

# Quels outils pour appréhender et analyser les mobilisations de Gilets jaunes et les données issues du Débat national ?



© Brice Le Gall, <https://www.bricelegall.com/>  
Rond-point de Chevières dans l'Oise, le 19 décembre 2018

Diego ANTOLINOS-BASSO<sup>1</sup>, Stéphane BACIOCCHI<sup>2</sup>,  
Frédéric CASSOR<sup>3</sup>, Flora CHANVRIL<sup>4</sup>, Estelle CZERNY<sup>5</sup>,  
Guillaume GARCIA<sup>6</sup>, Bénédicte GARNIER<sup>7</sup>, Viviane LE HAY<sup>8</sup>,  
Victor LEPAUX<sup>9</sup>, Jean-Baptiste PRESSAC<sup>10</sup>, Paul SALZE<sup>11</sup>,  
Yannick SAVINA<sup>12</sup>

1. Sciences Po, Centre de recherches politiques (CEVIPOF), CNRS, Paris Sciences Po, médialab, Paris
2. CRH-LaDéHis, EHESS
3. Sciences Po, Centre de recherches politiques (CEVIPOF), CNRS, Paris
4. Sciences Po, Centre de recherches politiques (CEVIPOF), CNRS, Paris,  
Adresses de contact : [flora.chanvril@sciencespo.fr](mailto:flora.chanvril@sciencespo.fr), [mate-shs@services.cnrs.fr](mailto:mate-shs@services.cnrs.fr)
5. SAGE, Université de Strasbourg
6. Sciences Po, Centre de données socio-politiques (CDSP), CNRS, Paris
7. Institut national d'études démographiques (INED)
8. CNRS, Centre Emile Durkheim, Sciences Po Bordeaux
9. SAGE, CNRS
10. Centre de recherche bretonne et celtique, Université de Bretagne Occidentale
11. LIVE (laboratoire Image, Ville, Environnement), UMR 7362 CNRS/Université de Strasbourg
12. Sciences Po, Observatoire sociologique du changement (OSC), CNRS, Paris

**TITLE**

Which tools (do we need) to understand the Yellow Vests movement and analyze the data of the French National Debate?

**RÉSUMÉ**

Dans cet article introductif, nous revenons sur la genèse des journées d'études sur le mouvement des Gilets jaunes et le Débat national qui se sont déroulées les 16 et 17 janvier 2020.

Après avoir replacé ces journées dans leur contexte – une mobilisation sociale inédite et un foisonnement d'analyses « à chaud » – ce texte présente les questions méthodologiques posées par l'étude de ce mouvement, de l'analyse des réseaux sociaux à l'épistémologie du débat. Nous détaillons ensuite quelques principaux résultats démontrés dans les articles du numéro (articles tirés d'expériences empiriques généralement qualifiées de « quantitatives »).

**Mots-clés :** *Gilets jaunes, Débat national, méthodes quantitatives, mouvement social, réseaux sociaux, analyse de données textuelles, apprentissage automatique, transparence.*

**ABSTRACT**

This special issue is drawn from a workshop held in Paris on January 16 and 17, 2020, about empirical ways of studying the French Yellow Vests movement and the “National Debate” it sparked.

In this introductory article to this special issue, we take a look back at the origin of the workshop. After putting it back in context – an unprecedented social mobilization and an abundance of analyses in the heat of the moment – we present the methodological questions raised by the study of this movement, from the analysis of social networks to the epistemology of the debate. We then present some of the main results drawn from the articles in the issue (articles that, given the journal, are drawn from empirical experiments usually described as “quantitative”).

**Keywords:** *Yellow Vests, French National Debate, quantitative methods, social movement, social networks, textual data analysis, machine learning, transparency.*

## 1. Introduction

### 1.1 Un mouvement difficile à cadrer

À l'automne 2018, en réaction notamment à la hausse du prix des carburants suite à l'augmentation de la taxe intérieure de consommation sur les produits énergétiques (TICPE), des manifestations voient le jour sur l'ensemble du territoire national. Alimentée par les réseaux sociaux, cette mobilisation se veut apolitique, non structurée et sans leader. Face au maintien du mouvement et à ses revendications, l'exécutif répond entre autres par l'organisation d'un Grand débat national. En parallèle, un collectif de Gilets jaunes met en place le Vrai débat.

Nous retrouvons ici un processus classique de constitution en problème public d'une question sociale (Gusfield, 2009 ; Neveu, 2015) : un ou plusieurs problèmes sociaux se coagulent en mobilisation protestataire, laquelle va avoir comme débouché notable l'organisation de consultations publiques. Ceci illustre le *processus de mise sur agenda* de problèmes sociaux qui vont devenir ou pas des politiques publiques (Favre, 1992 ; Lagroye, 2003). Dans le cadre spécifique de la mobilisation des Gilets jaunes, des contraintes particulières apparaissent pour définir les contours d'une mobilisation qui se révèle protéiforme, à la fois dans son mode d'expression et dans ses revendications. Des difficultés surgissent aussi lors des débats institutionnels, avec l'organisation de consultations « parallèles » qui ont eu des échos non négligeables dans l'espace public.

### 1.2 Florilège d'événements scientifiques « à chaud »

Le mouvement des Gilets jaunes a fortement mobilisé les chercheur·e·s, en particulier en sciences sociales. Un nouvel objet de recherche a émergé, fruit d'analyses issues d'approches disciplinaires multiples : sociologique, géographique, politique, économique, historique, ethnographique, juridique, mathématique, etc. Parmi les grandes questions qui se sont posées et qui sont encore aujourd'hui pour certaines en débat, nous pouvons citer, entre autres, celles ayant trait à la mobilisation sociale (Jeanpierre, 2019), à la géographie sociale (Depraz, 2019), au rapport au politique et aux institutions (Neyrat, 2019 ; Mauger, 2019), au rapport au populisme (Rouban, 2019), aux classes populaires ou à la stratification sociale (Bedock *et al.*, 2019) et/ou parfois dans une visée de comparaison internationale (Printemps arabes, Indignés, etc.) et historique (Noiriél, 2019). Les exemples de travaux cités ici donnent un aperçu de la variété des études publiées.



**Figure 1** – Bandeau réalisé avec les quelques 150 couvertures d'ouvrages consacrés aux Gilets jaunes en 2019

Plusieurs compilations d'analyses et événements scientifiques ont fleuri depuis le début du mouvement et au cours de l'année 2019 pour faire dialoguer les travaux en cours durant

cette période. Sans prétention à l'exhaustivité, prenons ici quelques exemples. Ainsi, le site géoconfluences (ENS Lyon) a mis en ligne dès novembre 2018 une sélection d'analyses portant sur ce « Que disent les sciences sociales sur le mouvement des Gilets jaunes ? »<sup>13</sup>. Toujours à l'ENS Lyon, la revue d'analyses « Les sciences sociales et le mouvement des Gilets jaunes »<sup>14</sup> proposée par le site SES-ENS s'articule avec la précédente. Au laboratoire lyonnais Triangle, le séminaire « L'ordinaire dans les sciences sociales : des vies en jaune »<sup>15</sup> a rythmé la saison 2019-2020. À Lille, en février 2019, une conférence-débat organisée par le CERAPS s'intitulait « Les Gilets jaunes. Enigmes et premières pistes d'analyse »<sup>16</sup>.

Du côté des juristes, une demi-journée d'études intitulée « Les "Gilets jaunes" au prisme du droit »<sup>17</sup> a été organisée en septembre 2019 par l'Université d'Evry et le Centre de Recherche Léon Duguit. En novembre 2019, l'ethnographie a été mise à l'honneur lors de deux journées à l'EHESS « Approches ethnographiques des gilets jaunes : enquêtes et expériences »<sup>18</sup>. Du côté des mathématicien·ne·s citons également le séminaire « Systèmes complexes en sciences sociales » organisé par le Centre d'analyse et de mathématiques sociales qui a consacré en 2019 deux de ses séances<sup>19</sup> à la présentation de projets d'analyse de données issues du Grand débat national. Ces différents événements scientifiques ont permis de faire dialoguer chacune des disciplines sur le mouvement des Gilets jaunes et le Débat national.

Pour autant, il a peu été question jusqu'ici des aspects strictement méthodologiques dans une optique interdisciplinaire. C'est ce que propose, à l'appui de travaux empiriques inédits, ce numéro de *Statistique et Société*, en se concentrant sur les démarches quantitatives.

## 2. Regard méthodologique

### 2.1 Genèse et structuration des journées

Ce numéro s'appuie sur une sélection de six contributions aux journées d'études « Quels outils pour appréhender et analyser les mobilisations de Gilets jaunes et les données issues du Débat national ? » organisées les 16 et 17 janvier 2020 à Paris<sup>20</sup>. Ces journées ont été portées par un groupe de 12 ingénieur·e·s spécialisé·e·s dans les méthodes de recherche en sciences sociales, co-auteur·e·s de cette introduction, organisateur·rice·s par ailleurs du séminaire MetSem à Sciences Po et/ou membres du réseau Mate-SHS, un réseau professionnel de spécialistes des méthodes en sciences sociales et humaines. Le séminaire MetSem est un lieu d'échanges et de partages autour de la pluralité des outils et des méthodes des sciences humaines et sociales. Il existe depuis novembre 2016 et l'archive des thématiques abordées et des supports de présentation est disponible en libre accès sur le blog hypotheses<sup>21</sup>. Mate-SHS « est un réseau métier initié et porté par des ingénieur·e·s qui travaillent à la production, au traitement, à l'analyse et à la représentation des données dans la recherche en Sciences Humaines et Sociales (SHS). Ce réseau est un réseau de l'InSHS (Institut national des Sciences Humaines et Sociales) du CNRS (Centre National de la Recherche Scientifique) »<sup>22</sup>.

Tou·te·s engagé·e·s pour nos métiers dans des structures différentes<sup>23</sup>, nos intérêts convergents pour un mouvement social inédit, dont la pervasivité l'amenait jusqu'aux portes de l'Université

13. <http://geoconfluences.ens-lyon.fr/actualites/veille/revues-de-presse/gilets-jaunes>

14. <http://ses.ens-lyon.fr/actualites/rapports-etudes-et-4-pages/les-sciences-sociales-et-le-mouvement-des-gilets-jaunes>

15. <http://triangle.ens-lyon.fr/spip.php?rubrique819>

16. <https://www.afsp.info/les-gilets-jaunes-enigmes-et-premier-pistes-danalyse/>

17. <https://www.univ-evry.fr/evenements/les-gilets-jaunes-au-prisme-du-droit.html>

18. <https://www.ehess.fr/journ%C3%A9es-d%C3%A9tude/approches-ethnographiques-gilets-jaunes-enqu%C3%AAtes-et-exp%C3%A9riences>

19. Voir les séances du 8 mars et du 10 mai 2019 : <http://cams.ehess.fr/seminaire-systemes-complexes-en-sciences-sociales-2018-2019/>

20. Le programme détaillé et les résumés des communications sont disponibles ici : [https://metsem.hypotheses.org/files/2020/06/ProgrammeEtResumes\\_JEGJDN1617012020.pdf](https://metsem.hypotheses.org/files/2020/06/ProgrammeEtResumes_JEGJDN1617012020.pdf)

21. <https://metsem.hypotheses.org/>

22. <http://mate-shs.cnrs.fr/>

23. Nos affiliations respectives témoignent ainsi de la diversité de nos cadres de recherche : Universités, Institut national d'études démographiques (INED), Ecole des hautes études en sciences sociales (EHESS), Centre National de la Recherche Scientifique (CNRS), Sciences Po...

ou de centres de recherche, nous ont réunis autour de l'organisation de journées d'études au moment du rendu des comptes du Grand débat national. La genèse de ces journées date de mars 2019. L'appel à communication a été lancé en mai 2019, après l'organisation du Grand débat national et à un moment où l'on pouvait déjà lire un grand nombre d'enquêtes et d'analyses du mouvement des Gilets jaunes. En témoigne la trentaine de propositions de communications que nous avons reçues et soigneusement examinées durant l'été 2019. Initialement prévues les 5 et 6 décembre 2019, soit environ un an après le début du mouvement, les journées ont été reportées aux 16 et 17 janvier 2020 en raison de la grève du secteur des transports. Des journées d'études sur un mouvement social « perturbées » par un autre mouvement social, quelle belle mise en abîme ! Les contributions qui en sont issues, ainsi que ce numéro qui leur fait suite, visent à montrer comment les sciences sociales se saisissent de l'actualité d'un objet complexe, soudain et inédit.

Ce numéro porte sur une partie des communications faites lors de ces journées et traite plus spécifiquement de méthodologies quantitatives. Avant de présenter plus en détail son contenu, revenons sur le cadre général de ces journées. Elles furent organisées en six sessions complémentaires, en fonction à la fois du type des méthodologies mobilisées et de différentes facettes de cet objet complexe (mobilisation sociale, débat national ou regard épistémologique, utilisation d'un ou plusieurs types de méthodes, comparaison de plusieurs terrains, etc.). Ainsi une première session pluraliste, ouvrant les journées, s'intitulait « Observer et compter (ethnographie, analyses spatiales et temporelles) » et avait pour objectif de présenter des travaux en lien direct avec les manifestations de Gilets jaunes. Une session portait entièrement sur la thématique des groupes en ligne et réseaux sociaux numériques. Deux sessions furent consacrées à la phase du Débat national : un premier temps « dissociant » Grand et Vrai débats (comment analyser qui a dit quoi et comment... dans le Grand ou le Vrai débat), un second les reliant dans une session comparative (comment comparer et quelle différence entre Grand et Vrai débats ?). Enfin deux sessions portèrent sur l'analyse et le regard critique portés sur les sources et matériaux d'analyse : le premier temps étant plutôt consacré à l'épistémologie et le second aux aspects plus techniques, informatiques et/ou documentaires.

## 2.2 Pluralisme et nouveautés méthodologiques

Les articles regroupés ici visent à fournir aux lecteur·rice·s un panorama des méthodes mobilisées ou adaptées à l'occasion de l'étude de ce mouvement et de ce qu'il produit dans l'espace social mais aussi des questionnements qui entourent leurs usages, notamment à l'occasion des Grand et Vrai débats. Comment étudier un tel mouvement social, de grande ampleur et évolutif ? Comment traiter, analyser et présenter les données multiples et variées qui émergent du Débat national ? Comment comparer les données issues de la plateforme officielle du Grand débat national<sup>24</sup> et celles issues du Vrai débat<sup>25</sup>, plateforme créée par un collectif de Gilets jaunes ?

Lors des journées, s'est immédiatement posée la question de la pluralité des méthodes : comment utiliser et adapter des approches quantitatives, qualitatives ou mixtes, dans l'urgence d'une telle mobilisation ? Aux côtés des méthodes qualitatives fortement et traditionnellement mobilisées pour l'étude des mouvements sociaux (Neveu, 2019 ; Della Porta, 2014), les approches quantitatives permettent, par la force du nombre, de proposer des mesures du niveau de la mobilisation (et de son évolution), d'investiguer la composition, les caractéristiques et les opinions des populations concernées, de les comparer, etc., et par les propos tenus d'identifier des champs lexicaux afin de déterminer les thèmes récurrents.

24. <https://granddebat.fr/>

25. <https://www.le-vrai-debat.fr/>

De manière transversale aux différents objets de recherche et disciplines, nombreux sont ceux·elles qui se sont interrogé·e·s sur le caractère novateur de ce mouvement, voire de « disruption » avec les mobilisations précédentes. Les différents articles présentés dans ce dossier illustrent comment cet objet vient questionner les méthodes utilisées par les sciences sociales, et pour ce qui nous concerne ici, celles ayant recours à la démarche statistique. Un fait social « nouveau » nécessite-t-il l'emploi de méthodologies de recueil, de traitement ou d'analyse de données elles aussi nouvelles, ou au contraire renouvelées ?

Parmi les points d'achoppement figurent notamment la relative nouveauté (Jeanpierre, 2019) du mouvement dans ses modes d'expression, avec l'usage massif des réseaux sociaux numériques comme outil d'organisation et comme espace virtuel de mobilisation, sa nature protéiforme et son refus de se situer sur les clivages politiques habituels. Il ne s'agit pas ici de débattre du degré réel de nouveauté d'une mobilisation comme celle des Gilets jaunes, objet par ailleurs de débats de spécialistes, en lien par exemple avec les mouvements des « places » dans la période contemporaine. Soulignons simplement la filiation avec les coordinations (Fillieule et Péchu, 1993 ; Neveu, 2019) et leur usage militant des « nouveaux médias », ancêtres des réseaux sociaux actuels (Kergoat, 1994 ; Granjon, 2001), pour se limiter au seul cas français et ne remonter que jusqu'aux années 1980. De nouveaux outils sont nécessaires pour saisir ces nouveaux modes d'expression et de nouvelles méthodes le sont pour les analyser. Les modalités d'enquête « habituelles » apparaissent difficiles à appliquer, particulièrement dans un temps court. On peut dénombrer deux principales difficultés, qui sont au cœur des problématiques auxquelles se confrontent par essence les méthodes dites « quantitatives ».

La première renvoie au questionnement originel de la statistique : comment dénombrer les populations (Desrosières, 2010 ; Martin, 2020) ? Mis à l'épreuve d'une mobilisation comme celle des Gilets jaunes, ce questionnement s'actualise d'abord dans le débat récurrent autour du comptage des manifestant·e·s, débat sur lequel la littérature en langue française est, à notre connaissance, étrangement assez peu développée (Fillieule, 1997 ; Fillieule, Tartakowsky, 2013 ; Muller, 2015 ; Suesser, 2015). Il s'actualise également dans un second débat portant sur la délimitation des populations à enquêter. La difficulté est redoublée quand on a affaire à une mobilisation aussi protéiforme que celle des Gilets jaunes, sans ancrage objectif, qu'il s'agisse d'une structure organisationnelle formelle ou de critères d'appartenance ou d'identification relativement clairs. Le standard des enquêtes en manifestations, qui est encore aujourd'hui l'étalon de la sociologie des mouvements sociaux (Andretta, Della Porta, 2014 ; Fillieule, Blanchard, 2008), trouve ici ses limites, notamment dans le comptage des foules. On peut mentionner ici le cabinet d'études Occurrences mandaté par les médias pour cette fonction depuis 2018<sup>26</sup>. L'option consistant à saisir la population mobilisée en lien avec l'espace de mobilisation le plus « fédérateur », à savoir les réseaux sociaux numériques, a ainsi été privilégiée.

Le mouvement des Gilets jaunes est né d'initiatives individuelles en ligne : pétition contre la hausse du prix des carburants, création de la page Facebook « officielle » du mouvement « La France en Colère » et diffusion d'une vidéo sur le même média pour dénoncer la « traque aux automobilistes ». Parallèlement, les Gilets jaunes ont utilisé ces espaces pour partager leurs revendications, se raconter, se filmer en direct et parler à des milliers de suiveurs déposant à leur tour des questions ou des commentaires, ce qui est une manière de faire du « web dans la rue » (comme l'est dans un tout autre contexte un panneau publicitaire connecté sur un abribus). Partout en France, des groupes Gilets jaunes se sont constitués. Près de 900 de ces groupes ont été recensés dans l'Atlas élaboré par Morales *et al.* Ils ont annoncé leurs actions, invité les abonnés à s'y joindre grâce aux pages « Événement » et ont coordonné leur mouvement. Ainsi au début de la mobilisation, chaque semaine, plus de 1500 événements Facebook, à l'échelle

26. Dominique Seux, « Manifestations : des médias lancent leur propre système de comptage » [<https://www.lesechos.fr/2018/03/manifestations-des-medias-lancent-leur-propre-systeme-de-comptage-987181>], sur [lesechos.fr](http://lesechos.fr), 2018 (consulté le 7 janvier 2019).

d'un quartier, d'un village, d'une ville ont été organisés.

Les Gilets jaunes se sont appropriés les plateformes numériques comme Facebook, Twitter, Discord, des chaînes YouTube, des messageries cryptées ou le moteur de recherche Google pour identifier les terrains de mobilisation, témoigner et trouver des personnes qui partagent les mêmes revendications ou les mêmes souffrances. Les Gilets jaunes ont d'ailleurs bénéficié d'une configuration favorisant leurs échanges : dans une optique de lutte contre les *fake news*, Facebook a en effet changé d'algorithme en janvier 2019, augmentant la visibilité des discussions et favorisant les contacts et échanges entre personnes se reconnaissant comme « ami·e·s » et ceux organisés dans des groupes. Cette évolution d'algorithme a agrégé/aidé à la mobilisation. Aujourd'hui les Gilets jaunes ont leur propre « portail collaboratif »<sup>27</sup>.

Par cet usage nouveau et intensif, des questions de méthode se posent : l'approche par les réseaux sociaux permet-elle de dépasser réellement certains problèmes de diversité ou de représentativité des échantillons abordés classiquement (sur l'exemple de Twitter : Morales, Cointet, Laborde, 2020 ; Barbera, 2015) ? Quels sont ses avantages pratiques, logistiques, face à la difficulté d'organiser une enquête classique avec peu de moyens sur une mobilisation en cours et fortement territorialisée, malgré tous les biais que ce médium comporte ? Comment mettre en œuvre des analyses spatialisées fines et nouvelles ?

Par ailleurs, les plateformes numériques ont été et sont encore le support de contenus expressifs personnels, notamment textuels. Les chercheur·e·s se sont par exemple emparé·e·s des commentaires rédigés sur les groupes Facebook pour en faire des analyses lexicométriques identifiant les champs lexicaux de leurs revendications. Cela nous amène à la seconde problématique : comment saisir des expressions, spontanées ou organisées, de l'activité militante en ligne ou de dispositifs participatifs, à partir de masses d'informations volumineuses et peu structurées ? Cette quantité d'information est telle qu'elle nécessite *a priori* la médiation de l'outil informatique, et suscite des interrogations quant aux opérations « invisibles » intervenant parfois au cœur de ce type d'outils dans lesquels les « boîtes noires » ne sont pas toujours maîtrisées. Mise à l'épreuve du mouvement des Gilets jaunes et des consultations publiques qui lui ont succédé, cette problématique s'actualise au cœur d'un débat déjà ancien pour les habitué·e·s de l'analyse de données textuelles, à propos des méthodes de classification et de catégorisation employées, et leurs limites.

Par exemple, peut-on et comment qualifier politiquement, idéologiquement, une mobilisation qui refuse au moins formellement de se laisser ramener aux grilles de lectures habituelles ? A quels référents, quels critères recourir, notamment lorsque les marqueurs sociographiques habituellement utilisés dans d'autres types d'enquêtes (questions fermées des enquêtes par questionnaire) font défaut ? Une approche inductive, véritablement libérée des catégories pré-construites est-elle réellement possible lorsqu'on doit traiter de grandes masses de traces numériques, dans un espace qui est très loin d'être circonscrit *a priori* ? Quels sont les effets des outils utilisés ? Le recours à l'intelligence artificielle est-il réellement de nature à dépasser ces limites des opérations humaines ?

### 3. Présentation du numéro

Les articles de ce numéro spécial portent sur ces aspects méthodologiques riches et variés abordés jusqu'ici et apportent également plusieurs résultats significatifs à l'étude du mouvement des Gilets jaunes.

27. <https://giletsjaunes-coordination.fr/>

### 3.1 De l'analyse des réseaux sociaux à l'épistémologie du débat

Ce numéro spécial s'ouvre sur une série d'articles s'attelant à caractériser les acteur·rice·s de cette mobilisation. Ainsi le texte « Enquêter sur les Gilets jaunes. Sociologie politique d'un mouvement social à partir d'une enquête diffusée sur les réseaux sociaux » analyse les résultats d'une enquête par questionnaire administrée auprès des membres de plusieurs groupes Facebook. Dans une première partie, les auteur·e·s interrogent la définition même de la population des Gilets jaunes et reviennent sur les difficultés rencontrées. Ils détaillent ainsi les différents « gradients de l'engagement » permettant de cibler cette population et le rôle joué par les réseaux sociaux dans l'émergence et la construction de la mobilisation des Gilets jaunes. Ils précisent ensuite la méthodologie d'enquête par questionnaire mise en place. Ainsi, au-delà du contenu du questionnaire et de la mobilisation du score EPICES et de différentes mesures du niveau d'engagement, les auteur·e·s reviennent sur la diffusion du questionnaire : ils se sont abonnés à différents groupes de Gilets jaunes « nationaux et locaux » afin de pouvoir « distribuer le lien vers le questionnaire ». Ils expliquent également l'évolution du nombre de répondants à l'enquête. Dans une troisième partie, ils présentent les premiers résultats portant sur l'analyse du lien entre précarité et distance à la politique. Enfin, ils terminent avec une réflexion sur l'importance des terrains d'enquête sur les réseaux sociaux pour analyser les nouveaux types de mobilisation sociale.

L'« Atlas multi-plateforme d'un mouvement social : le cas des Gilets jaunes » s'attache à « délimiter les contours de l'espace de revendication du mouvement et sa dynamique d'agrégation ». Les auteur·e·s décrivent tout d'abord leur corpus, portant sur un « écosystème » de trois plateformes numériques : Facebook, Twitter et Youtube. Après un rappel de l'importance du réseau Facebook dans la naissance du mouvement, ils expliquent comment récupérer les signaux émis et les contextualiser à partir des données de Twitter et de Youtube. Dans un second temps ils présentent les résultats de leurs analyses selon différentes dimensions : descriptive et spatiale d'abord, politique ensuite, puis étudient les thématiques abordées dans les groupes, et enfin précisent le rapport des Gilets jaunes aux médias. Les auteur·e·s développent enfin une réflexion sur les intérêts de l'utilisation des liens hypertexte dans la méthode employée, qui permet ainsi de parvenir à « une compréhension plus précise de la diversité du mouvement en permettant de positionner les groupes Facebook dans un espace idéologique et de les situer dans l'espace médiatique » mais également d'apporter un éclairage intéressant sur la « dynamique temporelle » du mouvement.

Dans l'article suivant « De quelle(s) couleur(s) sont les Gilets jaunes ? Plonger des posts Facebook dans un espace idéologique latent », les auteur·e·s cherchent à comprendre « Comment les pratiques de citation en ligne trahissent-elles non pas la couleur politique du mouvement, mais l'espace politique dont [les Gilets jaunes] se nourrissent et qu'ils alimentent. ». Pour répondre à cette question, ils utilisent plusieurs méthodes d'inférence idéologique. Après avoir passé en revue les méthodes d'inférence idéologique existantes, ils expliquent ensuite leur stratégie d'analyse « en trois étapes » qui se base sur des données issues de Twitter : d'abord la construction d'un espace idéologique latent avec une analyse des correspondances, puis le calcul de scores des liens partagés par les comptes Twitter correspondants et enfin la coloration idéologique des publications Facebook individuelles en fonction des liens cités. Ils peuvent ainsi procéder à une analyse dynamique de l'évolution de cette couleur politique au fil de l'avancée du mouvement à l'aide d'une extraction terminologique précise. Les auteur·e·s concluent leur article en portant un regard distancié sur les méthodes mises en application, et en particulier sur leur caractère inductif, souple et stable.

L'article « Les multiples agendas médiatiques des Gilets Jaunes sur YouTube. Exploration d'un corpus de vidéos avec les topics models » interroge également les transformations de l'espace et de l'agenda médiatique, à partir de la plateforme YouTube. Il étudie les proximités thématiques de contenus médiatiques produits par différents acteur·rice·s. Les analyses

probabilistes du corpus de sous-titres de vidéos comparent, entre autres, les productions des chaînes Gilets jaunes, des médias *mainstream* et des chaînes de vulgarisation politique ou de contre-information. Elles révèlent ainsi des points communs et des divergences ou spécificités, comme l'importance des sujets de citoyenneté pour les chaînes des Gilets jaunes ou la violence spectaculaire du côté des médias traditionnels. Elles témoignent aussi de la réorganisation de l'espace médiatique à la suite du Grand débat. Cet article montre enfin la pertinence du matériau mobilisé pour caractériser cet espace et la richesse des traitements que la quantification et l'analyse des *topics models* permet : classification, visualisation en réseaux et profils temporels notamment.

L'article suivant porte sur une comparaison des données issues des Grand et Vrai débats : « Que peuvent les algorithmes de plongement de mots pour l'analyse sociologique des textes ? Analyser les discours et caractériser les locuteurs des plateformes "Grand Débat National" et "Vrai Débat" ». L'objectif de cet article est double : méthodologique d'abord en comparant des algorithmes de plongement de mots aux méthodes de statistique textuelle plus classiques afin de mesurer leur apport en termes d'interprétation ; sociologique ensuite en déterminant les profils sociaux et géographiques des participants aux deux plateformes. Dans la comparaison des approches sont mises en balance les résultats obtenus d'une part avec une méthode textométrique (calcul de spécificités entre les deux corpus et classification descendante hiérarchique sur l'ensemble des deux corpus afin de repérer des univers lexicaux sous le logiciel IRaMuTeQ) et d'autre part une technique de plongement de mots de type modèle d'apprentissage appelée Word2Vec. Enfin un troisième corpus issu de la plateforme « Entendre la France » sert de point de contrôle et est utilisé pour caractériser sociologiquement et géographiquement les participants avec une méthode de type *machine learning*. Les auteur·e·s concluent leur article en revenant sur les résultats obtenus et en posant deux questions méthodologiques sur la problématique des « données absentes » et sur les usages des algorithmes.

Enfin, l'article « Pour plus de transparence dans l'analyse automatique des consultations ouvertes : leçons de la synthèse du Grand Débat National » s'intéresse à l'analyse officielle des contributions au Grand Débat National. Celle-ci constitue une « boîte noire » tant la méthodologie employée est peu documentée. Les auteur·e·s mènent l'enquête en procédant à une ingénieuse rétro-analyse : en partant des résultats publiés dans la synthèse officielle, ils tentent de démêler les fils des traitements automatiques opérés. Quels algorithmes de catégorisation des textes ont été employés ? Et au-delà, quel est le niveau d'intervention humaine ? Ils constatent l'impossibilité de toute réplique, la diversité des résultats obtenus selon la méthode employée (notamment le choix de représentation vectorielle des textes) et plusieurs anomalies. Ils en tirent la leçon d'une indispensable plus grande transparence dans le traitement automatique des consultations participatives.

Les intervenant·e·s, dans ces articles ou durant les journées, ont pris soin d'être réflexifs voire critiques sur ce que la nouveauté de l'outil ou des données peut ôter à l'analyse sociologique : au-delà des outils, la recherche de résultats interprétables.

### 3.2 Au-delà des méthodes, les résultats

Plusieurs points ont particulièrement retenu notre attention dans les conclusions des articles qui suivent.

Guerra *et al.* reviennent sur la définition de ce qu'est un Gilet jaune : leur schéma présentant leur population d'enquête montre ainsi les différents cercles de la mobilisation (participation active, identification, soutien) et la transversalité des participant·e·s aux groupes Facebook (des passifs aux actifs dans la rue, en passant par les actifs sur les réseaux sociaux en ligne). Cette question de la délimitation de la population GJ a traversé les questionnaires d'enquêtes et plus largement les chercheur·e·s et analystes durant tout le mouvement. En plus de ces différentes

couches de mobilisation, il est à noter la dispersion (environ 900 groupes Facebook repérés par Morales *et al.*) de celle-ci.

De nombreux enseignements sont à tirer sur la composition de la « nébuleuse » des Gilets jaunes. Guerra *et al.* montrent qu'il s'agit d'une population particulièrement précaire, avec des variantes selon le degré de mobilisation. En revanche, s'ils notent une certaine stabilité du positionnement politique, ils constatent une différence entre les Gilets jaunes de terrain et ceux actifs uniquement sur les réseaux sociaux. Cointet *et al.* reviennent sur l'hétérogénéité de l'espace idéologique des Gilets jaunes. Plus spécifiquement sur le positionnement gauche/droite, Morales *et al.* observent un déplacement vers la gauche au fur et à mesure de l'avancée dans le temps de la mobilisation. Avec leur algorithme, Suignard *et al.* apportent des éléments sur le profil des participants au Grand débat, avec une population soutenant au tiers le mouvement des Gilets jaunes et très masculine.

Benbouzid et Guérin présentent des résultats principalement liés au traitement médiatique des thématiques discutées par les Gilets jaunes. Ils montrent ainsi la différence entre ce qui est « produit par les médias » *mainstream* et ce qui est produit par de l'analyse de contenus. Sur les thèmes ressortant de leur analyse, on retrouve l'importance de la citoyenneté, mais aussi des thèmes fondateurs du mouvement, en particulier sur les questions de taxation et de pouvoir d'achat, qui restent plus importants que ceux ayant trait aux questions d'écologie.

Guerra *et al.* analysent également dans leur article deux types de sources de l'engagement des Gilets jaunes : plus de justice sociale et un meilleur fonctionnement démocratique. Suignard *et al.* comparent les discours sur les deux plateformes des Grand et Vrai débats. Le principal résultat est la différence d'échelle entre les deux plateformes : les participants au GDN se placent à une échelle macro et parlent du dérèglement climatique à l'échelle mondiale ou de macro-économie et de croissance, tandis que ceux du VDN parlent d'enjeux plus locaux, en citant des noms de villes par exemple, et concrets comme la thématique de la route (vitesse, radars, etc.), de la pollution, de la souffrance animale, des taxes, du pouvoir d'achat.

Enfin Bellet *et al.* posent le problème de la reproductibilité des analyses. Ils ne sont pas parvenus à « retrouver des effectifs comparables à ceux de la synthèse officielle » malgré les différentes méthodes utilisées. L'impossible reproductibilité aboutit à une demande de davantage de transparence technique, en particulier dans les algorithmes utilisés pour le traitement des données, et de davantage d'intervention humaine dans les traitements effectués.

L'ensemble de ces articles donne un riche aperçu de ce qui s'est dit lors des journées des 16 et 17 janvier 2020. Pour terminer, nous souhaitons remercier les autres auteur·e·s des journées d'étude, dont les présentations ont nourri les échanges et le débat, les invité·e·s qui ont apporté des regards connexes sur ces journées ainsi que les différentes institutions qui ont participé à leur financement : Institut national d'études démographiques (INED), Centre de données socio-politiques (Sciences Po, CNRS), médialab (Sciences Po), Centre de recherches politiques (Sciences Po, CNRS), Observatoire sociologique du changement (Sciences Po, CNRS), laboratoire Sociétés, Acteurs, Gouvernement en Europe (Université de Strasbourg, CNRS), Centre Emile Durkheim (Sciences Po Bordeaux, CNRS), Centre de recherche bretonne et celtique (Université de Bretagne Occidentale, CNRS)<sup>28</sup>.

28. L'ensemble des auteur·e·s et leurs contributions sont détaillés dans le programme des journées : [https://metsem.hypotheses.org/files/2020/06/ProgrammeEtResumes\\_JEGJDN1617012020.pdf](https://metsem.hypotheses.org/files/2020/06/ProgrammeEtResumes_JEGJDN1617012020.pdf). Les invité·e·s ont également apporté des regards connexes sur ces journées : Natalia La Valle Torres (référénte sociologie à la bibliothèque de Sciences Po) qui a tenu un stand bibliographique tout au long des journées ; Brice Le Gall (photographe, auteur de *Justice et Respect. Le soulèvement des Gilets Jaunes*, Édition Syllepse, décembre 2019) pour présenter son livre ; Pascal Perrineau (garant du Grand débat national) qui est intervenu en tant que garant du Grand Débat ; et Luc Rouban (Sciences Po, Centre de recherches politiques (CEVIPOF), CNRS) qui a animé une des six sessions de ces journées.

## Références

- Andretta M. and D. Della Porta (2014), « Surveying Protestors: Why and How », in D. Della Porta (ed.), *Methodological Practices in Social Movement Research*, Oxford, Oxford University Press, pp. 308-334.
- Barberá P. (2015), « Birds of the Same Feather Tweet Together: Bayesian Ideal Point Estimation Using Twitter Data », *Political Analysis*, vol. 23, n° 1, pp. 76-91.
- Collectif d'enquête sur les Gilets jaunes, Camille Bedock, Zakaria Bendali *et al.* (2019), « Enquêter *in situ* par questionnaire sur une mobilisation. Une étude sur les gilets jaunes », *Revue française de science politique*, vol. 69, n° 5, pp. 869-892.
- Della Porta D. (ed.) (2014), *Methodological Practices in Social Movement Research*, Oxford, Oxford University Press.
- Depraz S. (2019), « La géographie est-elle une science engagée ? Fracture(s) territoriale(s) et Gilets jaunes », *Historiens & Géographes*, n° 446, pp. 25-29.
- Desrosières A. (2010), *La politique des grands nombres. Histoire de la raison statistique*, Paris, La Découverte, coll. « Poche / Sciences humaines et sociales ».
- Favre P. (1992), « L'émergence des problèmes dans le champ politique », in P. Favre (éd.), *Sida et politique. Les premiers affrontements, 1981-1987*, Paris, L'Harmattan, pp. 5-37.
- Fillieule O. (1997), *Stratégies de la rue. Les manifestations en France*, Paris, Presses de Sciences Po.
- Fillieule O. and Ph. Blanchard (2008), « INdividual SURveys in RAllies (INSURA). A New Tool for Exploring Transnational Activism? », *University of Lausanne: Political Science Working Paper Series*, n° 35.
- Fillieule O. et C. Péchu (1993), *Lutter ensemble. Les théories de l'action collective*, Paris, L'Harmattan.
- Fillieule O. et D. Tartakowsky (2013), *La manifestation*, Paris, Presses de Sciences Po.
- Granjon F. (2001), *L'Internet militant : Mouvement social et usage des réseaux télématiques*, Rennes, Apogée.
- Gusfield J. (2009), *La culture des problèmes publics. L'alcool au volant : la production d'un ordre symbolique*, Paris, Economica, coll. « Études Sociologiques ».
- Jeanpierre L. (2019), *In Girum Les leçons politiques des ronds-points*, Paris, La Découverte.
- Kergoat D. (1994), « De la jubilation à la dérégulation. L'utilisation du Minitel dans les luttes infirmières (1988-1989). Note de travail » in P. Rozenblatt *et al.*, *Les coordinations de travailleurs dans la confrontation sociale*, Paris, L'Harmattan.
- Lagroye J. (éd.) (2003), *La politisation*, Paris, Belin.
- Le Gall B. (2019), *Justice et Respect. Le soulèvement des Gilets Jaunes*, Paris, Éditions Syllepse.
- Martin O. (2020), *L'empire des chiffres*, Paris, Armand Colin.

Mauger G. (2019), « Gilets jaunes », *Savoir/ Agir*, n° 47, pp. 109-117.

Ramaciotti Morales P., J.-Ph. Cointet, and J. Laborde (2020), « Your most telling friends: Propagating latent ideological features on Twitter using neighborhood coherence », *2020 IEEE/ ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*.

Muller P. (2015), « Compter le nombre de manifestants sur la voie publique : une problématique statistique mais aussi et surtout politique », *Statistique et Société*, vol. 3, n° 3, pp. 41-48.

Neveu É. (2015), *Sociologie politique des problèmes publics*, Paris, Armand Colin, coll. « U ».

Neveu É. (2019), *Sociologie des mouvements sociaux*, Paris, La Découverte.

Neyrat Fr. (2019), *La politisation des Gilets jaunes : retours sur la dynamique d'une mobilisation*, Collectif de réflexion #Gilets Jaunes. *La politique au rond-point*, Deux plumes, pp. 133-151.

Noiriel G. (2019), *Les Gilets jaunes à la lumière de l'histoire. Dialogue avec Nicolas Truong*, Editions de l'Aube.

Rouban L. (2019), « Les "Gilets jaunes", une transition populiste de droite », *The Conversation*, janvier 2019.

Suesser J.-R. (2015), « Nombre de manifestants : "Information partagée" ou "composante de la confrontation" », *Statistique et Société*, vol. 3, n° 3, pp. 49-52.