. L'algorithme est itératif et procède donc par améliorations successives. Au point courant, un déplacement est effectué le long d'une direction de descente, de manière à faire décroître la fonction. Le déplacement le long de cette direction est déterminé par la technique numérique connue sous le nom de recherche linéaire. Toujours dans le but de réduire le sur-apprentissage, l'algorithme utilisé intègre des pénalités, on parle de régularisation L1 et L2 (L2: Ridge/ L1: Lasso). Ce modèle correspond à l'**Elastic-Net Classifier** (Blondel, Mathieu, Seki, Uehara, 2013)

### Schéma expliquant le fonctionnement de l'Elastic-Net Classifier – via DataRobot



Nous avons formalisé une **relation inverse entre la probabilité moyenne d'être un leader go-between et le nombre d'abonnés du compte (ce qui confirme nettement les travaux de Granovetter, 1973 et de Crozier, 1977)** ainsi qu'entre la concentration de la communauté à laquelle il appartient et sa probabilité moyenne d'être un leader go-between.

#### 4. Nos pistes de travail pour perfectionner le modèle IA

L'échantillon d'entraînement est limité. S'il a permis de valider les hypothèses et d'affiner la méthodologie, il est nécessaire de l'enrichir pour améliorer le modèle, en doublant à minima les sets de données. Ceci permettra de soumettre le modèle à différents échantillons d'entraînements et de validation via la méthode de « validation croisée », tout en s'assurant d'avoir des sous échantillons de corpus différents dans l'entraînement et la validation de manière à éviter le surapprentissage et donc avoir un modèle extensible à de nouvelles prédictions. De plus, **on pourra également diminuer les faux positifs en disposant d'une plus grande variété de cibles.**

Pour l'exploitation des résultats de l'IA, l'extension des corpus tests aura un autre intérêt : cela permettra de définir les sets de données correspondant à ce que l'on souhaite détecter, afin de définir un seuil de leaders go-between (cibles) par rapport à l'ensemble d'un corpus qui doit déclencher une alerte. En effet, puisque l'on arrive à retrouver ces leaders go-between quel que soit le corpus étudié, il faut définir une valeur de déclenchement d'alerte qui sera intégrée à un système de veille.

Pour cela DataRobot met à disposition la matrice de confusion qui permet de relever les faux négatifs et positifs sur les jeux de données de validation et de holdout. On peut visuellement définir dans la plateforme ce seuil d'alerte qui permet de discriminer les cibles positives et négatives selon les résultats escomptés.

Un algorithme de machine-learning non supervisé (qui ne s'intéresse donc à aucune cible à prédire) de Datarobot permet également d'identifier les « anomalies » dans les résultats obtenus, c'est-à-dire les cibles considérées comme positives mais dont les variables sont très éloignées de la moyenne des autres cibles. Ceux-ci seront très probablement des faux positifs. En vérifiant cette hypothèse il sera donc possible d'affiner les résultats.

## Conclusion

Ce projet de recherche a été financé dans le cadre d'un appel à projets du Centre des Hautes Études du Ministère de l'Intérieur et réalisé en lien avec le Centre de Recherche de l'Ecole Nationale Supérieure de Police. L'objectif était d'utiliser l'IA pour aider à la compréhension des échanges sur les réseaux sociaux. Comme nous l'avons rappelé au début de l'article, les solutions existantes ne permettent au mieux que de détecter des signaux déjà caractérisés (via des mots-clés, des personnes connues), tout en étant sujet à des biais de manipulation, un fort volume de mentions pouvant correspondre à l'utilisation de robots, à une opération d'influence ou de communication ...

Nous avons démontré que l'analyse de la dynamique des échanges Twitter, et non du contenu de ces échanges ou des caractéristiques personnelles des comptes Twitter, **est une piste sérieuse pour arriver à anticiper la viralité d'un nouveau sujet sur Twitter**. Sur des corpus de données variés, correspondant à différents acteurs et contextes, à portée nationale ou locale, le profil des leaders go-between, ceux qui transmettent l'information d'une communauté à une autre, ce qui permet à un sujet de devenir viral, varie peu. Ces leaders sont identifiables via une méthodologie spécifique.

Leur détection via les seuls calculs statistiques de Gephi n'est pas possible, puisqu'il faut une vérification humaine des résultats. Mais en fournissant des corpus de données qualifiés à une IA d'apprentissage supervisée, nous pouvons les retrouver à partir d'une vingtaine de calculs proposés par Gephi. **Nous avons constaté que tous n'étaient pas pertinents, et qu'il fallait ne fournir que des variables dont l'interprétation n'amènerait pas de biais.**

Nous allons poursuivre notre travail pour mettre un point un système d'alerte sur les sujets devenant viraux sur Twitter à partir de cette démarche de Social Graph Learning. L'enjeu est double : **éviter la surcharge d'alertes non pertinentes aux systèmes de veille actuels**, qui au final ne sont pas pris en compte faute de temps suffisant pour les traiter ; et **permettre d'interpréter correctement les signaux remontés avec une grille de lecture caractérisant les typologies de communautés** sur les réseaux sociaux.

Enfin, nous souhaitons rappeler qu'un système de veille ne peut se limiter aux données Twitter. Ce réseau social sera important pour identifier de nouvelles tendances, de nouveaux mouvements, des opérations d'influence ... mais il faudra également croiser les informations remontées avec des recherches sur d'autres réseaux sociaux et sur le terrain, auprès de sources humaines.

## V. Bibliographie

- Alloing C., Pierre J. (2017). « Le Web Affectif : Une économie numérique des émotions », INA.
- Amisse S., Baulant C., Muller P., Vargas A. (2008). « The Concept of Cluster and its Two Logics ». EAEPE Conference Labour, Institution in a Global Knowledge Economy, Rome.
- Antolinos-Basso D., Flaminia Paddeu N., Douay N., Blanc N. (2017). « Pourquoi le débat #EuropaCity n'a pas pris sur Twitter ? : Analyse de la mobilisation autour d'une controverse environnementale sur le réseau social. » *Reset* no. 7:1-1.
- Baulant C. (2015). « The Role of Networks for Helping Firms and Countries Invent New Competitive Strategies Well Adapted to the World Knowledge Economy », *Journal of Economics Issues*, 49-2, p 563-573.
- Baulant C. (2017). « How Happiness can lead to more Efficiency: A New Paradigm Adapted to the World Knowledge Economy », *The American Review of Political Economy*, pp110-125.
- Baulant C. (2017). « Rethinking the links between Human Relationships and Economic Efficiency using the Local Micro institutions: the case of two emerging countries », *Journal of Economics Issues*, 51-3, pp 651-662.
- Beauvisage T., Beuscart J-S. Couronné T., Mellet K. (2011). « Le succès sur Internet repose-t-il sur la contagion ? Une analyse des recherches sur la viralité », *Tracés. Revue de Sciences Humaines*, pp 151-166.
- Beauvisage T., Beuscart J-S. Couronné T., Mellet K., « Le succès sur Internet repose-t-il sur la contagion ? Une analyse des recherches sur la viralité », *Tracés. Revue de Sciences humaines*, pp 151-166, 2011
- Berney E. (1928). *Propaganda : Comment manipuler l'opinion en démocratie* (trad. Oristelle Bonis, préf. Normand Baillargeon), première édition 1928, Paris, La Découverte, 2007.
- Blondel M., Kazuhiro S., Kuniaki U. (2013). « Block coordinate descent algorithms for large-scale sparse multiclass classification. » *Machine Learning* 93.1 :31-52.
- Blondel V., Guillaume J-L, Lambiotte R., Lefebvre E. (2008). « Fast unfolding of communities in large networks », *J. Stat. Mech.* P10008.
- Bouba-Olg O., Grossetti M. (2008). «Socio-économie de proximité », *Revue d'économie régionale et urbaine*, n°3.
- Boudon R. (1990). *L'art de se persuader*, Seuil.
- Boullier D., Lohard A. (2012). « Opinion mining et Sentiment analysis », *Collection Sciences Po-médialab*.
- Boyadjian J., Velcin J. (2017). « L'analyse quantitative des médias sociaux, une alternative aux enquêtes déclaratives ? La mesure de la popularité des personnalités politiques sur Twitter », *Questions de communication*, 2017/1 (n° 31), p. 111-135.
- Cha M., Haddadi H., Benevenuto F., Gummadi K. P. (2010). Measuring User Influence in Twitter: The Million Follower Fallacy., *Conference: Proceedings of the Fourth International Conference on Weblogs and Social Media, ICWSM 2010, Washington, DC, USA, May 23- 26, 2010*.

- Cialdini R. (2004). *Influence et manipulation : comprendre et maîtriser les mécanismes de persuasion*, Editions française First-Grund.
- Cialdini R., Wosinksa W., Barrett D. W., Butner J., Gornik-Durose M. (199). Compliance with a request in two cultures: The differential influence of social proof and commitment/consistency on collectivists and individualists. *Personality and Social Psychology Bulletin* 25, 1999, 1242–1253.
- Cialdini R., Vincent J.E., Lewis S.K., Catalan J., Wheeler D. & Darby B.L. (1975). Reciprocal Concessions Procedure for Inducing Compliance: The door-in the face Technique. *JSPS* 31, 206-215.
- Cohendet P., Creplet F. & Dupouët O. (2000). Organisational Innovation, Communities of Practice and Epistemic Communities. in A Kirman & J-B Zimmermann (Eds), *Economics with Heterogeneous Interacting Agents*, Heidelberg: Springer Verlag, p 303-326.
- Crozier, Michel. (1977). *L'Acteur et le Système*. (en collaboration avec Erhard Friedberg), Paris, Le Seuil.
- Dupuy F. (2005). *la fatigue des élites*, Seuil, 2005.
- Dupuy F. (2015). *La faillite de la pensée managériale*. *Lost in management* 2, Seuil, 2015.
- Rosnay JDe. (2006). *La révolte du pronétariat : de mass média aux média de masse*, chez Fayard, janvier, 2006.
- Rosnay JDe. (2012). *Surfer la vie: comment survivre dans une société fluide*, Edition Babel, les liens qui libèrent, 2012.
- Foray D., Lundvall B. Å. (1996). "The knowledge-based economy: from economics of knowledge to learning economy." in Foray & Lundvall (Eds) *Employment and Growth in the Knowledge Economy*, Paris, France: OECD.
- Gigerenzer G. (2009). *Le Génie de l'Intuition - Intelligence et pouvoirs de l'inconscient* - Edition Belfond, 2009.
- Goleman D. (2011). « What Makes a Leader? », *On leadership, HBR's 10 Must Reads*, Boston, MA, Harvard Business Review Press, 1996, second edition, 2011, p 1-22.
- Grandjean M. (2016). A social network analysis of Twitter: Mapping the digital humanities community, *Cogent Arts & Humanities* 3: 1171458
- Grandjean M., Jacomy M., Girard M. (2017). Visual Network Analysis with Gephi, *Digital Humanities*, Krakow, Poland. pp.930-931
- Granovetter, MS. (1973). « The strength of weak ties », *American Journal of Sociology*, 1973, 78(6): 1360-1380.
- Guilain Y. (2015). Le Seeking Mantra au service de l'e-reputation, *I2D – Information, données & documents* 2015/2 (Volume 52), p 46-48.
- Kotter, John P. (2007). « Leading Change: Why Transformation Effort Fails. » *Harvard Business Review*, January, 96-103.
- Jacomy M, Venturini T, Heymann S, Bastian M. (2014). ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software. *PLoS ONE* 9(6): e98679. doi:10.1371/journal.pone.0098679.

- Jammet, T. (2018). Vers une communication de marque dictée par les algorithmes ? Les relations publiques 2.0 face au Big Data. *Communication et organisation*, n°54, p.93-105.
- Karaszewski, R., Lis A. (2013) « The role of leadership to stimulate pro-developmental positive organisation potential. » *Positive management: managing the key areas of positive organisational potential for company success*. Torun, Marek Jacek Stankiewicz ed: 59-88.
- Massé G., Thibaut F. (2000). *Intelligence économique : Guide pour une économie de l'intelligence*, Paris, Editions De Boeck Université.
- Menger PM. (2002). *Portrait de l'artiste en travailleur : métamorphose du capitalisme*, Seuil, La République des Idées.
- Muldoon R. (2013). "Diversity and the Division of Cognitive Labor." *Philosophy Compass* 8/2, pp 117-125.
- Muller P., Pénin J. (2008). Why Do Firms Disclose Knowledge and How Does It Matter?. *Journal of Evolutionary Economics*, pp.85-108.
- Nye J. S. (2004). *Soft Power: The Means to Success in World Politics, Public Affairs*.
- Nye J. S. (2013). *Hard, Soft, and Smart Power, The Oxford Handbook of Modern Diplomacy*. Edited by Andrew F. Cooper, Jorge Heine, and Ramesh Thakur.
- Pecqueur B., Zimmerman JB. (2004). *Economie de proximités*, Edition Hermès-Lavoisier.
- Phillips L., Dowling C., Shaffer K, Hodas N., Volkova S. (2017). Using Social Media to Predict the Future: A Systematic Literature Review, Arxiv
- Rallet A., Torre, A. (2004), Proximité et localisation, *Economie rurale* n°280, p 25- 41.
- Rieder B., Smyrnaioi N. (2012). Pluralisme et infomédiation sociale de l'actualité : le cas de Twitter, *Réseaux* 2012/6 (n° 176), 105-139.
- Rogers R. (2015). *Digital Methods for Web Research, Emerging Trends in the Social and Behavioral Sciences*.
- Sadin E. (2013). *L'humanité augmentée, l'administration numérique du monde*, Paris, éditions l'Echappée.
- Sadin E. (2018) *l'Intelligence artificielle ou l'enjeu du siècle : anatomie d'un anti-humanisme radical*, Paris, éditions l'Echappée.
- Smith M., Rainie L., Shneiderman B., Himelboim I. (2014). *Mapping Twitter Topic Networks: From Polarized Crowds to Community Clusters*, Pew Research Center.
- Sebbah B., Marchand P., Souillard N., Loubère L., Smyrnaioi N., Renard J. (2019). *Les gilets jaunes: le pari gagné de l'existence médiatique ?*, 15 novembre 2019, Rapport de recherche LERASS – Laboratoire d'Études et de Recherches Appliquées en Sciences Sociales, Université de Toulouse
- Sylvestre G. (2016). Détecter des signaux faibles en cartographiant des tweets, *I2D – Information, données & documents* 2016/3 (Volume 53), p 12-15.
- Sylvestre G. (2020). *La Social Graph Due Diligence sur Twitter, une méthodologie d'analyse stratégique qui combine algorithmes de sciences sociales, datavisualisation et expertise humaine*, *Cahiers de la Documentation*, 2020/1.

## VI. Annexes

1. Analyse des échanges Twitter sur les manifestations anti-spécistes en septembre 2018
2. Analyse des échanges Twitter sur #Lapurge en octobre 2018
3. Analyse des échanges Twitter sur le Pacte de Marrakech en novembre – décembre 2018
4. Analyse des échanges Twitter sur Extinction Rebellion à Italie 2 en octobre 2019
5. Analyse des échanges Twitter sur Notre Dame des Landes en janvier 2016
6. Analyse des échanges Twitter sur Europacity en octobre 2019
7. Analyse des échanges Twitter sur la ZAD Dune en janvier 2020
8. Analyse des échanges Twitter sur Extinction Rebellion et Lafarge en février 2020

## 1. Analyse des échanges Twitter sur les manifestations anti-spécistes en septembre 2018

La multiplication des scandales alimentaires entraîne une prise de conscience et une modification des habitudes du grand public : développement du véganisme, création du cuir végan, développement de mouvements de lutte pour le bien-être animal ... avec ces tendances, se développent des groupes d'appartenance où les membres cherchent à se faire entendre au niveau national par des manifestations ou des opérations « coup de poing ». Certains sujets qui soutiennent la polémique vont alors nourrir et faire émerger de nouveaux mouvements allant dans le sens des thématiques abordées.

**L'association L214 est créée en 2008 Brigitte Gothière et Sébastien Arzac.** L214 fonctionne en diffusant des vidéos choquantes sur le fonctionnement des abattoirs pour attaquer l'industrie agroalimentaire. La société civile se mobilise de plus en plus, sur fond de problématique du réchauffement climatique, contre la consommation de viande : **outre l'éthique vis-à-vis de toutes les espèces vivant sur la terre se pose le problème de son impact écologique.**

Les bouchers seront visés par de premières actions non violentes au mois de septembre 2018, amenant diverses réactions que nous allons étudier à travers les mentions de « antispécisme » et « antispécistes » sur Twitter

### Un sujet polémique au sein même des antispécistes

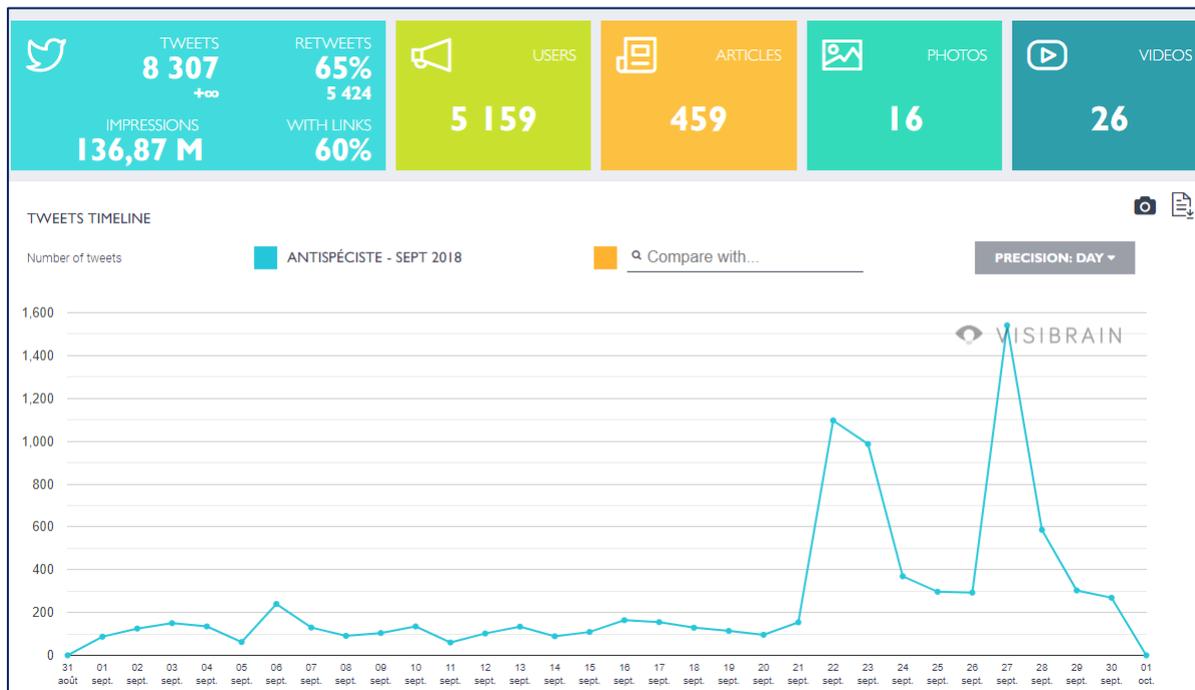
Le classement via Visibrain des tweets les plus repris mentionnant le sujet en septembre 2018 montre des réactions principalement négatives suite aux manifestations anti bouchers organisées à Paris le 22 septembre 2018.

### Classement via Visibrain des Tweets mentionnant l'antispécisme entre le 1er septembre et le 1er octobre 2018

	<b>AymericCaronOfficiel</b> @CaronAymericoff On commence à entendre dans les médias des militants vegans aux discours agressifs qui ne reflètent pas la philosophie antispéciste non-violente que je prône depuis des années. Merci aux médias ne pas tomber dans le piège du sensationnalisme en privilégiant les extrémistes. 27 sept. 2018 11h44 · Twitter for Android		<b>584</b> (RE)TWEETS
	<b>franceinfo</b> @franceinfo "Barrez-vous" : Un boucher répond aux antispécistes qui manifestent devant sa boutique <a href="https://franceinfo.fr/sante/alimenta...pic.twitter.com/zWylXWfEl">franceinfo.fr/sante/alimenta...pic.twitter.com/zWylXWfEl</a> 22 sept. 2018 19h10 · TweetDeck		<b>482</b> (RE)TWEETS
	<b>Verlaïne</b> @_Verlaïne_ Plusieurs dizaines d'actions antispécistes organisées devant des boucheries Personne n'oblige personne à consommer de la viande, ça aussi c'est une liberté acquise tout comme la liberté d'expression de Zemmour. Laissez travailler les bouchers, merci <a href="https://t.co/AiYz5mv7Uf">https://t.co/AiYz5mv7Uf</a> 23 sept. 2018 13h51 · Twitter for Android		<b>310</b> (RE)TWEETS
	<b>LICRA</b> @LICRA_ On a touché le fond de l'odieux. Ce genre de comparaison est une injure aux 6 millions de victimes de la Shoah. Les comparer ainsi à des animaux, c'est une abjection totale. <a href="https://t.co/9lrlJpqsFD">https://t.co/9lrlJpqsFD</a> 27 sept. 2018 17h39 · Twitter Web App		<b>250</b> (RE)TWEETS
	<b>Eugénie Bastié</b> @EugenieBastie Pour cette militante végan, abattre des animaux dans des conditions éthiques, c'est comme "violer avec empathie". L'antispécisme dans toute sa bêtise. (Tout le débat est surréaliste). <a href="https://t.co/Y3kPd1gOf">https://t.co/Y3kPd1gOf</a> 27 sept. 2018 19h42 · Twitter Web Client		<b>154</b> (RE)TWEETS
	<b>Libération Animale</b> @BoohVegan Les #animaux ne sont pas : Des ressources. Des produits. Des objets. De la nourriture. Des vêtements. Des divertissements. Des moyens de transport. Des sujets d'expérience. À notre disposition. À nos ordres. Destinés à nous servir. #antispécisme #veganisme 16 sept. 2018 22h16 · Twitter Web App		<b>146</b> (RE)TWEETS

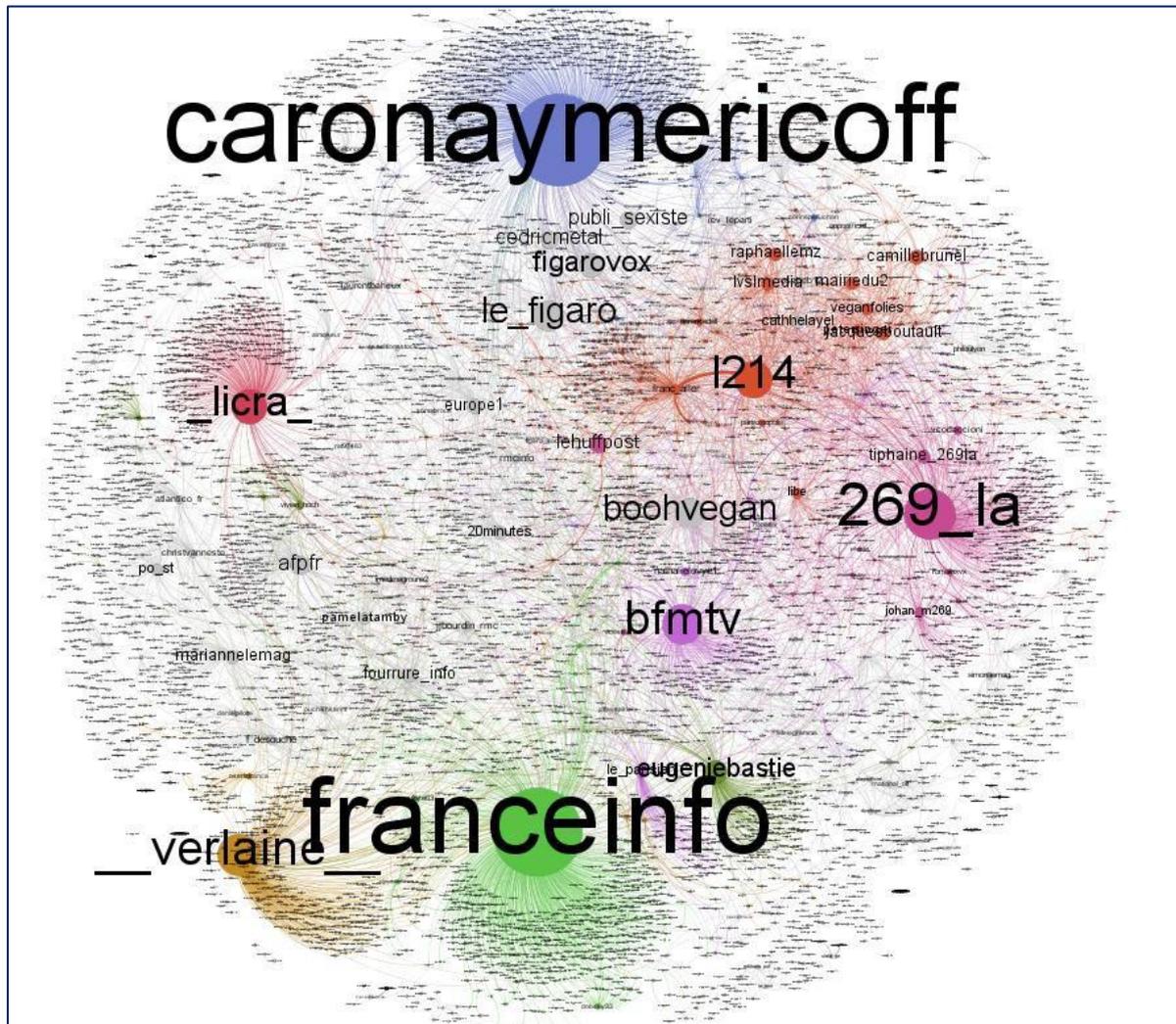
La Licra s'est notamment indignée d'une comparaison effectuée par une militante antispéciste entre l'Holocauste et la situation des animaux dans les abattoirs. Pour autant, les comptes militants antispécistes @BoohVegan et @269\_LA sont également très repris. Avant l'action du 22 septembre, il n'y avait guère que les militants qui étaient actifs comme on le voit sur l'évolution des tweets :

**Timeline via Visibrain des mentions Twitter sur l'antispécisme entre le 1er septembre et le 1er octobre 2018**



Le sujet suscite 65 % de retweets sur le mois de septembre, et 57 % avant l'action du 22 septembre, suggérant que les militants sont des acteurs plutôt actifs. On identifie avec Gephi plus de 700 communautés sur les tweets échangés en septembre, les 8 premières regroupant 50 % des échanges. On retrouve ici le rôle des communautés jouant un rôle pro-actif de leader go-between tant chez les militants anti-spécisme que chez les militants pro-spécisme (Nye, 2004, Goleman, 2011, Baulant, 2015).

**Cartographie via Gephi et Visibrain des comptes les plus mentionnés par leurs communautés sur l'antispécisme entre le 1er septembre et le 1er octobre 2018**



La cartographie des comptes les plus influents dans les échanges illustre cette fragmentation, avec des communautés qui échangent peu entre elles. Il existe une multitude de petits courants au sein même des deux grands courants spécisme et anti-spécisme. Cette scission s'explique par l'apparition explicite de problème d'éthique qui se double et renforce les problèmes de souffrance ou bien-être animal. Cet émiettement des discours se retrouve dans l'analyse des tweets. Les leaders go-between sont en effet peu repris en dehors de leurs communautés propres.



## Identification des leaders go-between

Nous avons analysé les échanges du 1<sup>er</sup> au 21 septembre inclus, soit la veille des actions militantes. Ce corpus correspond à 2565 tweets, dont 1127 uniques, envoyés par 1534 utilisateurs. Le volume de tweets est assez régulier sur la période étudiée :

### Timeline via Visibrain des mentions Twitter sur l'antispécisme entre le 1er et le 21 septembre 2018



On identifie ainsi les comptes Twitter qui étaient au centre des échanges avant les manifestations du 22 septembre 2018, et ont été au cœur de la viralité du sujet.





## Classement via Visibrain des Tweets mentionnant l'antispécisme entre le 19 avril et le 11 mai 2019

	<p><b>Brigitte Benkemoun</b> @bribenkemoun</p> <p>Et ce matin Steven, seul boucher bio du #marchesaintquentin est fermé. Il a été blessé hier par le commando vegan qui ont attaqué son stand. Steven doit avoir 30 ans, il est installé depuis moins d'un an. Il se bat pour avoir la meilleure viande dans un respect animal. #Paris10 pic.twitter.com/9Xq8g1jwwM</p> <p>05 mai 2019 11h31 · Twitter for iPhone</p>		<p><b>1 402</b> (RE)TWEETS</p>
	<p><b>Brigitte Benkemoun</b> @bribenkemoun</p> <p>Et ce matin Steven, seul boucher bio du #marchesaintquentin est fermé. Il a été blessé hier par le commando vegan qui a attaqué son stand. Steven doit avoir 30 ans, il est installé depuis moins d'un an. Il se bat pour avoir la meilleure viande dans un respect animal. #Paris10 pic.twitter.com/w4va6Ra0yY</p> <p>05 mai 2019 11h53 · Twitter for iPhone</p>		<p><b>218</b> (RE)TWEETS</p>
	<p><b>Jean-Pierre Denis</b> @jeanpierredenis</p> <p>Un jeune et vaillant artisan bio, qui se trouve être mon boucher, attaqué par un commando vegan, qui agit comme une secte fasciste. Il va falloir s'interroger sur ces actions de + en + aveugles et violentes, se demander quel monde elles promettent, quels intérêts elles servent.</p> <p><a href="http://twitter.com/BFMParis/statuses/1125306232368173056">http://twitter.com/BFMParis/statuses/1125306232368173056</a></p> <p>06 mai 2019 12h52 · Twitter for iPhone</p>		<p><b>143</b> (RE)TWEETS</p>
	<p><b>FranceNews24</b> @FranceNews24</p> <p>● INFO - #Paris : Un boucher bio du #marchesaintquentin a été blessé hier par un commando #vegan qui ont attaqué son stand Boulevard de Magenta dans le 10<sup>e</sup> arrondissement." pic.twitter.com/ya4B4ngciZ</p> <p>05 mai 2019 14h50 · Twitter Web Client</p>		<p><b>98</b> (RE)TWEETS</p>
	<p><b>Didier GUILLAUME</b> @dguillaume26</p> <p>Soutien total au jeune artisan #boucher attaqué lâchement ce week-end par des militants #antispecistes. Ces actes sont intolérables ! pic.twitter.com/GTbFIGvSZ</p> <p>07 mai 2019 10h09 · Twitter for Android</p>		<p><b>91</b> (RE)TWEETS</p>

Comme pour le cas de la viralité du projet d'Europacité, on s'aperçoit que le discours transversal des communautés organisées autour d'un leader social est bien plus efficace que le discours politique d'un Ministre dépositaire de la loi ou des partis politiques ou des syndicats qui sont eux aussi organisée de façon hiérarchisée et « top down » (Goleman, 2011, Muldoon, 2013, Baulant, 2017). Dans une société complexe (Morin, 2011), l'image (mise en avant d'exemples extrêmes) prime sur la rationalité et sur les discours structurés qui deviennent alors non lisibles, non seulement en raison de la désaffectation du grand public pour ses structures organisées (baisse taux de syndicalisation) mais aussi en raison de la lenteur et, parfois, de la lourdeur de la réponse apportée, devant en principe être validée d'en haut, mais qui trop souvent devient non validée et crée alors des déclarations qui peuvent être contraires entre les membres d'un même gouvernement, d'un même parti ou d'un même syndicat.

Le mouvement anti-spécisme est donc particulièrement atypique sur Twitter, car il mobilise très fortement une communauté de défenseur de droits des animaux (L214), et ponctuellement d'autres communautés comme les vegans. Le lien entre les deux communautés n'est pas géré par l'utilisation efficace des « liens faibles » (Granovetter, 1973, Baulant 2017) et les leaders go-between des deux communautés échouent à employer « les bons termes » sur les réseaux sociaux et les « bonnes vidéos sur les actions menées » pour pouvoir être suffisamment vastes pour que l'ensemble des communautés puissent s'y reconnaître (Girard, 1982, 2007 ;). L'aspect violent des deux mouvements est un élément qui explique l'absence de viralité. L'aspect éthique propre à chaque communauté explique aussi cet absence de contagion mimétique (Dupuy, 1982, 2005). Pourtant le mouvement anti-spécisme est amené à prendre de l'ampleur avec la lutte contre le réchauffement climatique. En réaction, les opposants au mouvement vont être amené à se structurer, d'autant plus facilement qu'ils ont réussi à intégrer le monde agricole, qui s'est fortement investi sur Twitter en 2019.

## 2. Analyse des échanges Twitter sur #LaPurge en octobre 2018

### Analyse des tweets

#LaPurge est un appel à agresser les policiers pendant la nuit d'Halloween 2018. Un signalement a été effectué le 27 octobre, et le lendemain le hashtag a été partagé : <https://twitter.com/cestsafyo/status/1056142102386696192>

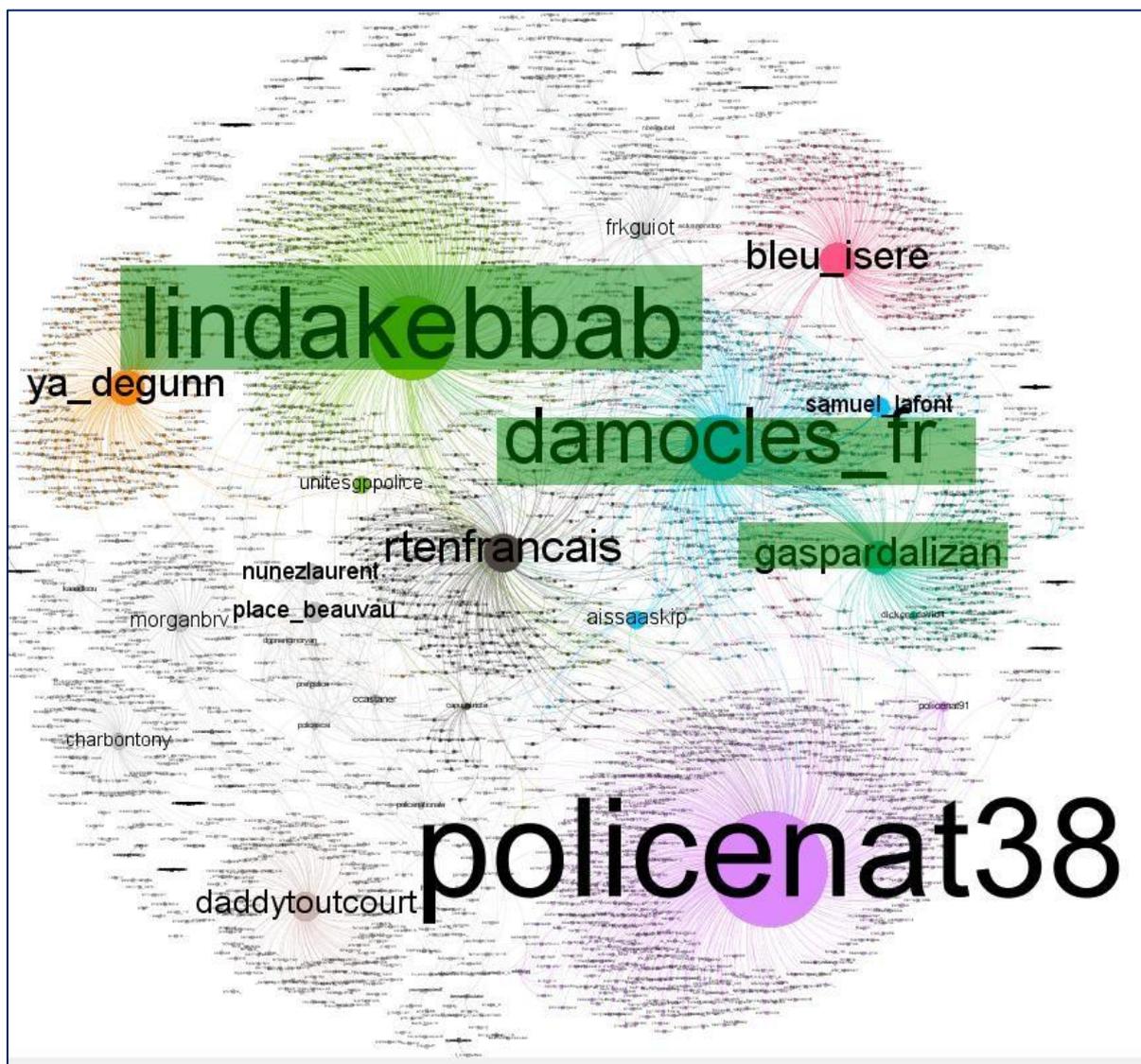
Il a généré 3458 mentions Twitter du 28 octobre 2018 au 12 novembre 2018, dont 88 % de retweets, par 2679 comptes. Les tweets les plus repris sont des soutiens aux forces de l'ordre :

### Classement via Visibrain des Tweets mentionnant #LaPurge du 27 octobre 2018 au 12 novembre 2018

 <p><b>Police nationale 38</b> @PoliceNat38</p> <p>⚠️ #Interpellation Un suspect qui a appelé à #LaPurge à #Grenoble vient d'être placé en #GardeàVue pour provocations à la violence contre les forces de l'ordre. ⚠️ Ne relayez pas ce type de message sous peine d'être poursuivi à votre tour. <a href="https://pic.twitter.com/coamw7uFRP">pic.twitter.com/coamw7uFRP</a></p> <p>29 oct. 2018 17h24 Twitter Web Client</p>		<p><b>530</b> (RE)TWEETS</p>
 <p><b>Linda Kebbab</b> @LindaKebbab</p> <p>Bon courage à nos collègues 👮👮 qui vont devoir faire face à des voyous rassemblés ce soir pour la #lapurge. Force et honneur face à ceux dont le passe-temps est de pourrir la vie d'autrui. Unité SGP Police, de tout cœur avec les policiers #Halloween #Halloween2018 #LaPurge</p> <p>31 oct. 2018 22h03 Twitter for iPhone</p>		<p><b>460</b> (RE)TWEETS</p>
 <p><b>France Bleu Isère</b> @bleu_isere</p> <p>🔴 INFO FRANCE BLEU - Le jeune Isérois qui avait appelé à "la purge" des policiers pendant la nuit d'Halloween veut porter plainte contre le ministre de l'Intérieur. Il juge la réaction de Christophe Castaner, "disproportionnée". #grenoble #lapurge → <a href="https://francebleu.fr/infos/faits-di...">francebleu.fr/infos/faits-di...</a> <a href="https://pic.twitter.com/OBMIFXhMxA">pic.twitter.com/OBMIFXhMxA</a></p> <p>07 nov. 2018 18h03 Twitter Web Client</p>		<p><b>185</b> (RE)TWEETS</p>
 <p><b>yadegdeg</b> @Ya_degunn</p> <p>Tous les attardés demain soir ils seront comme ça dans les rues de paname pour la purge 😞😞</p> <p>#lapurge #purge <a href="https://pic.twitter.com/fmicHKMijz">pic.twitter.com/fmicHKMijz</a></p> <p>30 oct. 2018 21h38 Twitter for iPhone</p>		<p><b>185</b> (RE)TWEETS</p>
 <p><b>Gaspard Alizan</b> @GaspardAlizan</p> <p>Il n'aura fallu qu'une simple vidéo d'un demeureur pour déclencher #LaPurge de hier soir, imaginez votre pays quand ils seront encore plus nombreux et que l'appel proviendra d'une personne qu'ils respectent...</p> <p>01 nov. 2018 09h30 Twitter Web Client</p>		<p><b>173</b> (RE)TWEETS</p>
 <p><b>Damocles</b> @Damocles_Fr</p> <p>🔴 Des racailles appellent à tout casser, brûler et à attaquer les forces de l'ordre dans la nuit du 31 octobre. Le message est pris très au sérieux par les autorités. #lapurge <a href="https://pic.twitter.com/EjGfcsyHL">pic.twitter.com/EjGfcsyHL</a></p> <p>29 oct. 2018 12h46 Twitter Web Client</p>		<p><b>153</b> (RE)TWEETS</p>

On identifie via Gephi 221 communautés, dont les 8 premières totalisent plus de 70 % des comptes Twitter. La cartographie correspond à un groupe de marques, avec différentes communautés qui n'échangent pas entre elles, en dehors des communautés vertes, bleu clair et noir qui elles sont des groupes de communautés. Les leaders de ces communautés, surlignés en vert, sont au centre des échanges :

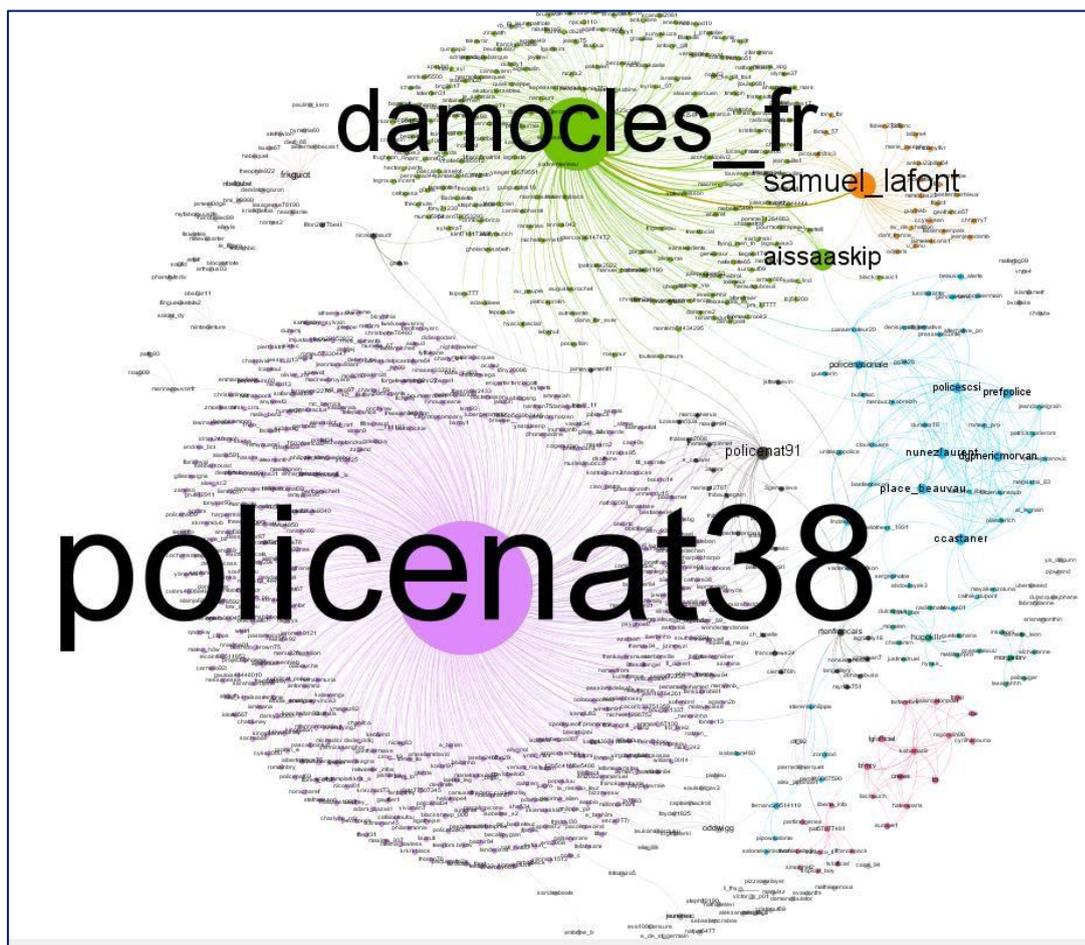
**Cartographie via Gephi et Visibrain des comptes au centre des échanges sur #LaPurge entre le 27 octobre 2018 et le 12 novembre 2018**



**Identification des leaders go-between avant le 31 octobre**

Il y a eu 1137 tweets publiés par 898 comptes Twitter avec le hashtag #lapurge du 28 au 30 octobre inclus. On identifie via Gephi 40 communautés, les deux premières totalisant plus de 70 % des comptes Twitter. Il s'agit là aussi de groupes de marques, le compte @Damocles\_fr étant au centre des échanges :

## Cartographie via Gephi et Visibrain des comptes au centre des échanges sur #LaPurge entre le 27 et le 31 octobre 2018



A noter que la communauté en bleu clair, autour des comptes @PoliceSCSI et @PoliceNationale, obtient un score de centralité important, contrairement au compte plus repris @PoliceNat38 (interpellation d'un suspect ayant appelé à #LaPurge à Grenoble). L'analyse des comptes au centre des échanges révèle 5 leaders go-between sur 933 comptes Twitter.

On notera également, outre l'importance de la source d'extrême droite Damoclès, la forte visibilité du média Russia Today les 30 et 31 octobre, les médias nationaux français étant absents :

**Filtrage via Visibrain des Tweets de @RTenfrancais mentionnant #LaPurge du 27 octobre 2018 au 12 novembre 2018**

	<p><b>RT France</b> @RTenfrancais</p> <p>• «#Purge» anti-#police : «mobilisation renforcée» des forces de l'ordre, premiers incidents à #Lyon #lapurge</p> <p><a href="https://francais.rt.com/france/55035-p...pic.twitter.com/7IICIGvBnH">francais.rt.com/france/55035-p...pic.twitter.com/7IICIGvBnH</a></p> <p>31 oct. 2018 22h38 Twitter Web Client</p>		<p><b>136</b> (RE)TWEETS</p>
	<p><b>RT France</b> @RTenfrancais</p> <p>• «#Purge» anti-#police : dans l' #Essonne, «une vingtaine de personnes masquées, cagoulées, ou encapuchonnées» ont #pillé un magasin #Intersport, selon Le Parisien #lapurge</p> <p><a href="https://francais.rt.com/france/55038-p...pic.twitter.com/Tz89TL3G8U">francais.rt.com/france/55038-p...pic.twitter.com/Tz89TL3G8U</a></p> <p>31 oct. 2018 23h10 Twitter Web Client</p>		<p><b>68</b> (RE)TWEETS</p>
	<p><b>RT France</b> @RTenfrancais</p> <p>• «#Purge» anti-#police : des #incendies ont été déclenchés sur la voie publique dans plusieurs communes du #Rhône, selon Le Progrès #lapurge</p> <p><a href="https://t.co/wTVnmeoM3k">https://t.co/wTVnmeoM3k</a></p> <p>31 oct. 2018 23h19 Twitter Web Client</p>		<p><b>57</b> (RE)TWEETS</p>
	<p><b>RT France</b> @RTenfrancais</p> <p>#Lapurge : Une «blague» revendiquée par un internaute <a href="https://francais.rt.com/france/54968-b...">francais.rt.com/france/54968-b...</a></p> <p>30 oct. 2018 07h30 TweetDeck</p>		<p><b>9</b> (RE)TWEETS</p>

### 3. Analyse des échanges Twitter sur le Pacte de Marrakech en novembre – décembre 2018

Le « Pacte mondial pour des migrations sûres ordonnées et régulières » (pacte de Marrakech) a été signé lors d'une conférence intergouvernementale qui s'est tenue du 10 au 11 décembre 2018 à Marrakech. Il a pour objectif de clarifier les questions de la migration à l'échelle internationale. Il est le fruit de plusieurs mois de travail des Etats membres des Nations unies.

Dans plusieurs pays, des politiques et partis opposés à l'immigration vont propager des rumeurs sur les conséquences de ce pacte : un abandon de souveraineté de la France à l'ONU sur les questions migratoires, l'ouverture totale des frontières, etc ... en s'appuyant sur le faible echo médiatique du sujet.

#### Analyse des tweets sur le pacte de Marrakech

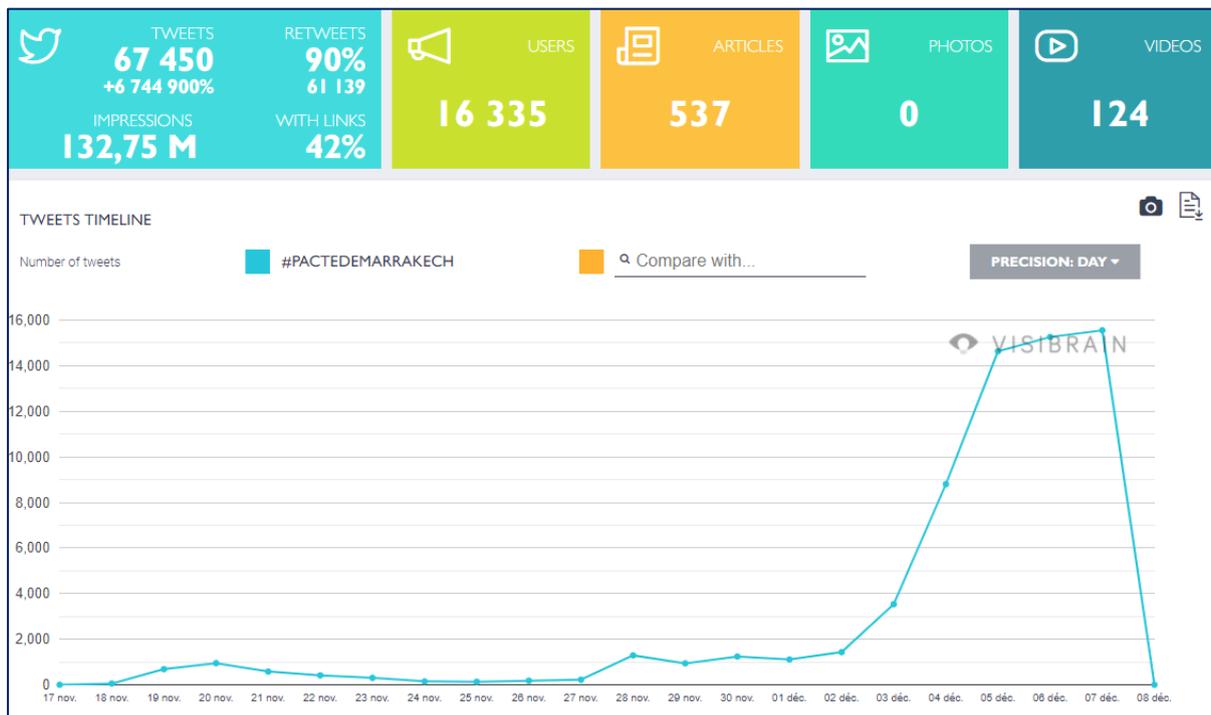
Sur Twitter, le hashtag #Pactedemarrakech est apparu le 18 novembre via le parti belge d'extrême droite Vlaams Belang, opposé à la signature du Pacte. En France, avant l'apparition du hashtag, [c'est le militant d'extrême droite Julien Rochedy qui a le premier donné de la visibilité au sujet](#). Le pacte de Marrakech restera ensuite relativement peu mentionné en France, jusqu'à la reprise par des élus nationaux à partir du 3 décembre, devenant ensuite viral jusqu'à une première manifestation organisée en Belgique le 8 décembre. En tout, près de 300 000 tweets en français vont mentionner le pacte de Marrakech en 3 semaines, [amenant d'ailleurs le site officiel du gouvernement à démentir les rumeurs](#). Le classement des tweets les plus repris sur le Pacte de Marrakech montre une domination des élus du Rassemblement National et de Républicains entre le 17 novembre et le 8 décembre :

#### Classement via Visibrain des Tweets les plus repris sur le pacte de Marrakech du 17 novembre au 8 décembre 2018 :

	<b>Valérie Boyer</b> @valerieboyer13 En pleine manifestation des #GiletsJaunes, la majorité a fait le choix la semaine dernière d'avoir un débat sur les fessées mais refuse un grand débat sur le #Pactedemarrakech en matière migratoire. Nous marchons sur la tête! #DirectAN pic.twitter.com/8ZojBqoMHT 05 déc. 2018 12h03 · Twitter Web Client		<b>2 651</b> (RE)TWEETS
	<b>Damien Rieu</b> @DamienRieu Contre le dangereux #Pactedemarrakech pro-migrants, mobilisons-nous! #PacteMondialSurLesMigrations #GiletsJaunes pic.twitter.com/vO3x7MH6kd 05 déc. 2018 21h05 · Twitter Web Client		<b>2 234</b> (RE)TWEETS
	<b>Marine Le Pen</b> @MLP_officiel Le Président pense camoufler la trahison du pays que représente la signature du #Pactedemarrakech en envoyant un sous-fifre signer à sa place. Pourquoi refuse-t-il d'entendre que les Français ne veulent plus de cette immigration de masse? MLP #PacteMondialSurLesMigrations 07 déc. 2018 11h40 · Twitter for iPhone	⌵	<b>1 543</b> (RE)TWEETS
	<b>Gilbert Collard</b> @GilbertCollard Si le #Pactedemarrakech est aussi inoffensif que le chantent les journalistes français, pourquoi USATLRSK CZCHAUHTBGRDOPLEEDKNL refusent de le signer? 07 déc. 2018 12h30 · Twitter for iPhone	⌵	<b>1 512</b> (RE)TWEETS
	<b>Marine Le Pen</b> @MLP_officiel Et au fait, il en pense quoi @laurentwauquiez du pacte de Marrakech, qui va organiser la submersion migratoire de notre pays? MLP 04 déc. 2018 09h25 · Twitter for Android	⌵	<b>1 027</b> (RE)TWEETS
	<b>Les Républicains</b> @lesRepublicains « Emmanuel Macron ne doit pas signer le #Pactedemarrakech. Depuis des années, l'immigration de masse que connaît la France est soustraite à la délibération démocratique. Chaque peuple a le droit de décider souverainement qui il souhaite accueillir. » republicains.fr/actualites_pre... 05 déc. 2018 15h13 · Twitter Web Client	⌵	<b>1 024</b> (RE)TWEETS

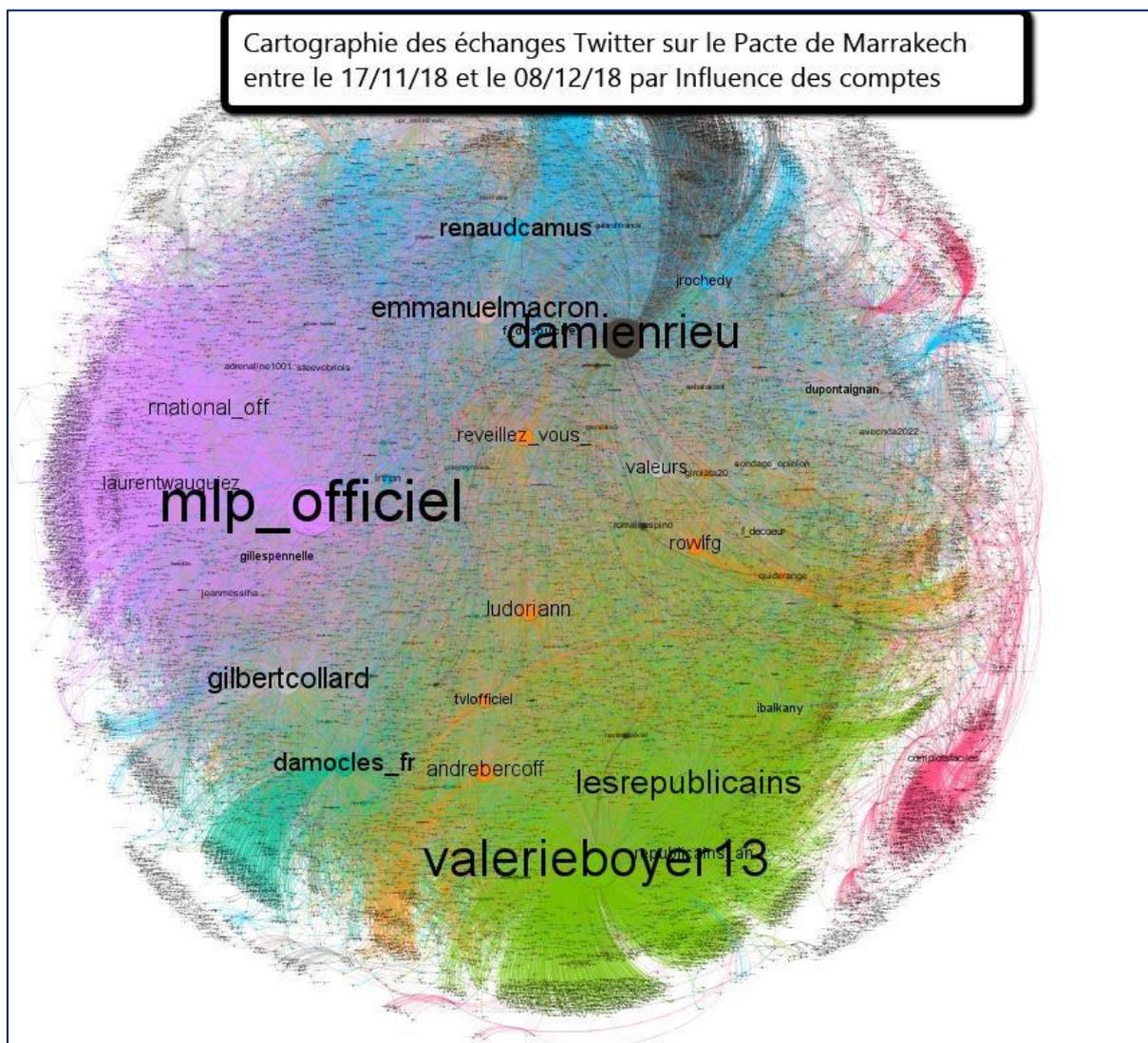
Pour autant, on notera qu'aucun de ces acteurs ne s'intéressait au sujet avant le 2 décembre, bien qu'il fasse déjà l'objet d'échanges. Si on suit la courbe d'évolution des tweets, on note en effet une évolution très rapide du volume des tweets, dû à un nombre très forts de retweets, 90 % ici :

**Timeline via Visibrain des mentions Twitter sur le pacte de Marrakech du 17 novembre au 8 décembre 2018 :**



On identifie avec Gephi 700 communautés sur l'ensemble des 16 335 comptes Twitter actifs, les 8 premières représentant plus de 70 % des comptes. Les deux communautés les plus importantes ont un poids similaire, avec environ 16 % des comptes Twitter leur appartenant.

**Cartographie via Gephi et Visibrain des comptes les plus mentionnés par leurs communautés sur le pacte de Marrakech du 17 novembre au 8 décembre 2018 :**

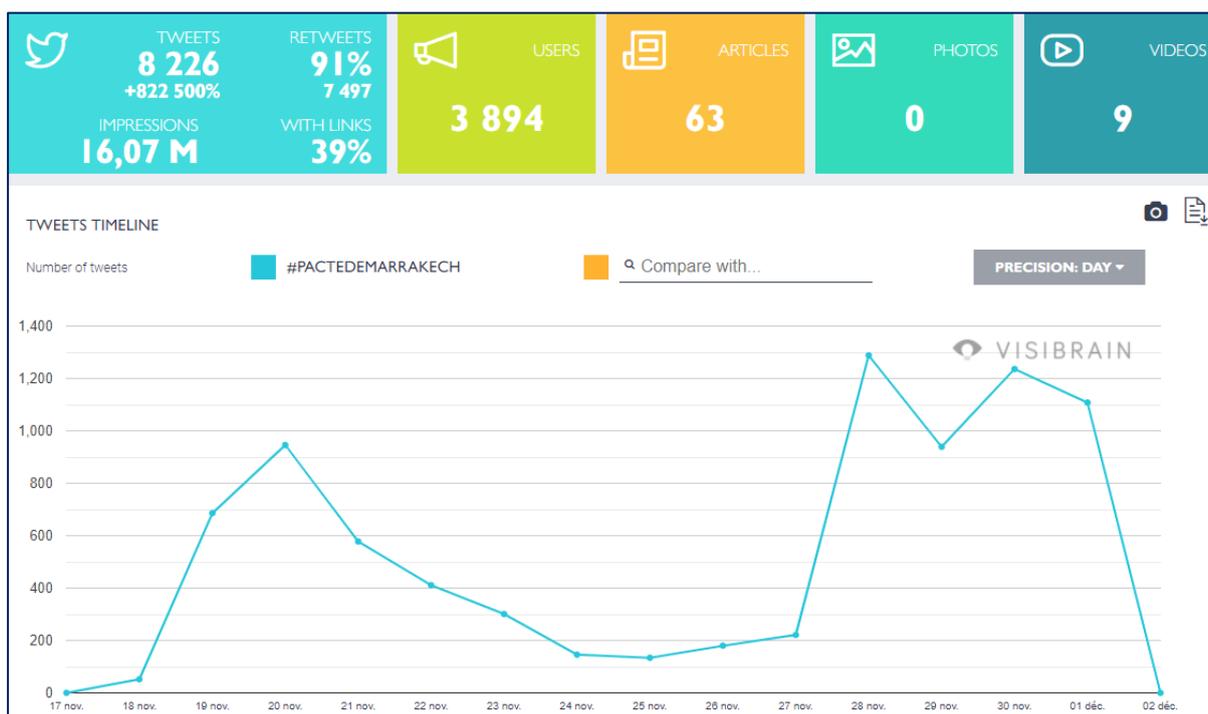


**Identification des leaders go-between**

Pour ce corpus, nous avons considéré que la période de viralité commençait au 3 décembre, jusqu'à la première manifestation du 8 décembre. L'export Gephi comprend les tweets jusqu'au 2 décembre, soit la veille de l'intervention de Marine Le Pen.

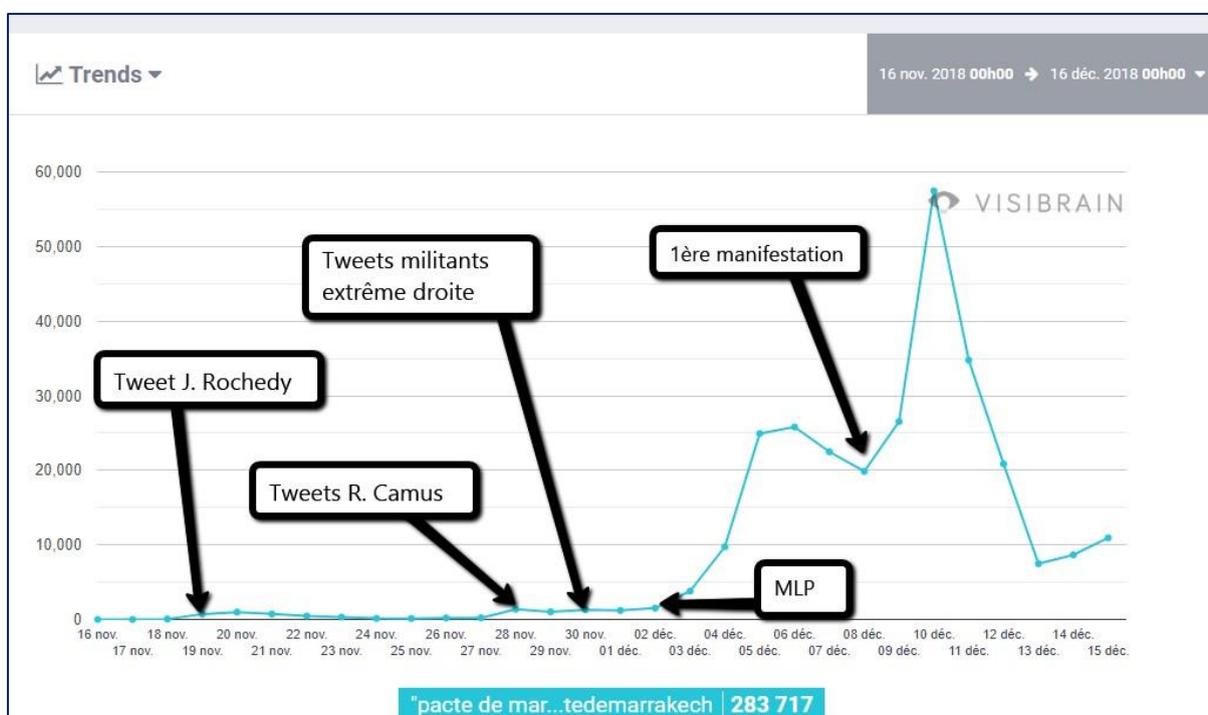
Du 17 novembre 2018 au 2 décembre 2018, 8226 tweets ont été partagés, par 3894 utilisateurs, avec toujours 90 % de retweets. Après des échanges nombreux suite aux premiers tweets, il y a eu un tassement des échanges, jusqu'à une reprise plus forte de fin novembre à début décembre. Marine Le Pen sera la première personnalité politique à mentionner le sujet à l'antenne le 2 décembre.

**Timeline via Visibrain des mentions Twitter sur le pacte de Marrakech du 17 novembre au 1er décembre 2018 :**



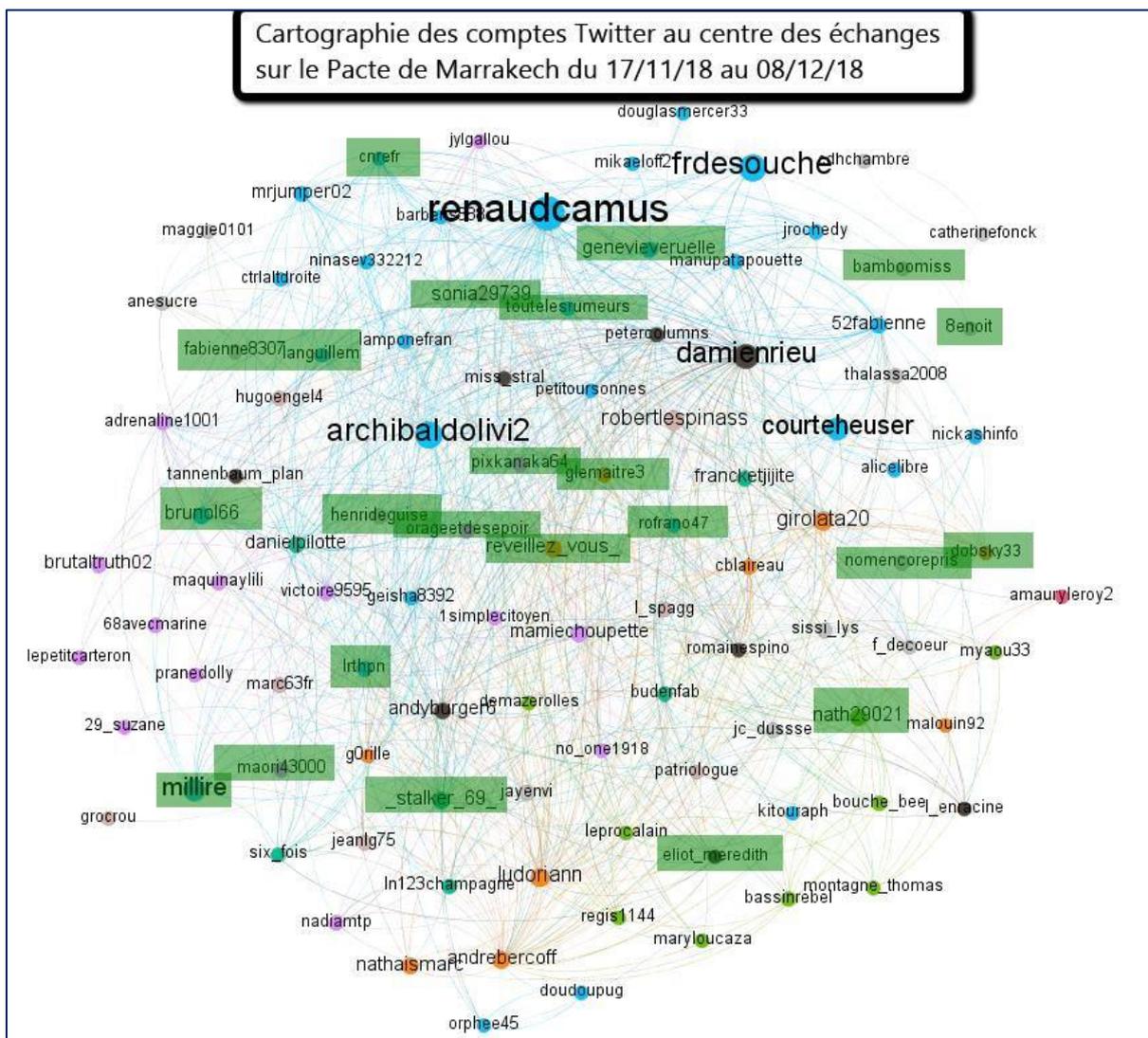
Si on reprend l'évolution globale des tweets sur le sujet, on constate que c'est sur une très courte période, la semaine de la signature du traité, que le sujet a pris de l'ampleur :

**Timeline via Visibrain des mentions Twitter sur le pacte de Marrakech du 16 novembre au 16 décembre 2018 :**



En reprenant la cartographie des comptes au centre des échanges du 17 novembre au 8 décembre, on met en évidence les leaders go-between actifs avant le 2 décembre, surlignés en vert. **Ceux-ci ont donc fortement contribué au succès de cette Fake News, et correspondent principalement aux communautés militantes en dehors des réseaux LR et FN.**

**Cartographie via Gephi et Visibrain des comptes au centre des échanges sur le pacte de Marrakech du 17 novembre au 8 décembre 2018 – application d’un filtre pour garder les comptes ayant un score élevé**



**Des communautés militantes locales et structurées à l’origine de la viralité**

Selon la méthodologie du Pew Research Center, les échanges sur le Pacte de Marrakech du 17 novembre au 2 décembre 2018 correspondent à une foule étroite : plusieurs petits groupes qui échangent entre eux. **C’est en arrivant à influençant les communautés politiques liées aux dirigeants du FN et de LR que la Fake News arrivera à être très reprise.**

Cette forte diffusion correspond à un important travail militant visant à exploiter spécifiquement Twitter pour obtenir une forte visibilité. La création très rapide d'un hashtag dédié a vite permis au sujet d'arriver en trending topic, en exploitant très peu d'articles ou d'informations concrètes. Plusieurs communautés militantes reprennent le sujet en se présentant comme des « médias alternatifs », avec le soutien de branches locales de partis politiques. **Il n'y a pas ici de contre discours avant que le sujet devienne viral.**

#### 4. Analyse des échanges Twitter sur Extinction Rebellion à Italie 2 en octobre 2019

Mouvement né au Royaume-Uni en automne 2018 et qui est présent dans 55 pays, Extinction Rebellion prône la désobéissance civile via des actions coup de poing, tout en restant non-violent. Ce sont notamment des actions illégales de blocage d'institutions, d'entreprises, de lieux publics symboliques ou stratégiques comme des ponts, des routes ou des aéroports, pour dénoncer l'inaction face au dérèglement climatique et à l'extinction de la biodiversité.

Extinction Rebellion France a été lancé le 24 mars 2019 et a recruté très vite plusieurs milliers de militants prêts à s'engager pour ses actions. Ceux-ci n'ont souvent jamais milité, et se retrouvent dans une structure très horizontale, sans réelle hiérarchie. A partir de 3 militants, des actions peuvent être organisées. Cela en fait un mouvement capable de monter très vite en puissance et dont il est difficile d'anticiper la portée.

Ainsi, l'action « [Notre cœur brûle](#) » sur les incendies de forêt en Amazonie, organisée à Paris le 31 août, a réuni 250 militants qui se sont coordonnés en moins d'une semaine. La militante qui a proposé l'action avait rejoint le mouvement un mois plus tôt.

#### Analyse des tweets sur la première grosse action de Extinction Rebellion France

L'occupation du centre commercial Italie 2, organisée dans le cadre d'une semaine de mobilisation internationale du mouvement, a eu un très fort écho sur Twitter, via des médias spécialisés, des militants traditionnels et des politiques de l'opposition qui ne font pas partie d'Extinction Rebellion :

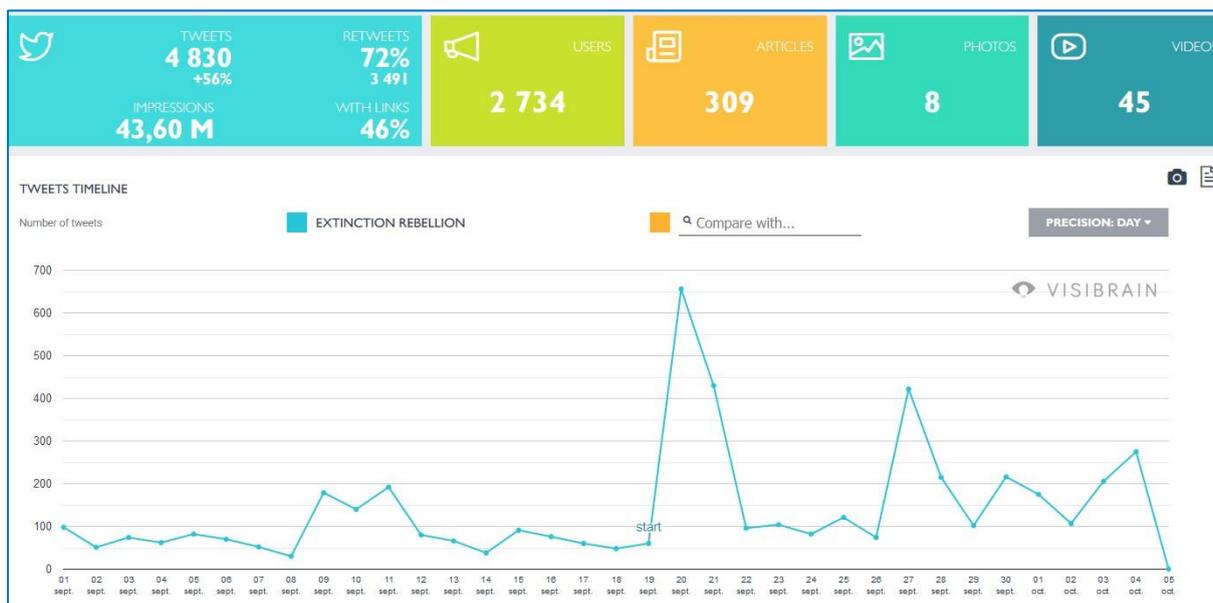
#### Classement via Visibrain des Tweets les plus repris sur Extinction Rebellion en France du 1er septembre au 6 octobre 2019 :

	<b>Là-bas si j'y suis</b> @LabasOfficiel DIRECT - #Paris 300 militants écologistes et Gilets jaunes occupent actuellement le centre commercial #Italie2 à Paris pour dénoncer la société de consommation et l'urgence climatique. #ExtinctionRebellion images : @T_Bouhafis pic.twitter.com/FH4zv39QjC 05 oct. 2019 10h39 Twitter Web App		<b>532</b> (RE)TWEETS
	<b>Anonyme Citoyen</b> @AnonymeCitoyen Les #GiletsJaunes envahissent le centre commercial des Grands Hommes à #Bordeaux. #Acte47 #ExtinctionRebellion #Macron #CastanerDemission pic.twitter.com/Jl8Dq7VAM0 05 oct. 2019 15h14 Twitter for Android		<b>384</b> (RE)TWEETS
	<b>Mathilde Larrere</b> @LarrereMathilde C'est dingue ce qui se passe à Italie 2. Convergence des luttes en pratique Comité adama Extinction rébellion Gilets jaunes Des hong-kongais sont venus témoigner Et le homard qui garde l'entrée ( photo @MarionLpz ) pic.twitter.com/RZkVYoVtvV 05 oct. 2019 15h41 Twitter for iPhone		<b>286</b> (RE)TWEETS
	<b>Реми ф</b> @Callystor "Travaille, consomme, et ferme ta gueule." Occupation par des #GiletsJaunes et militants #ExtinctionRebellion du centre commercial #Italie2 à #Paris. #Acte47 #ActeXLVII pic.twitter.com/JdLUGxxFTq 05 oct. 2019 11h14 Twitter for Android		<b>270</b> (RE)TWEETS
	<b>Marcel Aiphan</b> @AiphanMarcel C'est sûrement une coïncidence mais n'y a toujours ni intervention policière ni casse au centre commercial #ItalieDeux que les #GiletsJaunes, #ExtinctionRebellion et #JusticePourAdama occupent pacifiquement depuis 4 heures... Étonnant, non ? #Acte47 #Paris pic.twitter.com/VsM61on8y3 05 oct. 2019 14h24 Twitter for iPhone		<b>268</b> (RE)TWEETS
	<b>Manon Aubry</b> @ManonAubryFr La mobilisation climatique passe un cap ce matin ! L'occupation par des centaines d'activistes d'un temple de la sur-consommation en plein Paris n'est que la première étape, continuons la lutte contre ce système qui détruit nos vies et la planète ! #ExtinctionRebellion <a href="https://t.co/SvYUCnqpGd">https://t.co/SvYUCnqpGd</a> <a href="http://twitter.com/xrFrance/statuses/1180414080735174656">http://twitter.com/xrFrance/statuses/1180414080735174656</a> 05 oct. 2019 12h51 Twitter for Android		<b>222</b> (RE)TWEETS





## Timeline via Visibrain des mentions Twitter sur Extinction Rebellion en France du 1er septembre au 4 octobre 2019 :



Il y a 4830 mentions d'Extinction Rebellion du 1<sup>er</sup> septembre au 4 octobre 2019 inclus, dont 72 % de retweets, via 2734 utilisateurs. Le top des tweets les plus partagés permet de constater une reprise du mouvement par des Gilets Jaunes :

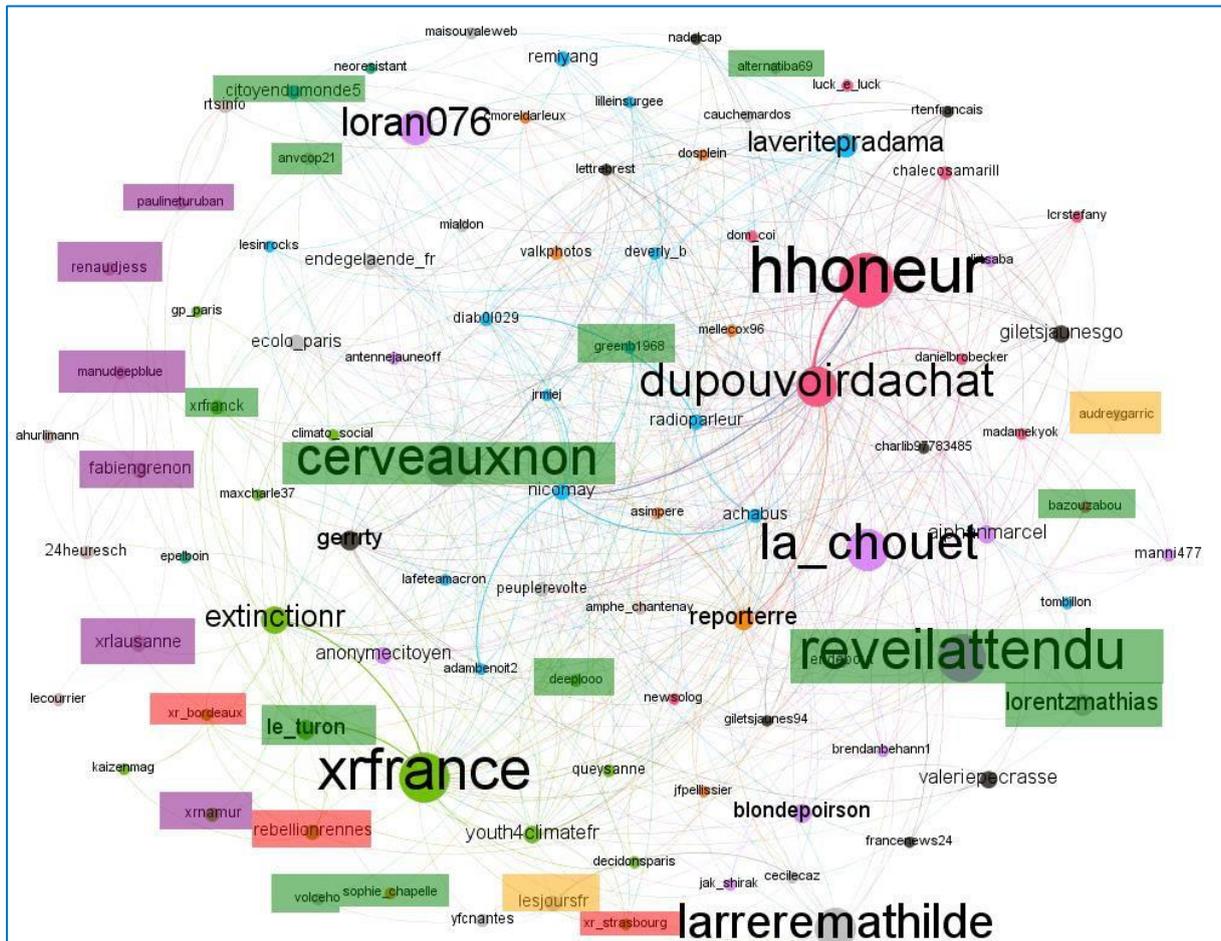
## Classement via Visibrain des Tweets les plus repris sur Extinction Rebellion en France du 1er septembre au 4 octobre 2019 :

<p><b>Lorentz mathias</b> @LorentzMathias</p> <p>Vidéo d'une main arrachée par une #Grenade lors de l'ActeXIII à #Paris ⚠️ Rien d'Irréparable ⚠️ selon #Macron! #FoutageDeGueuleAbsolu! #GiletsJaunes #CitoyensEnColere #ExtinctionRebellion #ConvergenceDesLuites #21Septembre #Acte45 #ViolencesPoliçieres #FinDePartie #GameOver pic.twitter.com/T9zDSZ8A17</p> <p>20 sept. 2019 09h33 Twitter for iPhone</p>		<p><b>217</b> (RE)TWEETS</p>
<p><b>Lorentz mathias</b> @LorentzMathias</p> <p>Vidéo très rare de l'Essoufflement des #GiletsJaunes pour cet #Acte45 à #Paris! Les #Merdias parlait de nous au passé, et la, ben c'est le drame pour eux! #PlusCredible 🤔 #ConvergenceDesLuites 🟡 #CitoyensEnColere #ActionClimat #ExtinctionRebellion pic.twitter.com/7k63rtIVLs</p> <p>21 sept. 2019 13h57 Twitter for iPhone</p>		<p><b>174</b> (RE)TWEETS</p>
<p><b>Extinction Rebellion France</b> 🇫🇷 @XtinctionRebel</p> <p>📢 En octobre, #ExtinctionRebellion mènera une rébellion internationale dans de nombreuses villes de par le monde ! À Paris, pour la Rebellion Internationale d'Octobre, il y aura de nombreuses actions 🟡 Tous les détails par ici <a href="https://xtinctionrebellion.fr/evènements/2019-10">xtinctionrebellion.fr/evènements/2019-10</a></p> <p>#InternationalRebellion pic.twitter.com/29zAKTEBAs</p> <p>25 sept. 2019 09h08 Twitter Web Client</p>		<p><b>49</b> (RE)TWEETS</p>
<p><b>Julien Olivier</b> @JulienOlivier45</p> <p>🌿🌱 Adeptes de la #medecine par les plantes ? Découvrez le siphium, #plante de Cyrénaïque aux vertus fameuses dans l'Antiquité. 📖📖 Surexploitée, c'est aussi une des premières espèces végétales éteintes par l'homme ! <a href="https://antiquitebnf.hypotheses.org/8676">antiquitebnf.hypotheses.org/8676</a> #ecologie</p> <p>#ExtinctionRebellion pic.twitter.com/7Un7zxYJm2</p> <p>05 sept. 2019 10h18 Twitter Web App</p>		<p><b>46</b> (RE)TWEETS</p>
<p><b>On se réveille ou pas?</b> 🕒 @ReveilAttendu</p> <p>#Luxembourg Noir jaune vert de monde!!!! #OnEstLa chanté en coeur, les pois!!!! #Paris #ActeXLV #GiletsJaunes #ExtinctionRebellion pic.twitter.com/h2aKkCdAUl</p> <p>21 sept. 2019 14h01 Twitter Web App</p>		<p><b>46</b> (RE)TWEETS</p>
<p><b>Alternatiba ANV Rhône</b> @alternatiba69</p> <p>BREAKING ⚡ Dans la nuit du 25 au 26.09 des dizaines de panneaux publicitaires recouvrent les rues de la Métropole #Lyon la déclarant en "Etat d'urgence climatique" 🟡 Une action #ExtinctionRebellion #Lyon soutenue par @alternatiba69 @YouthClimateLY Lire <a href="https://facebook.com/alternatiba.rh">facebook.com/alternatiba.rh</a> pic.twitter.com/a5EnX1hSb</p> <p>27 sept. 2019 11h40 Twitter Web App</p>		<p><b>43</b> (RE)TWEETS</p>

On identifie via Gephi plus de 500 communautés, les 8 premières totalisant plus de la moitié des comptes actifs. La première communauté, en rose, est celle des comptes officiels du mouvement XR France avec 14 %. Elle est suivie avec 7 % environ chacune par les communautés vert et bleu clair. Cette dernière, autour de @LorentzMathias, correspond aux gilets jaunes.

Après qualifier sur le fichier d'export Gephi les comptes au centre des échanges qui ont contribué à diffuser l'information, nous pouvons identifier les leaders go-between actifs avant le 5 octobre (surlignés) :

**Cartographie via Gephi et Visibrain des comptes au centre des sur Extinction Rebellion en France du 1er septembre au 6 octobre 2019 – application d'un filtre pour garder les comptes ayant un score élevé**

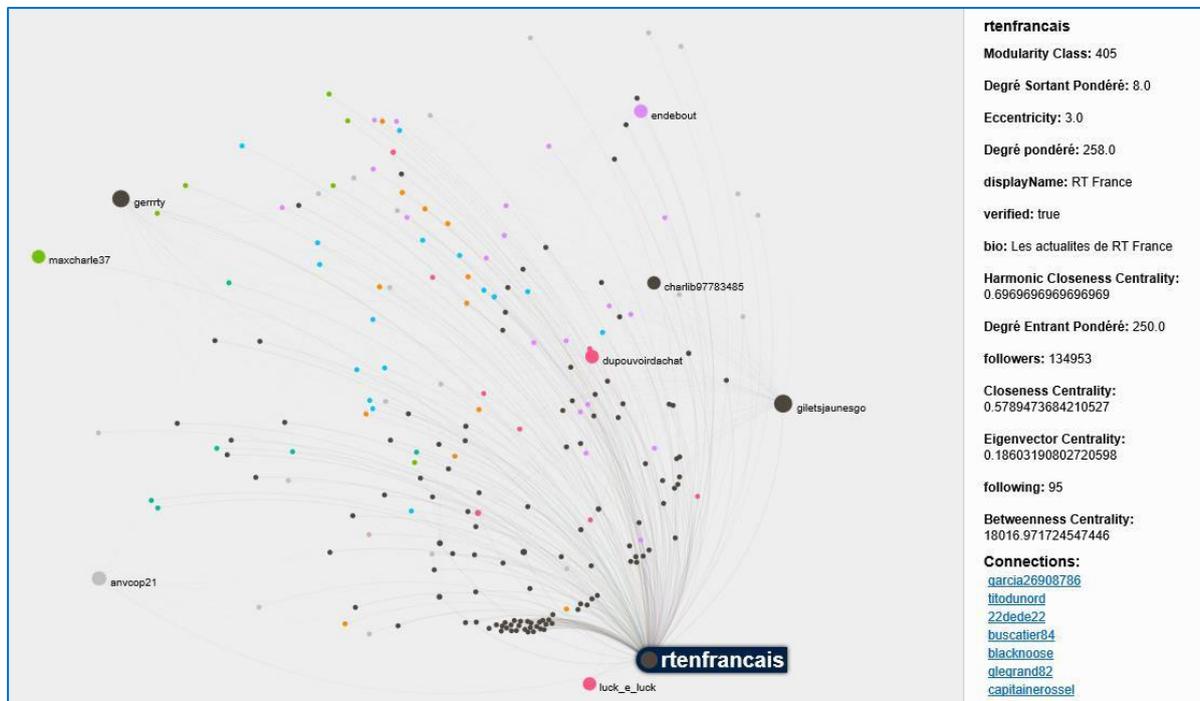


Sur les 42 leaders go-between identifiées comme étant critiques dans la propagation de l'information, nous en retrouvons 24 sur la cartographie des comptes les plus influents après le 6 décembre. Ceux-ci sont répartis en différentes catégories, spécifiques au mouvement Extinction Rebellion :

- Les acteurs XR francophones en dehors de la France (surlignés en violet), à l'écart des autres communautés, mais présents néanmoins car repris par les comptes XR hexagonaux, notamment les suisses ;
- Les émanations locales de Extinction Rebellion en France, surlignés en rouge ;
- Les militants XR / Gilets Jaunes, en vert ;
- Les journalistes et médias, en jaune.

On notera également la présence de Russia Today, moins influent avant sa couverture du 6 octobre, mais présent via une de ses journalistes. C'est le média le plus central dans les échanges après Reporterre (cité plusieurs fois par des militants XR mais pas actif avant le 6 octobre).

### Cartographie dynamique via Sigma - Gephi des échanges autour du compte @RTenfrançais sur Extinction Rebellion du 1er septembre au 6 octobre 2019



### **Des communautés locales qui se reprennent pour accélérer la viralité**

Selon la méthodologie du Pew Research Center, les échanges autour d'Extinction Rebellion du 1<sup>er</sup> septembre au 4 octobre 2019 correspondent à des groupes de communautés, influentes et qui échangent entre elles, complétés de groupes de marques.

Plusieurs personnes se revendiquant du mouvement des gilets jaunes, sans être des leaders historiques, ont tout de suite soutenues Extinction Rebellion. Des médias et journalistes intéressés aux questions sociales tels que Bastamag ont aussi repris le sujet, sans pour autant être très repris en dehors de leur communauté.

On note un fort activisme sur Twitter de comptes locaux repris par la communauté XR nationale. Extinction Rebellion a réussi à constituer en France en un laps de temps très court une dynamique sur Twitter, très probablement grâce à un fonctionnement horizontal favorisant l'animation de comptes sur les médias sociaux par des militants locaux.

## 5. Analyse des échanges Twitter sur Notre-Dame-des-Landes en janvier 2016

Cette analyse vise à identifier des acteurs spécifiques au sujet NDDL, qui ne seront pas forcément influents sur Twitter dans d'autres sujets. **L'analyse de la propagation des tweets via Gephi montre en effet que l'influence des acteurs traditionnels considérés comme tel : politiques, médias ... est toute relative (sociétés bottom up avec les NTIC).** Quand ils permettent de populariser une thématique, c'est souvent après qu'un important travail de diffusion des idées a déjà été effectué sur Twitter par des minorités actives.

**Nous allons ici analyser les échanges du 1<sup>er</sup> janvier 2016 au 10 janvier 2016 inclus, soit le lendemain de la première manifestation d'ampleur sur Notre-Dames-des-Landes.** Le volume des échanges Twitter va ainsi continuer de croître jusqu'à l'annonce du référendum le 12 février 2016 par François Hollande, avec une croissance continue des échanges Twitter, doublée de manifestations publiques ayant un succès de plus en plus important.

Notre-Dames-des-Landes est un cas très intéressant de basculement de l'opinion (hier homme politique et tribun, aujourd'hui leader réseaux sociaux) suite à un militantisme très actif d'un groupe minoritaire (logique réseau bouleverse la politique : une minorité structurée qui semble révéler à tous les enjeux globaux aux populations locales et citoyens (6x7) a désormais plus de pouvoir que la logique politique locale ou nationale avec une logique du secret : 10+10), qui finit par remporter son objectif malgré un référendum local perdu (minorités contre l'avis des populations locales à creuser).

D'enjeu de territoire, le débat a ainsi débordé sur la politique environnementale (**effet boule de neige d'indifférenciation et de déplacement du centre des préoccupations**), permettant d'obtenir **l'abandon d'un projet pourtant porté par les élus locaux depuis plusieurs décennies.** La récente décision d'abandonner le projet Europacity en est la suite logique, avec les mêmes enjeux : un projet très porté par les élus locaux, mais non soutenu par la population et donc politiquement risqué pour le gouvernement.

On notera que les élus locaux, pour Notre Dame des Landes comme pour Europacity, avaient mis en garde le gouvernement (rôle des méso-institutions : penser global et agir local) contre les risques d'abandon de ces projets, notamment en termes de rejet de la population locale.

### **La relance d'une dynamique de contestation sur Twitter après l'annonce de la reprise des travaux fin 2015**

Le projet d'aéroport du Grand Ouest visant à remplacer celui en place à Nantes par un aéroport plus grand à Notre-Dame-des-Landes a été lancé en 1963. **L'opposition, initiée en 1972, a mené à une bataille juridique et plusieurs recours, pendant que des opposants installaient une « Zone à Défendre » (ou ZAD) à Notre-Dames-des-Landes.** Les travaux bloqués entre 2012 et 2015, devaient être repris début 2016.

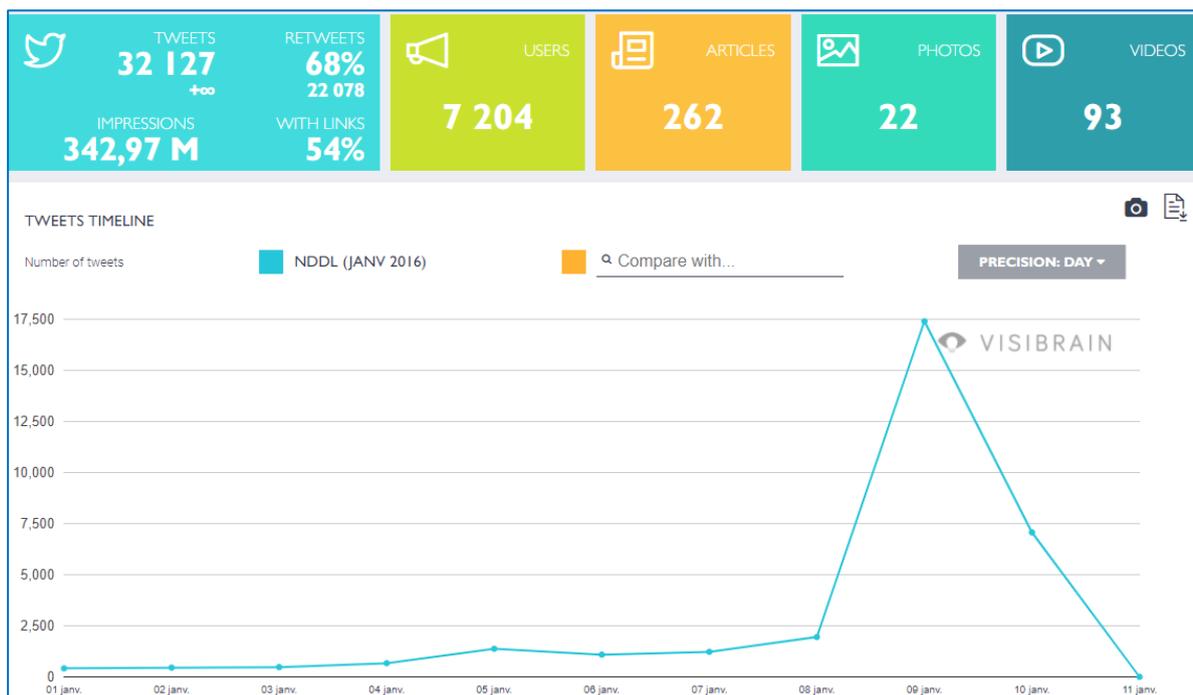
Si le sujet est très suivi par la presse locale et repris ponctuellement par la presse nationale, notamment avec l'essor de nouvelles « ZAD » ailleurs en France, **les échanges sur Twitter sont très peu nombreux avant l'annonce d'octobre 2015. Sur les mois précédents de l'année, il y a en moyenne 6000 tweets sur le sujet, avec quelques pics lors des nouvelles importantes** (18 000 tweets en juillet 2015 suite au rejet des recours des opposants, 14 000 tweets en décembre 2014 avec plusieurs articles sur les « zadistes ») et un énorme pic de 70 000 tweets à l'occasion de la manifestation du 22 février 2014 contre le projet à Nantes.

Ce qui va changer en 2016, c'est une relance des manifestations en prévision de l'annonce sur la reprise des travaux, qui va accélérer la dynamique Twitter jusqu'au référendum en juin 2016, avec une moyenne de 70 000 tweets par mois sur le premier semestre 2016, dont 130 000 entre la première manifestation de l'année le 9 janvier et l'annonce du référendum le 12 février.

Nous allons analyser et qualifier les acteurs Twitter les plus influents sur les échanges ayant mené à la première manifestation le 9 janvier 2016.

## Analyse des échanges Twitter entre le 1<sup>er</sup> janvier 2016 et le 10 janvier 2016

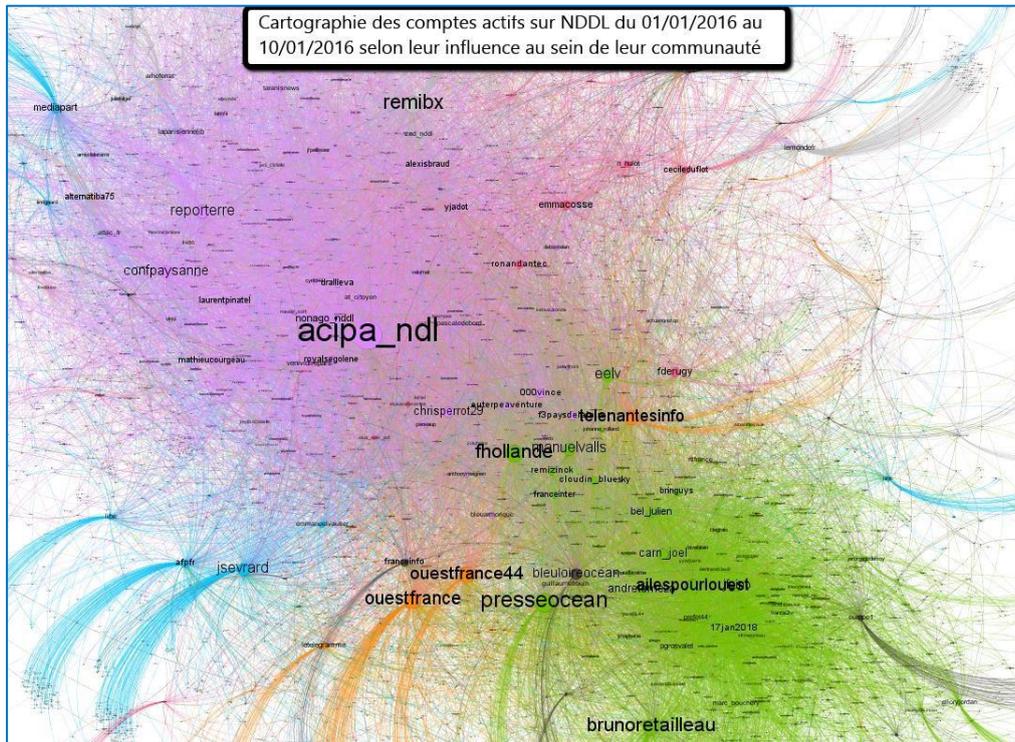
### Timeline via Visibrain des mentions Twitter sur Notre-Dame-des-Landes du 1er au 10 janvier 2016 :



Du 1<sup>er</sup> janvier au 10 janvier 2016, lendemain de la manifestation qui s'est tenue à Nantes, 32 127 tweets ont mentionné le sujet. On identifie via Gephi près de 1 000 communautés regroupant les comptes Twitter actifs sur ce sujet, les 8 plus importantes (leaders actifs) correspondant à 75 % des comptes actifs (les 25 % restants étant donc très peu visibles). La première communauté, à laquelle correspondent environ 20 % des comptes actifs, est celle des opposants à l'aéroport. Les défenseurs de l'aéroport arrivent en seconde position avec environ 15 % des comptes Twitter actifs sur le sujet.

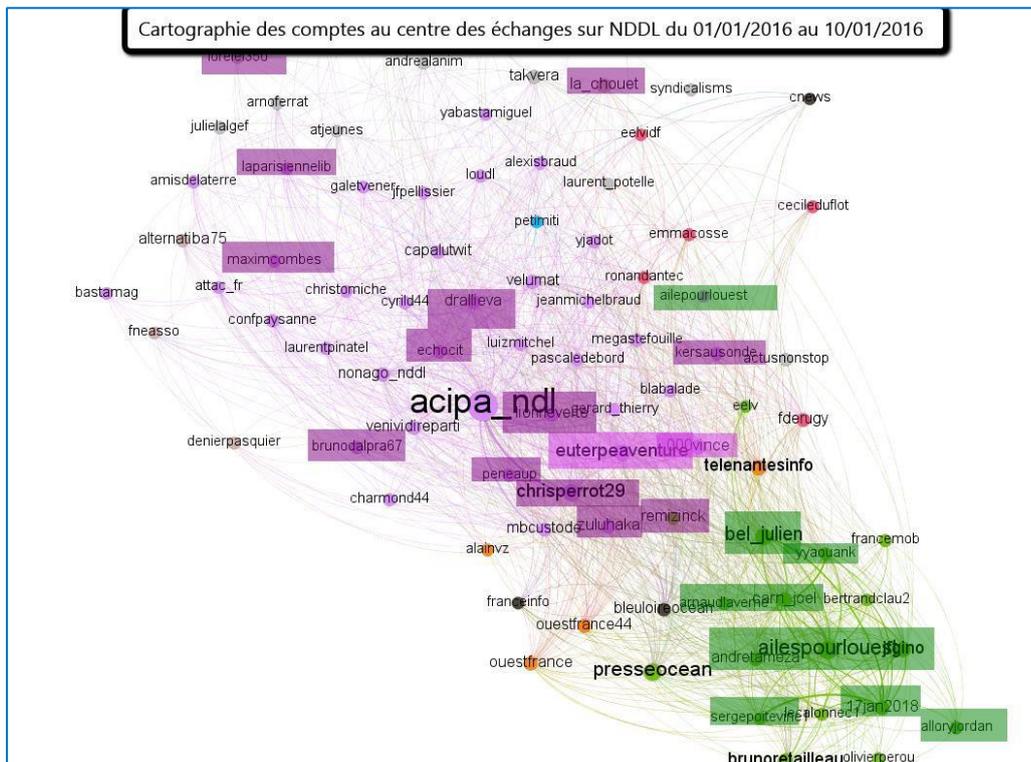
La cartographie des comptes les plus influents au sein de leur communauté montre une répartition assez bien distribuée des comptes Twitter entre les opposants à l'aéroport (en rose), et ses défenseurs (en vert). **Les médias locaux sont entre les deux communautés en orange, tandis que les médias nationaux (AFP ou Mediapart) sont à l'écart des échanges.** On notera que les membres du gouvernement et le compte national d'EELV sont reliés aux pros aéroport. Si les responsables verts se sont globalement opposés à l'aéroport sur cette période en diffusant notamment un appel de Nicolas Hulot à abandonner le projet, ils ne sont que peu repris par la communauté des opposants avec qui ils interagissent peu.

**Cartographie via Gephi et Visibrain des comptes les plus mentionnés par leurs communautés sur Notre-Dame-des-Landes du 1er au 10 janvier 2016 :**



**Identification des leaders go-between**

**Cartographie via Gephi et Visibrain des comptes au centre des échanges sur Notre-Dame-des-Landes du 1er au 10 janvier 2016 – application d'un filtre pour garder les comptes ayant un score élevé**



Si les scores les plus importants correspondent aux promoteurs de l'aéroport, **on retrouve autour du compte officiel des opposants @ACIPA\_NDL plusieurs utilisateurs au centre des échanges**, qui sont au final plus nombreux que les défenseurs de l'aéroport.

On notera que la moitié des comptes au centre des échanges n'étaient pas actifs sur NDDL entre le 1<sup>er</sup> janvier et le 8 janvier 2016 : **nous avons surligné ceux qui avaient contribué aux échanges sur cette période et dont les scores de betweenness centrality sont également élevés entre le 01/01/2016 et le 08/01/2016**, en excluant les comptes déjà très influents (@ACIPA\_NDL donc, mais également @BrunoRetailleau, les principaux médias et élus nationaux).

### **Une confrontation entre deux communautés opposées structurées autour de leaders locaux et nationaux**

Selon la méthodologie du Pew Research Center, **ce corpus de données correspond à une foule polarisée, avec deux communautés « leaders » représentant le tiers des comptes actifs et étant opposées sur le sujet. Chaque communauté est très structurée autour de ses leaders**, et reprend les mêmes messages et mots d'ordres : mise en avant des casseurs et du désordre chez les supporters de l'aéroport, rappel des enjeux de la COP 21 de l'autre. Le gouvernement semble incapable de se positionner, ce qui renforce les opposants à l'aéroport.

Les responsables politiques actifs dans chacune des communautés exploitent des arguments contradictoires à partir de données similaires : défense de la nécessité d'un nouvel aéroport d'un côté, proposition d'agrandir la piste existante de l'autre. **Sans utiliser de fausses informations, chacun exploitera donc les articles et rapports selon sa position : ce ne sont pas les « faits » qui amènent un positionnement, mais plutôt l'idéologie (1990 a été présenté comme la fin de l'histoire, la fin des idéologies mais c'est au contraire leur renouveau) qui modifie l'interprétation des faits.**

## 6. Analyse des échanges Twitter sur Europacity en octobre 2019

Les échanges sur twitters sur Europacity sont un bon exemple de viralité issue de différentes communautés aux objectifs opposés et aux pouvoirs différenciés. On retrouve des oppositions entre les grands groupes commerciaux et les citoyens, les partisans de la grande consommation et l'écologie, les acteurs individuels connus, les locaux et les acteurs nationaux.

### Analyse des tweets sur Europacity

L'analyse est effectuée du 1<sup>er</sup> septembre 2019 au 9 octobre 2019, juste après la dernière grosse manifestation du week-end du 5-6 octobre 2019. Sur cette période, il y a eu 11 160 mentions twitter, dont 82 % de retweets, par 5 252 utilisateurs.

### Timeline via Visibrain des mentions Twitter sur Europacity du 1er septembre au 9 octobre 2019 :

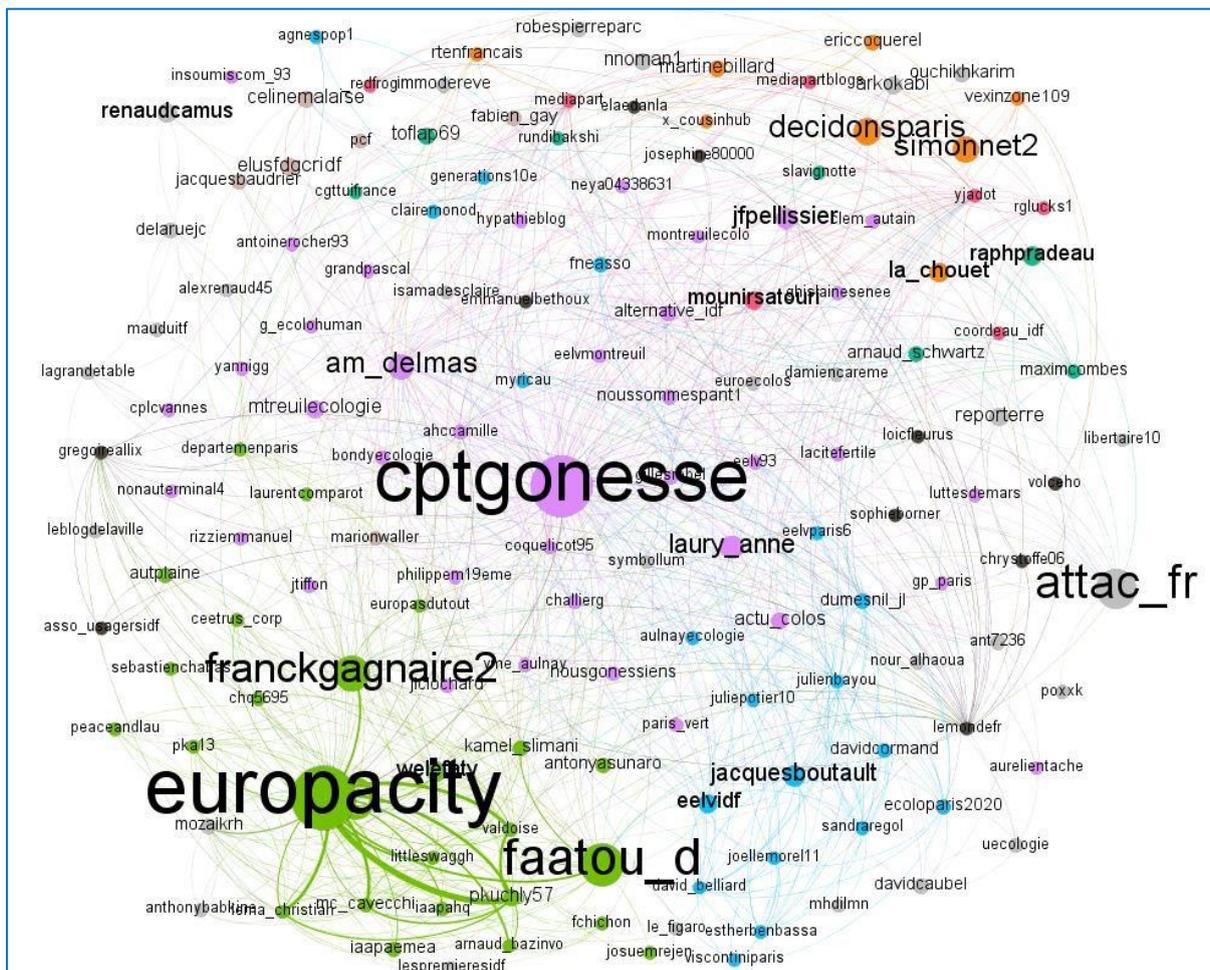


On identifie avec Gephi 285 communautés, les 8 premières totalisant plus de 70 % des comptes Twitter, dont 17 % pour la première (en rose).





**Cartographie via Gephi et Visibrain des comptes au centre des échanges sur Europacity du 1er septembre au 9 octobre 2019 – application d’un filtre pour garder les comptes ayant un score élevé**



On constate que les deux comptes opposés à ce projet, @EuropaCity et @CPTGonesse (auparavant @nonaeuropacity1) ont une influence relativement similaire au niveau local. Pour autant, les opposants sont les mieux connectés au niveau national, ce qui fait la différence dans la bataille d’influence.

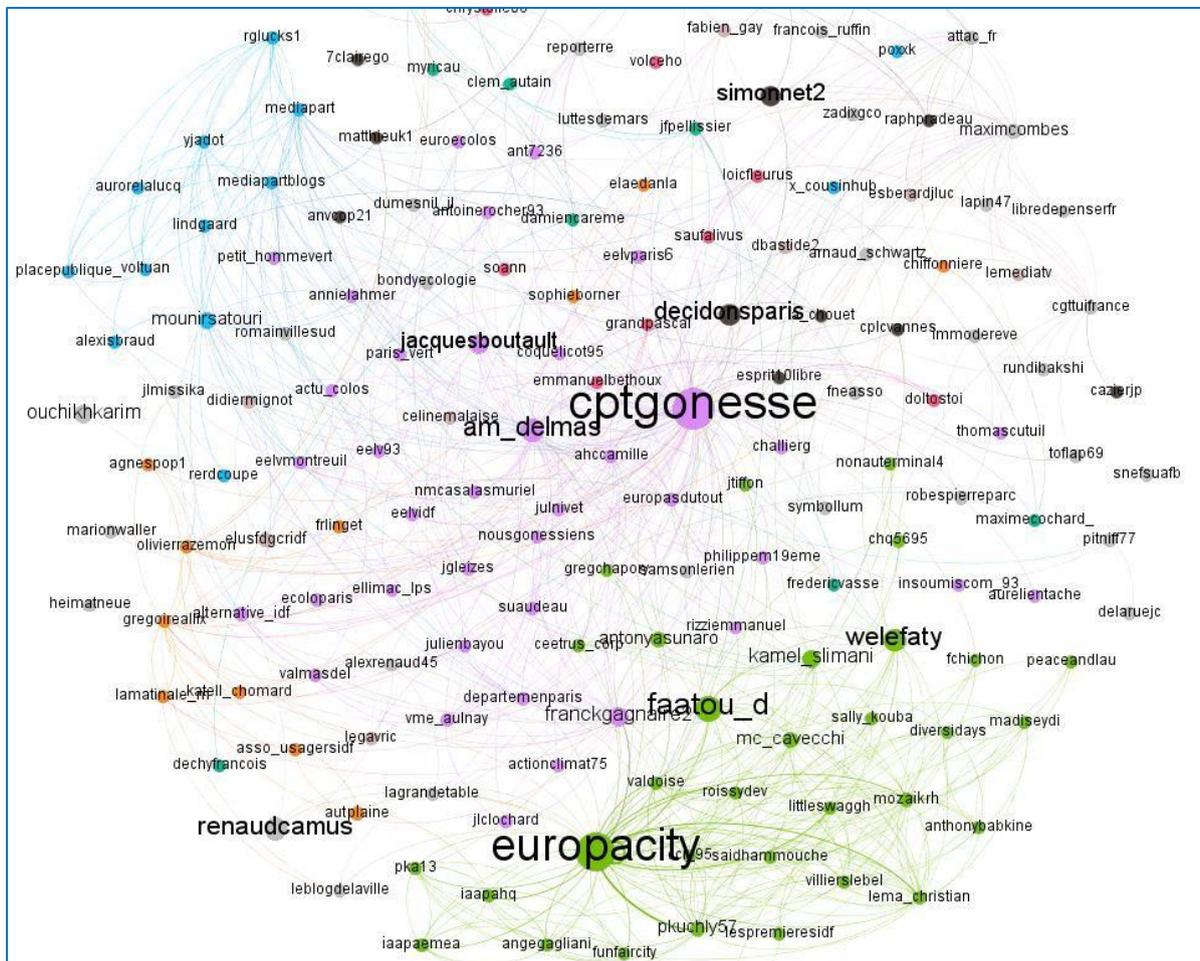
A noter que la manifestation du 5 octobre coïncide avec le premier gros évènement d’Extinction Rebellion, qui a éclipsé les autres sujets sur les réseaux sociaux. L’importance des acteurs locaux pour Europacity est donc également liée à un manque d’intérêt des influenceurs nationaux à ce moment précis.

**Identifications des leaders go-between avant la viralité**

Du 1<sup>er</sup> septembre 2019 au 4 octobre 2019, on obtient 6733 tweets, dont 84 % de retweets, par 3727 comptes Twitter. On identifie avec Gephi 155 communautés, dont les 8 premières totalisent plus de 70 % des comptes Twitter.

Les comptes au centre des échanges restent ceux des communautés locales impliquées dans le conflit. Elles restent les plus influentes sur les conversations concernant Europacity.

### Cartographie via Gephi et Visibrain des comptes au centre des échanges sur Europacity du 1er septembre au 9 octobre 2019 – application d'un filtre pour garder les comptes ayant un score élevé

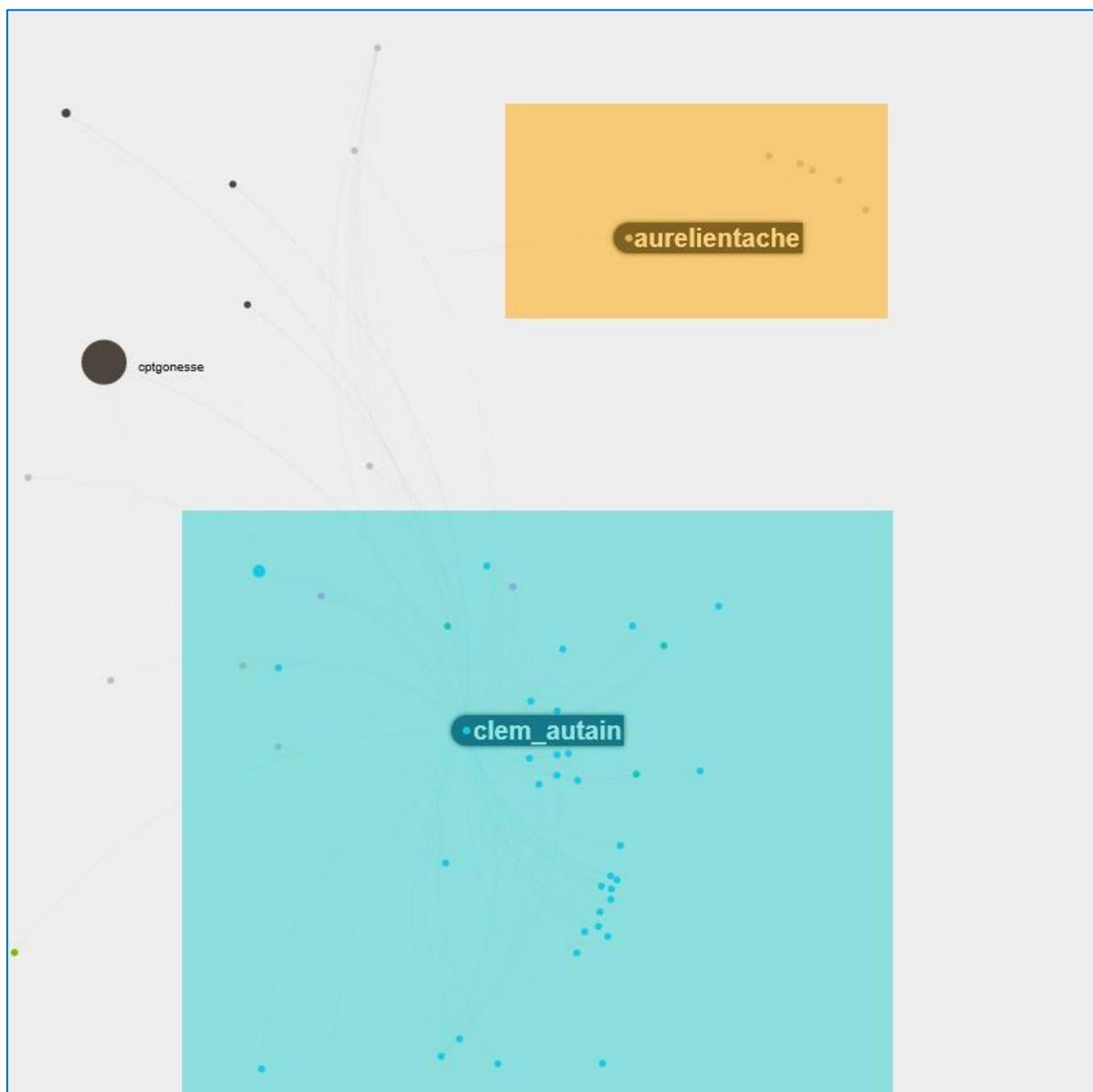


Les deux communautés vertes (pro Europacity) et rose (anti Europacity) ont un poids similaire, autour de 20 % des comptes les plus centraux.

### **Une confrontation entre deux communautés locales qui ignorent les leaders nationaux**

Une fois filtrés les « groupes de marques » liés à des acteurs nationaux, nous pouvons considérer que ce sujet est caractéristique d'une **polarisation entre deux camps selon la méthodologie du Pew Research Center**, de la même façon **que Notre-Dame-Des-Landes**, avec plusieurs comptes **politisés locaux** (mais **peu d'élus défendant le projet : Marie-Christine Cavecchi, présidente du Val d'Oise**, a ainsi surtout réagi après l'abandon du projet en novembre 2019). Seul **Aurélien Taché, député de la majorité**, a défendu le projet, mais il est bien moins visible que l'opposition (par exemple la députée FI Clémentine Autain).

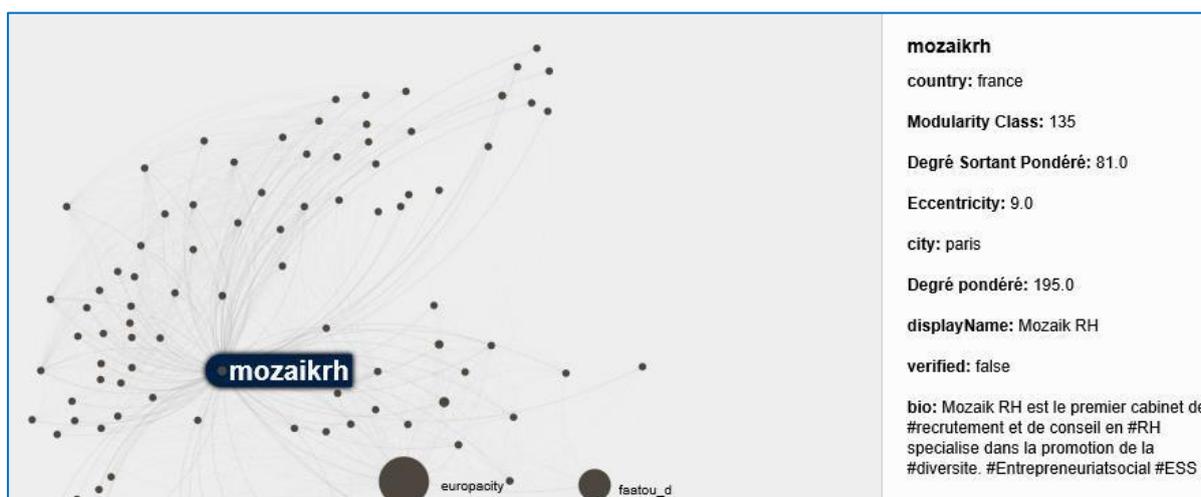
**Cartographie dynamique via Sigma - Gephi des échanges autour des comptes Twitter des députés Aurélien Taché et Clémentine Autain sur Europacity du 1er septembre au 9 octobre 2019**



Les leaders nationaux historiques, comme José Bové, sont peu visibles également car repris uniquement par leur communauté nationale – et pas des acteurs locaux. **L'analyse de la propagation des tweets montre en effet que l'influence des leaders traditionnels considérés comme tel au niveau politique (José Bové) est toute relative (Kotter, 2007, 2011, Drucker, 2011). Cela traduit le basculement d'une société top down (leader hiérarchique et/ou financier) à une société bottom up (leaders locaux et/ou stratégique) qui est propre à une société complexe (Foray & Lundvall, 1996, Muldoon, 2013).**

On notera l'intervention de l'extrême droite via Renaud Camus, et la présence de plusieurs comptes peu influents / mentionnés mais qui obtiennent un bon score de centralité car ils se reprennent entre eux pour défendre le projet Europacity. Les comptes d'acteurs visant à promouvoir les opportunités du projet, ici Mozaïk RH, qui peinent à dépasser la communauté des soutiens au projet.

**Cartographie dynamique via Sigma - Gephi des échanges autour du compte Twitter @mozaikrh sur Europacity du 1er septembre au 9 octobre 2019**



En réaction, on retrouve beaucoup d'acteurs locaux opposés au projet tel que le Collectif de commerçants EuropasDuTout ou encore l'association des usagers des transports d'IDF. Twitter est utilisé concrètement comme outil d'influence pour défendre deux visions radicalement opposées du territoire et de son avenir, et logiquement irréconciliables. La mise en évidence des leaders go-between qui étaient actifs avant la manifestation montre que la majorité d'entre eux étaient en effet déjà présents.



souvent local) d'avoir accès à une quantité plus élevée d'information (Bouba Olga & Grosseti, 2008, Baulant, 2015). Les liens faibles ont ainsi la possibilité d'avoir un très grand pouvoir d'influence dans la mesure où ils sont en relations avec deux (ou plusieurs) communautés. Cette analyse se rapproche de celle développée par Burt (2009) qui étudiait les « trous structurels entre les acteurs n'ayant aucune relation directe entre deux contacts d'un même acteur. Le leader go-between est capable de relier deux communautés isolées l'une de l'autre (Karaszewski & Lis, 2013, Baulant, 2017). Ainsi, par l'intermédiaire du leader go-between, les messages des acteurs locaux (élus, citoyens activistes écologiques) peuvent avoir une retombée nationale ou à l'inverse un leader écologique peut influencer le débat local. Dans les deux cas, le leader go between se trouve dans une position stratégique puisqu'il peut diffuser ou bloquer l'information de l'une des communautés.

Sur le territoire, l'influence des opposants dépasse ainsi les soutiens au projet Europacity . La « moindre médiatisation » du projet permet de façon un peu paradoxale aux opposants d'être relativement « plus fort » et de peser sur les solutions finales. Le combat des tweets est resté au niveau local, entre les élus locaux de la majorité (comme Aurélien Taché) et les opposants au projet (comme d'autres élus locaux, les écologistes, les citoyens).

Au niveau national, cette opposition rappelle constamment au gouvernement que ce projet risque de **s'ancrer durablement dans l'opinion**. Le phénomène de viralité, issu du développement des réseaux sociaux, permet également de faire apparaître une propriété intéressante, énoncée par les grandes analyses de la complexité (Koestler, 1988, Massé, 2000, Le Moigne et Morin, 2001, Damasio, 2010, Picq, 2012). Les opposants locaux au projet Europacity, souvent éparpillés et peu organisés, ont la possibilité désormais d'avoir un pouvoir qui est plus fort que celui des élus politiques (locaux ou nationaux) car ils en « touchent » les internautes par des mots clefs bien choisis (comme « *luttons contre la surconsommation* », « *stop aux dégâts écologiques* ») qui font sens pour les internautes non locaux (Goleman, 2013, Langer, 2014). Dans un monde devenu globalisé, le retour des petits acteurs stratégiques, comme les leaders go-between, à des moments où les élus politiques ne l'attendent pas est le propre des sociétés complexes où nous vivons depuis les années 2000. Le rôle du leader go-between a aussi été développé en sociologie par des auteurs comme Crozier (1977) qui a proposé une approche globale des conflits dans un réseau dans leur complexité en intégrant les aspects psychologiques qui sont propres à chaque acteur concerné. L'analyse du leader go-between a aussi été utilisée pour structurer des réseaux territorialisés, comme les clusters ou les pôles de compétitivité (Cohendet et al, 2000, Torre et Rallet, 2005, Amisse et al, 2008).

**Dans l'exemple d'Europacity, on observe ainsi que le pouvoir des échanges de tweets sur les réseaux sociaux issus de population locale, qui peut avoir un effet plus viral et des conséquences plus fortes que les discours politiques des élus locaux et nationaux. Dans le cas europacity, comme dans le cas de NDDL, le pouvoir des opposants locaux est structuré par les « leaders go-between locaux » (Muller, Pénin, 2008) ayant à la fois, une bonne connaissance du tissu local et des moyens modernes de façon à toucher l'opinion publique en dehors de la zone concernée (viralité),**

**ce qui donne à ces leader locaux un « temps d'avance » sur la communication souvent trop écrite et hiérarchisée des élus politiques nationaux (Collins, 2011, Goffee and Jones, 2011).**

Sur la période de septembre 2019, on note donc, conséquence logique de ce climat, un grand nombre de comptes Twitter au centre des échanges. Cela signifie que pour observer la viralité d'un sujet local, la détection de l'implication d'acteurs locaux, même peu suivis, est essentielle. A contrario, les leaders nationaux ont peu d'impact sur la diffusion de l'information et ne font que suivre les leaders locaux.

Sur des projets de sociétés comme Europacity (grand projet commercial avec 31 M de visiteurs étaient attendus), il est logique que les oppositions locales soient fortes entre différents acteurs ayant des vues controversées sur le projet.

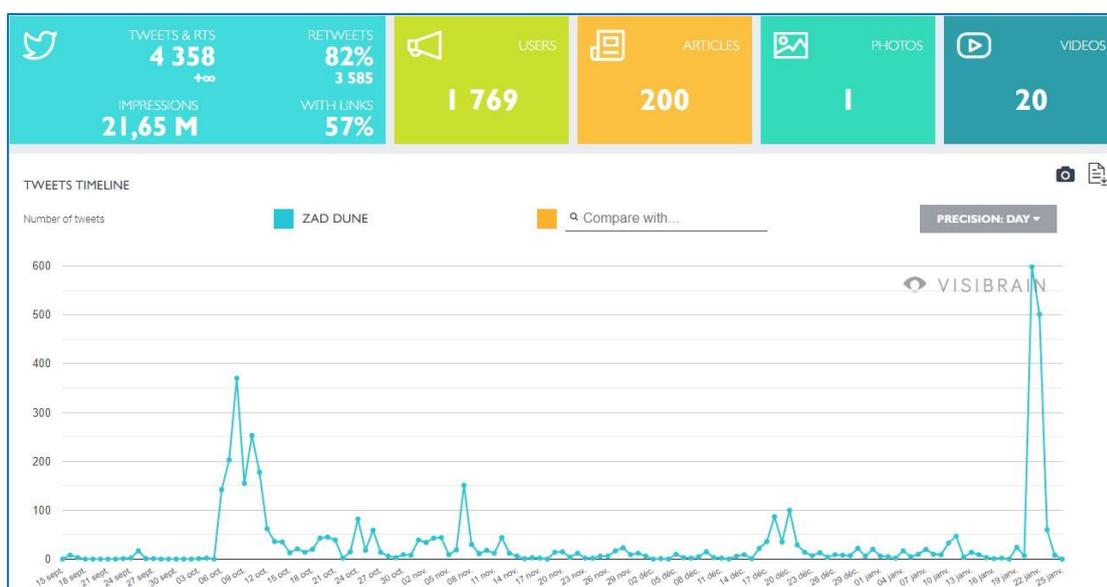
Dans ce cas, ce type d'opposition est difficile à trancher puisqu'il s'agit de mettre en valeur un département « défavorisé » (le Val-d'Oise) avec l'appui des acteurs locaux publics mais contre l'appui des citoyens et des écologistes eux aussi locaux. Cet exemple de tweets au sujet d'un projet local controversé montre bien que, **dans une société Bottom up ou horizontale, les approches verticales et dirigistes, même locales, ne sont plus adaptées pour convaincre des citoyens ou des élus désireux de promouvoir au contraire un développement économique et social plus soutenable et harmonieux** (Garvin & Roberto, 2013).

## 7. Analyse des échanges Twitter sur la ZAD Dune en janvier 2020

### Analyse des tweets sur la ZAD Dune

La « ZAD Dune » correspond à la contestation née en septembre 2019 après la validation par le préfet de Vendée du projet de port de plaisance de Brétignolles-sur-Mer, qui divise les habitants depuis 16 ans. C'est un enjeu local, qui a suscité 4358 mentions Twitter du 16 septembre 2019 au 25 janvier 2020 par 1769 comptes (82 % de retweets), quelques jours après l'opération d'expulsion de la ZAD le 21 janvier 2020.

### Timeline via Visibrain des mentions Twitter sur la ZAD Dune du 16 septembre 2019 au 25 janvier 2020 :



Ce sujet a surtout impliqué des militants locaux, même si la ZAD a été soutenue par [Extinction Rebellion](#) (via un des leaders du mouvement sur Twitter, cf analyse dédiée) :

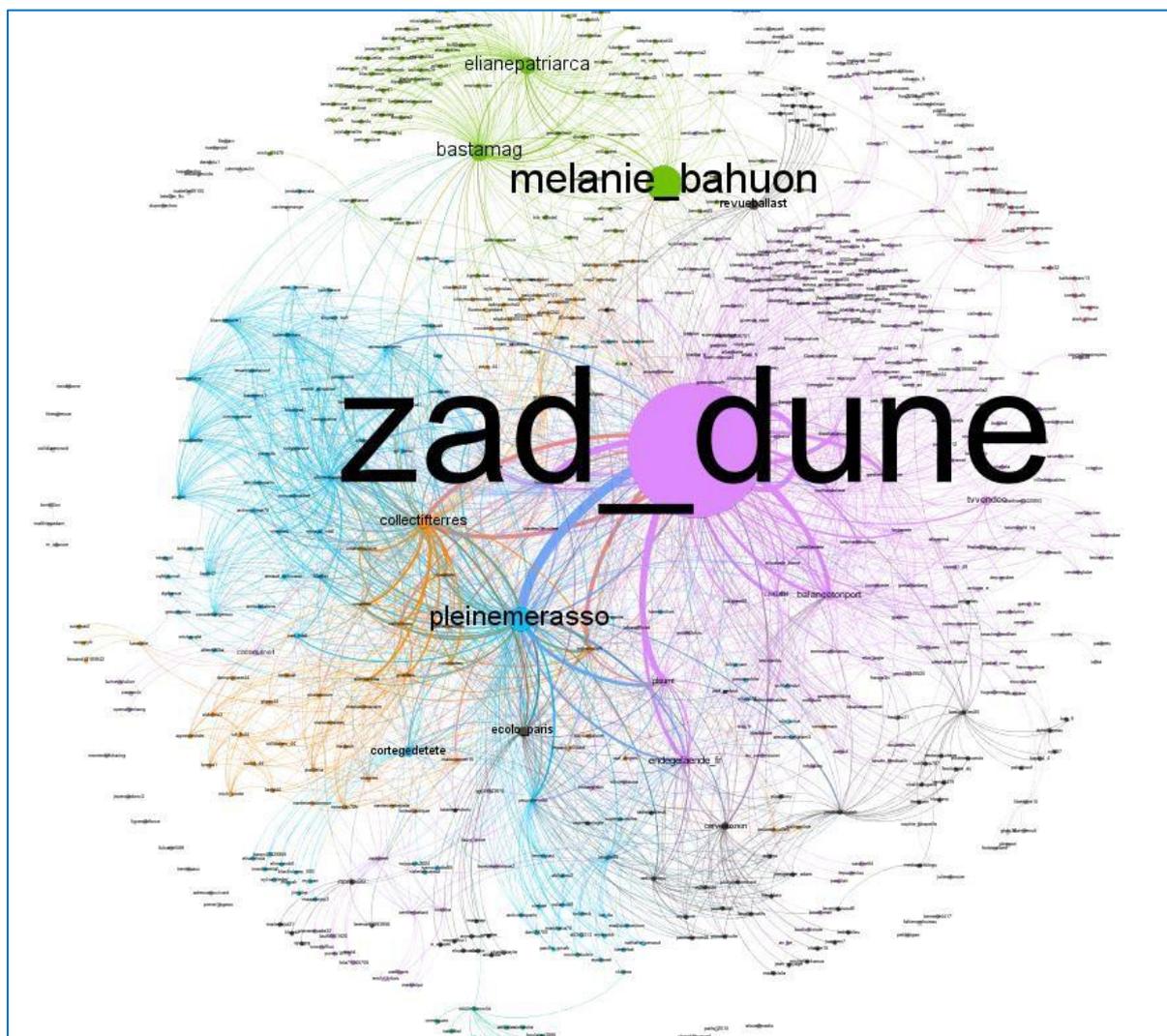
## Classement via Visibrain des Tweets les plus repris sur la ZAD Dune du 16 septembre 2019 au 25 janvier 2020 :

	<p><b>Nantes Révolte</b> @Nantes_Revoltee</p> <p>Pendant que les caméras sont braquées sur le Spectacle vert à Paris, la ZAD de Brétignolles, en Vendée, met en échec, depuis dimanche les Bulldozers qui veulent détruire une plage, une dune, et un réservoir d'eau. Bravo ! <a href="http://ouest-france.fr/pays-de-la-loi...">ouest-france.fr/pays-de-la-loi...</a></p> <p>08 oct. 2019 14h10 Twitter Web App</p>		<b>345</b> (RE)TWEETS
	<p><b>Peuple Révolté</b> @PeupleRevolte</p> <p>⚠️ Urgent une #ZAD ce met en place à #Brétignolles (85) pour lutter contre le projet de Port. 📦 Besoin de matériel(bois, outils,...), Nourriture, produits d'entretien et de renforts sur place #NonAuPort #BalanceTonPort #Vendee #ecologie #EcologieRadicate Crédits @MarionLaporte pic.twitter.com/4eevZ9DNHS</p> <p>06 oct. 2019 17h25 Twitter for Android</p>		<b>148</b> (RE)TWEETS
	<p><b>ZAD de la Dune</b> @zad_dune</p> <p>🚨 ALERTE EVACUATION EN COURS 🚨 #Brétignolles est vérouillé, un hélico survole la colo PTT, Gendarmerie mobile et une pelleuse sont en train d'évacuer des bâtiments ... VIDES 😞 suite à des informations parvenues hier soir nous étions prévenus et avons pris les devants [1/3]</p> <p>21 Janv. 2020 07h24 Twitter for Android</p>		<b>136</b> (RE)TWEETS
	<p><b>Nantes Révolte</b> @Nantes_Revoltee</p> <p>"Rendez-vous dès maintenant à la ZAD de la Dune, à Brétignolles, en Vendée, au lieu dit La Normandelière. Besoins de renfort et soutien physique, de photographie de vivres et de medics." <a href="https://nantes-revoltee.com/%f0%9f%94%b4-v...">nantes-revoltee.com/%f0%9f%94%b4-v...</a></p> <p>22 Janv. 2020 12h17 Twitter Web App</p>		<b>114</b> (RE)TWEETS
	<p><b>Street Zad Actionmédic</b> @ZadStreet</p> <p>📢 Appel à soutien 📢 L'expulsion de la #ZAD de la Dune @zad_dune à #brétignolles sur mer en Vendée, près des Sables-d'Olonne, a commencé ce mardi matin avec hélico &amp; tout le dispositif habituel des GM qui ne laisse rien présager de bon pour demain. Story 📱 <a href="https://t.co/TXuPY30APe">https://t.co/TXuPY30APe</a></p> <p>21 Janv. 2020 19h22 Twitter Web App</p>		<b>84</b> (RE)TWEETS
	<p><b>Pleine Mer</b> @PleineMerAsso</p> <p>📢 ZAD DE LA DUNE 📢 70 personnes sont venues en soutien ce matin. C'est super mais les personnes sur place ont besoin d'encore plus de monde, et de nourriture. Pas d'expulsion ce matin, mais vigilance : l'arrêt préfectoral a transformé la Normandelière en zone militarisée. <a href="https://pic.twitter.com/aEhixjgQc">pic.twitter.com/aEhixjgQc</a></p> <p>22 Janv. 2020 12h30 Twitter for Android</p>		<b>71</b> (RE)TWEETS

On identifie avec Gephi 95 communautés sur cette période, et les 3 premières communautés totalisent presque 50 des comptes Twitter. Si la communauté rose, autour du [compte dédié à la ZAD](#) est directement liée au sujet, les deux suivantes, en vert et bleu clair correspondent à des militants nantais qui se sont intéressés à cette action mais sont éloignés du cœur des échanges.



**Cartographie via Gephi et Visibrain des comptes au centre des échanges sur la ZAD Dune du 16 septembre 2019 au 25 janvier 2020 – application d’un filtre pour garder les comptes ayant un score élevé**



25 comptes twitter sont identifiés comme ayant contribué à la viralité des échanges, en enlevant les comptes très populaires. 14 d’entre eux sont des comptes Twitter militants, montrant une forte implication de mouvements locaux sur l’environnement. Par ailleurs, plusieurs associations / acteurs locaux sont également actifs spécifiquement sur le sujet. Quelques médias militants sont également très mentionnés et repris dans plusieurs communautés, à la différence du média local TV Vendée.

De la même façon que pour Europacity, les échanges sont très localisés et démontrent une forte capacité de mobilisation et de diffusion de l’information entre différentes communautés engagées. Très peu de personnalités / politiques nationaux obtiennent de la visibilité, et s’ils sont présents ils ne sont pas au centre des échanges.

## 8. Analyse des échanges Twitter sur Extinction Rebellion et Lafarge en février 2020

Il s'agit d'une action d'un collectif militant à portée nationale, qui a été réalisée en soutien à des riverains, eux-mêmes en contact avec des élus locaux dans le cadre des municipales. Le groupe Extinction Rébellion (XR) reproche ainsi les impacts environnementaux causés par les installations parisiennes du groupe Lafarge. Pourtant les installations de ce groupe respecteraient strictement les normes environnementales et n'entraîneraient pas de pollution dans la Seine. Nous avons ici deux discours de l'écologie. Le premier discours porté par les ONG, les citoyens et les élus locaux sur leur ressenti concret et local. Le second discours porté par Lafarge s'appuie sur le respect strict des normes ISO environnementales et se veut rassurant au niveau écologique et rationnel au niveau économique (emploi, croissance). Le fait de respecter les normes environnementales engendre par ailleurs un classement en termes de RSE important pour les autorisations de travaux et les financements.

Dans une société transversale portée par le paradoxe de la mondialisation (penser global et agir local) ou de nouvelle forme de proximité et d'acceptabilité des locaux apparaissent (Muldoon, 2013, Foray and B. Å. Lundvall, 1996, Muldoon, 2013,), le ressenti de la population et des élus locaux comptent plus que des règles comptables. Nous sommes passés dans la vision d'un « territoire aimant » (Markusen, 1996, Cohendet et al, 2000) où les acteurs locaux sont capables de co-construire ensemble leur territoire et créent ainsi une « proximité organisée » (Rallet et Torre, 2004,) leur permettant d'avoir un pouvoir supérieur à la simple « proximité géographique » et de lutter pieds à pieds contre la « proximité cognitive » des grands groupes mondiaux (Baulant, 2015). Un niveau local bien structuré (capacité du leader local à faire interagir ses membres) remet donc en question la notion même de seuil des normes autorisées (Berney, 1928, Baulant 2017). Il y a donc ici à la fois un combat de « mot » et « d'images » (« nous respectons les normes » selon Lafarge) et un combat de fonds (comment définir un seuil de normes qui réponde aux besoins des acteurs locaux) (Cialdini, 2004, Goleman, 2011).

### Analyse des tweets

**Comme pour NDL, la controverse entre pro-économie et pro-écologie est ancienne et revient régulièrement à chaque chantier. Mais les réseaux sociaux ont donné un pouvoir croissant aux opposants locaux face aux constructeurs accusés de détruire l'environnement (Massé et Thibaut, 2000, de Rosnay, 2012, Sadin, 2017).** Pour l'opposition aux travaux du groupe Lafarge, la problématique est ancienne : des riverains du quai de Javel à Paris se sont mobilisés contre un projet d'extension de l'entreprise Lafarge suite à une enquête publique en 2017 : <https://portdejavelriverains.wordpress.com/lafarge/>

On retrouve un total de 7095 mentions sur la période, la majorité correspondant à l'action d'XR menée le 17 février. Il y a 93 % de retweets, et 3463 comptes actifs sur le sujet.

## Classement via Visibrain des Tweets les plus repris sur l'action XR Lafarge du 26 janvier au 19 février 2020 :

	<p><b>Je suis Camille</b> @CamilleVernier1            Ces militants écologistes qui exigent plus d'énergies renouvelables, notamment des éoliennes, et qui, aujourd'hui, bloquent une usine Lafarge, géant du béton, ne doivent pas être au courant qu'il faut beaucoup, mais beaucoup de béton pour installer ne serait-ce qu'une éolienne. <a href="https://pic.twitter.com/NNXxjUu2lc">pic.twitter.com/NNXxjUu2lc</a>  <small>17 févr. 2020 15h41 Twitter for Android</small></p>		<p><b>934</b> (RE)TWEETS</p>
	<p><b>Franck XR</b> @XrFranck            ✨ action en cours ! 🚧 #FinDeChantiers 🚧 Nous bloquons maintenant un 2e site de Lafarge. À notre arrivée Lafarge vidange ses cuves... Et devinez où part le mélange d'adjuvants ? 🇫🇷 🇫🇷 🇫🇷 @xrFrance #ExtinctionRebellion <a href="https://pic.twitter.com/ULyLMeFutc">pic.twitter.com/ULyLMeFutc</a>  <small>17 févr. 2020 14h30 Twitter for iPhone</small></p>		<p><b>343</b> (RE)TWEETS</p>
	<p><b>Là-bas si j'y suis</b> @LabasOfficiel            Depuis ce matin, des militants @xrFrance occupent le site Cemex de #Lafarge à Paris, dans le cadre de l'action #FinDeChantiers dénonçant l'industrie du BTP "deuxième industrie la plus polluante". Site bloqué et prises de parole en cours. images : @T_Bouhafs <a href="https://pic.twitter.com/2ks6BEqWNe">pic.twitter.com/2ks6BEqWNe</a>  <small>17 févr. 2020 10h04 Twitter Web App</small></p>		<p><b>300</b> (RE)TWEETS</p>
	<p><b>Remy Buisine</b> @RemyBuisine            Blocage en cours d'une usine de Lafarge à Paris pour dénoncer l'industrie du BTP et l'impact de la bétonisation sur l'environnement. #FinDeChantiers @xrFrance <a href="https://pic.twitter.com/Asx57Cb3Rj">pic.twitter.com/Asx57Cb3Rj</a>  <small>17 févr. 2020 09h40 Twitter for iPhone</small></p>		<p><b>287</b> (RE)TWEETS</p>
	<p><b>Amar Taoualit</b> @TaoualitAmar            Le blocage de l'usine Lafarge est toujours en cours !! #ExtinctionRebellion #GREVE #FinDeChantiers #XRFrance <a href="https://pic.twitter.com/RQdSoXHlDs">pic.twitter.com/RQdSoXHlDs</a>  <small>17 févr. 2020 11h00 Twitter for iPhone</small></p>		<p><b>172</b> (RE)TWEETS</p>
	<p><b>Charles Baudry</b> @CharlesBaudry            🚧 Blocage en cours du site Cemex de Lafarge à Issy les Moulineaux par Extinction Rebellion (@xrFrance) pour une action « Fin de Chantiers » afin de dénoncer l'industrie de la construction sur l'environnement. #FinDeChantiers #Paris #ExtinctionRebellion #Cemex <a href="https://pic.twitter.com/jgseQ8GK7">pic.twitter.com/jgseQ8GK7</a>  <small>17 févr. 2020 08h25 Twitter for iPhone</small></p>		<p><b>162</b> (RE)TWEETS</p>

On identifie via Gephi 75 communautés, ce qui est assez faible, les 8 premières regroupant plus de 90 % des comptes Twitter.



## Identification des leaders go-between

Du 26 janvier au 16 février 2020 inclus, le sujet est peu mentionné, avec 1127 tweets. Il y a 86 % de retweets, et seulement 580 utilisateurs. Les principales mentions correspondent à l'appel à l'action d'XR, qui n'indique pas de quoi il s'agira.

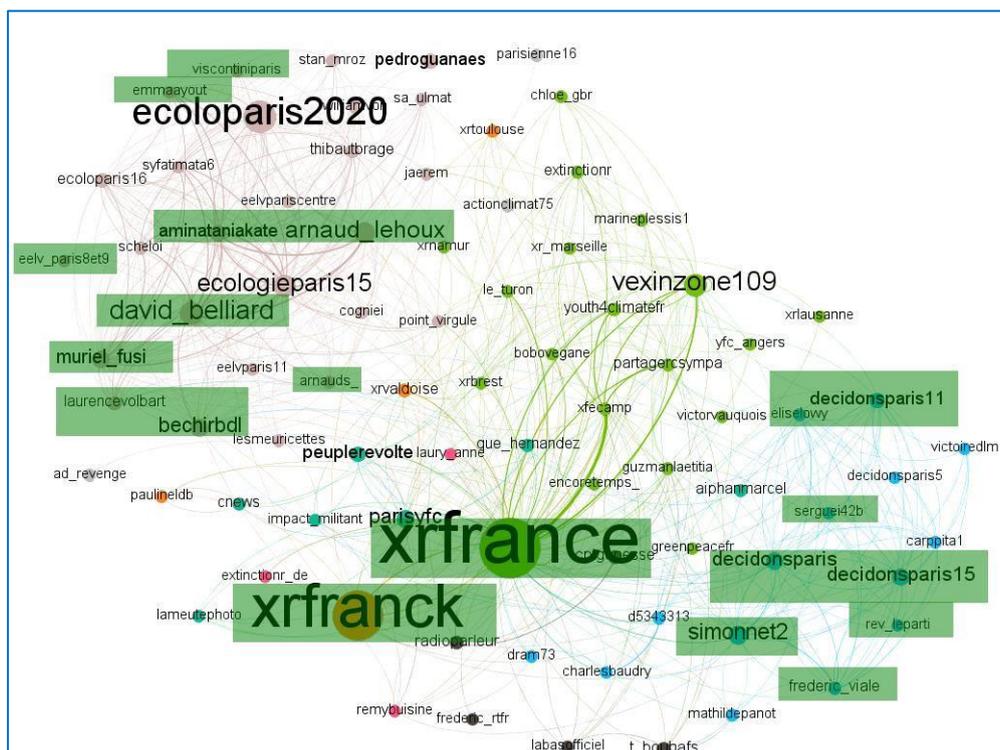
### Classement via Visibrain des Tweets les plus repris sur l'action XR Lafarge du 26 janvier au 16 février 2020 :

 <p><b>Extinction Rebellion France</b> 🦋🌿 @xrFrance            ⚠️ STOP à la bétonisation de notre avenir ! Rejoignez l'action #FinDeChantiers 🦋 le 17 février 📅 Inscription par ici : <a href="https://t.co/TiZPMc5SoH">https://t.co/TiZPMc5SoH</a> Rejoignez nous 📍 <a href="https://pic.twitter.com/w5pNqo0z5e">pic.twitter.com/w5pNqo0z5e</a>            05 févr. 2020 16h33 Twitter Web App</p>		<p><b>53</b> (RE)TWEETS</p>
 <p><b>Impact - média militant</b> @impact_militant            ⚠️ Plusieurs actions à Paris demain : 📍 Manif à l'appel de la coord nationale des grévistes et plusieurs AG interpro : 13h30 Opéra Garnier. #reformedesretraites #greve17fevrier 📍 Blocage #FinDeChantiers par @xrFrance @ParisYFC et d'autres collectifs. #ClimateAction #Ecologie <a href="https://pic.twitter.com/r8qB94bKcF">pic.twitter.com/r8qB94bKcF</a>            16 févr. 2020 10h34 Twitter for Android</p>		<p><b>49</b> (RE)TWEETS</p>
 <p><b>David Belliard</b> @David_Belliard            Stop au béton à Paris, stop à la cimenterie Lafarge et aux travaux d'extension du site. Dans le 15e comme partout, nous rendrons les berges de Seine à la nature et aux habitant.e.s. Enfin ! @AminataNiakate @Murel_FUSI #EcologieParis <a href="https://pic.twitter.com/vhBqq6n1KG">pic.twitter.com/vhBqq6n1KG</a>            15 févr. 2020 14h19 Twitter for Android</p>		<p><b>44</b> (RE)TWEETS</p>
 <p><b>Peuple Révolté</b> @PeupleRevolte            ⚠️ Demain @xrFrance @ParisYFC et d'autres collectifs organisent une action de blocage : #FinDeChantiers 📍 Cette action débutera à l'aube #ecologie #ClimateAction #desobeissancecivile <a href="https://pic.twitter.com/dsEi7ohZeB">pic.twitter.com/dsEi7ohZeB</a>            16 févr. 2020 10h22 Twitter for Android</p>		<p><b>40</b> (RE)TWEETS</p>
 <p><b>Extinction Rebellion France</b> 🦋🌿 @xrFrance            La France FR championne des écocides 🦋🌿🌿🌿 Rejoignez l'action #FinDeChantiers 🦋 le 17 février ! Et il est encore temps de s'inscrire par ici : <a href="https://framaforms.org/fin-de-chantie...">framaforms.org/fin-de-chantie...</a> #ExtinctionRebellion #SuperLocal <a href="https://pic.twitter.com/4lquo6MH4Q">pic.twitter.com/4lquo6MH4Q</a>            08 févr. 2020 20h53 Twitter Web App</p>		<p><b>34</b> (RE)TWEETS</p>
 <p><b>Extinction Rebellion France</b> 🦋🌿 @xrFrance            📍 Avec d'autres organisations, nous appelons à résister, partout, à la destruction programmée du vivant. 📅 Lundi 17/02   C'est #FinDeChantiers 🦋 ; nous bloquerons l'industrie de la construction. Join us ! 📍 <a href="https://framaforms.org/fin-de-chantie...">framaforms.org/fin-de-chantie...</a> <a href="https://reporterre.net/Nous-appelons-...">reporterre.net/Nous-appelons-...</a>            28 janv. 2020 09h15 Twitter Web App</p>		<p><b>34</b> (RE)TWEETS</p>

On identifie via Gephi 37 communautés, les 8 premières totalisant plus de 90 % des comptes Twitter. La cartographie des comptes les plus mentionnés montre que la communauté XR a très peu d'interactions avec les acteurs politiques locaux impliqués sur le quai de Javel, ce qui est cohérent avec le secret de l'opération. **Dans les deux cas le fonctionnement de chaque leader go-between diffère entre les leaders de XR (avec des « opérations chocs ») et les leaders locaux (qui opèrent plus avec un consensus plus large sur les opérations citoyennes et écologiques).** Pour autant, certaines interactions existent, montrant que des militants des différents groupes ont participé à la préparation de l'action.



**Cartographie via Gephi et Visibrain des comptes au centre des échanges sur l'action XR Lafarge du 26 janvier au 16 février 2020 – application d'un filtre pour garder les comptes ayant un score élevé**

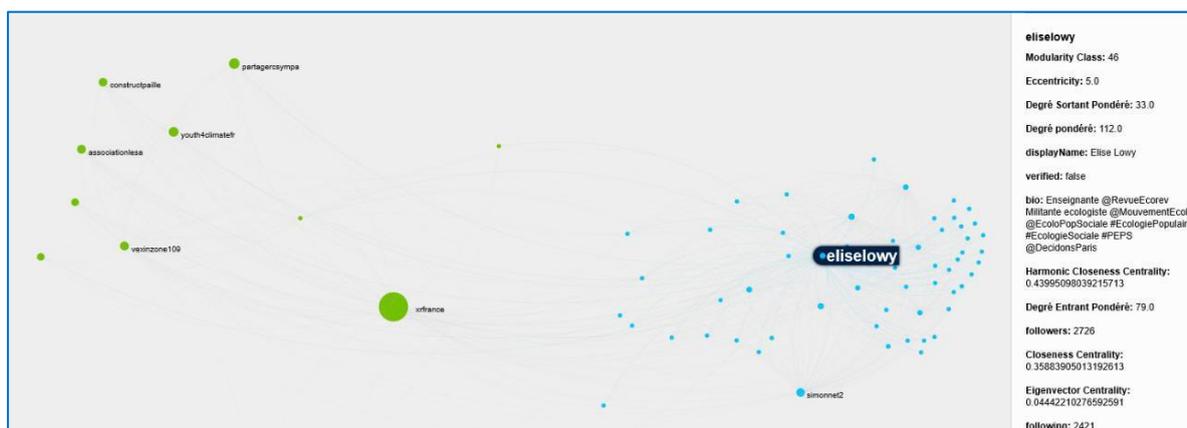


**Les leaders go-between locaux restent les plus influents**

Ce sujet montre l'intérêt d'effectuer des actions impliquant des acteurs locaux engagés afin de propager plus facilement l'information. **Ici le rôle des leaders go between l'emporte sur les opérations des leaders de XR car ils ont su mobiliser sur des actions plus larges et des mots d'ordre plus généraux ne heurtant pas les citoyens ou les élus locaux (Baulant , 2017).** On peut donc anticiper une montée en puissance de ces actions, à partir de mobilisations locales peu visibles, mais qui impliquerons divers acteurs aux buts convergents : militants d'ONG, élus locaux, et riverains donc.

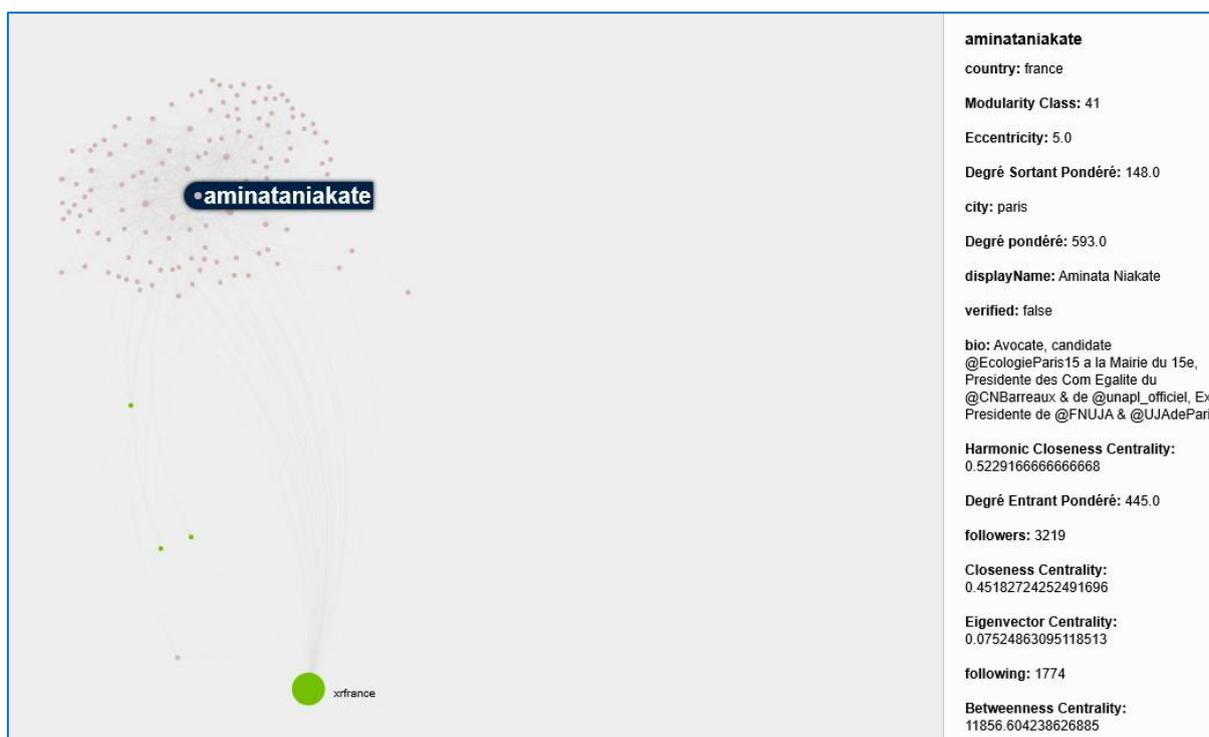
Si la majorité des acteurs communiquait de façon isolée sur ce sujet, puisque la mobilisation d'XR était secrète, et les deux communautés politiques sont en concurrence, on peut identifier plusieurs passerelles montrant la capacité de ces acteurs à échanger.

## Cartographie dynamique via Sigma - Gephi des échanges autour du compte Twitter d'Elise Löwy sur l'action XR Lafarge du 26 janvier au 19 février 2020



Elise Löwy, militante de Décidons Paris, est ainsi reliée à plusieurs membres de la communauté XR après une [action organisée contre Lafarge le 8 février 2020](#).

## Cartographie dynamique via Sigma - Gephi des échanges autour du compte Twitter d'Amina Niakate sur l'action XR Lafarge du 26 janvier au 19 février 2020



Amina Niakate, candidate sur les listes des Verts dans le 15<sup>ème</sup> arrondissement de Paris, communique elle sur [une action réalisée le 12 février](#).

Les Verts semblent néanmoins plus isolés que Décidons Paris de la communauté XR.