



HAL
open science

Panéliser Twitter, une mesure hybride de mesure des opinions en ligne

Julien Boyadjian

► **To cite this version:**

Julien Boyadjian. Panéliser Twitter, une mesure hybride de mesure des opinions en ligne. 13ème Congrès de l'Association française de science politique, Sciences-Po Aix, Jun 2015, Aix-en -Provence, France. hal-03346786

HAL Id: hal-03346786

<https://hal.umontpellier.fr/hal-03346786>

Submitted on 16 Sep 2021

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

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



Distributed under a Creative Commons Attribution - ShareAlike 4.0 International License

ST7 Les appropriations méthodologiques d'internet dans la recherche sur des objets politiques

Julien Boyadjian, Université de Montpellier, CEPEL, julien.boyadjian@hotmail.fr

Titre : Panéliser Twitter, une mesure hybride de mesure des opinions en ligne

En 2011, 35,6 % des internautes français âgés de 16 à 74 ans déclaraient publier du contenu sur les réseaux sociaux¹. La démocratisation de l'accès à internet² et le développement d'outils simplifiés de publication de contenu en ligne – plateformes de blogs, forums électroniques, réseaux sociaux, etc. – ont permis aux internautes ne disposant pas de compétences informatiques spécifiques de publier facilement et sans intermédiaire des messages sur les sujets et préoccupations de leurs choix³. Ainsi, en mars 2012, plus de deux millions de messages publiés sur le réseau social Twitter se référaient aux différents candidats à l'élection présidentielle française⁴.

À l'instar des réponses de questionnaire, des verbatims d'entretien, ou des observations ethnographiques, les messages publiés sur le Web 2.0 peuvent être appréhendés par les chercheurs en sciences sociales comme un véritable matériau d'enquête. Les milliers de messages publiés chaque jour sur le web constituent en effet autant d'indices de pratiques, d'attitudes et d'opinions exprimées par les internautes. Ce nouveau matériau présente néanmoins certaines spécificités :

- les messages publiés sur internet ne sont pas des productions générées artificiellement dans le cadre d'une interaction de recherche,
- les publications des internautes, notamment sur Twitter, sont systématiquement horodatées et peuvent être ainsi précisément situées dans le temps,
- ces messages, de par leur facilité d'accès et de collecte, permettent d'analyser certains phénomènes en ligne de manière exhaustive,
- les messages publiés sur internet ne sont pas systématiquement associés à des auteurs aux caractéristiques sociologiques et politiques clairement définies.

Une première singularité caractérisant ces données numériques concerne leur caractère non-artefactuel. Contrairement aux réponses enregistrées dans le cadre d'un questionnaire ou d'un entretien, les messages publiés par les internautes sur les réseaux sociaux ne sont pas des données générées artificiellement par une interaction de recherche. Les internautes ne publient pas des messages sur internet à la demande d'un chercheur ou d'un « sondeur ». Ils n'adaptent donc pas leur discours en fonction d'un protocole d'enquête. Pour les sciences sociales, cette différence est fondamentale, car tout leur « malheur »⁵ est précisément d'avoir à faire à des êtres qui peuvent s'approprier le discours des scientifiques et modifier leurs comportements en conséquence. L'analyse des messages publiés sur le web peut donc se prévaloir de ne pas

¹ Eurostat, « Statistiques sur la société de l'information », *DEPS*, 2013.

² Selon les données de la Banque Mondiale, 81,9% des Français disposaient, en 2013, d'un accès à internet. Cf. <http://donnees.banquemondiale.org/indicateur/IT.NET.USER.P2>.

³ Dominique Cardon *et alii*, « Présentation », *Réseaux*, 3/2006 (n° 137), p. 9-25.

⁴ Données recueillies à l'aide du logiciel *AMI software*, qui collecte et archive, de façon exhaustive et automatique, les messages publiés sur le réseau social.

⁵ Pierre Bourdieu, Jean-Claude Chamboredon et Jean-Claude Passeron, *Le métier de sociologue. Préalables épistémologiques*, Paris, Éditions de l'EHESS, 1968.

reproduire les biais inhérents aux enquêtes déclaratives, en premier lieu desquels le risque d'imposition de problématique⁶.

Un deuxième aspect caractérisant les messages publiés sur les réseaux sociaux concerne leur horodatage. Sur Twitter notamment, chaque tweet est associé à une date et une heure précise de publication. Les logiciels informatiques de *web-tracking*⁷, qui permettent de collecter et d'archiver des millions de données numériques, offrent alors la possibilité aux chercheurs d'étudier des messages plusieurs mois voire plusieurs années après qu'ils aient été publiés par leurs auteurs. Il devient dès lors possible d'étudier les réactions, les discours et les représentations des internautes sur des faits ou des événements passés au moment-même où se sont déroulés ces faits, cela sans faire appel à la mémoire, souvent défaillante, des enquêtés. Les logiciels de *web-tracking* permettent ainsi d'objectiver le niveau et la fréquence de publication de messages politiques sur internet, et de situer précisément cette production dans le temps.

Une troisième caractéristique que nous souhaitons mettre ici en lumière concerne le caractère potentiellement exhaustif conféré à l'analyse informatisée des autopublications numériques. Les logiciels de *web-tracking* permettent de collecter des quantités très importantes de données numériques. À titre d'exemple, le logiciel *Ami Opinion Tracker*⁸ nous a permis, dans le cadre de cette recherche, de collecter l'intégralité des tweets contenant l'occurrence « Hollande » sur le réseau social. Depuis le jour de création de la collecte, en mars 2010, tout message relatif à l'actuel Président de la République a été ainsi collecté et archivé par le logiciel, soit à ce jour plusieurs centaines de millions de tweets. Les logiciels de *web-tracking* permettent donc aux chercheurs d'opérer des analyses statistiques sur des populations potentiellement exhaustives (dans l'exemple donné, l'ensemble des individus ayant publié au moins un message à propos de François Hollande ces quatre dernières années sur Twitter). Il n'est dès lors plus nécessaire, pour obtenir des résultats statistiques représentatifs, d'échantillonner la population-cible étudiée. La nature numérique du terrain d'étude rend techniquement possible une analyse exhaustive des phénomènes sociaux étudiés en ligne.

Si ces quelques caractéristiques présentent donc de réels intérêts heuristiques et conditionnent en partie, nous le verrons, la manière dont ont été appréhendées ces données dans la littérature, les autopublications numériques présentent néanmoins l'inconvénient majeur de ne pas être systématiquement associées à des auteurs aux caractéristiques socio-démographiques clairement définies. Si sur les réseaux sociaux les individus indiquent parfois certaines informations concernant leur sexe, leur âge, leur profession ou leur lieu d'habitation, ces données sont parcellaires et ne permettent pas d'établir une sociographie fine des utilisateurs. Ainsi, à titre d'exemple, sur Twitter, seul un inscrit sur deux indique dans sa description de profil un indice concernant sa profession ou son statut social (*cf.* tableau 1).

⁶ Pierre Bourdieu, « L'opinion publique n'existe pas », art. cit.

⁷ Les logiciels de *web-tracking* permettent de collecter, archiver et analyser une masse importante de données numériques.

⁸ Ce logiciel a été conçu par la société *AMI software*, membre du projet ANR Imagiweb.

Tableau 1. Taux d'utilisateurs de Twitter indiquant dans leur description de profil diverses informations socio-démographiques (échantillon aléatoire de 620 comptes)

Information socio-démographique	%
Sexe	95,2
Âge	11,6
Lieu de résidence	60,2
Niveau de diplôme	36,3
PCS	52,6
Appartenance politique	42,6

Contrairement aux enquêtes par questionnaire, les analyses de *web-tracking* ne permettent pas de situer socialement les auteurs des messages collectés. Les chercheurs souhaitant analyser de manière quantitative les messages politiques publiés sur les réseaux sociaux, et notamment sur Twitter, peuvent alors adopter, schématiquement, deux stratégies de recherche :

- opter pour une analyse informatisée du réseau social (de type « *web-tracking* » ou « *digital trace data* ») et ainsi tirer profit des avantages heuristiques associés à l'étude des tweets (données non suscitées, précisément horodatées et exhaustivement collectées). Néanmoins, cette analyse informatisée, du moins à elle seule, ne permet pas de caractériser socialement les auteurs des messages étudiés,
- préférer une approche plus classique d'interrogation par questionnaire, et ainsi identifier les propriétés sociologiques des producteurs de messages. Cependant, cette approche ne permet pas de tirer profit des analyses de *web-tracking*.

Les récents travaux académiques consacrés aux usages politiques de Twitter privilégient majoritairement des analyses informatisées d'importants volumes de données⁹. Cette posture de recherche incite alors les chercheurs à adopter une analyse davantage centrée sur le contenu et le contexte de publication des messages que sur les logiques sociales qui sous-tendent leur production.

Des recherches académiques davantage centrées sur le contenu et le contexte d'émission des messages

Bien que participant d'un phénomène relativement récent (Twitter a été créé aux États-Unis en mars 2006), les usages politiques du réseau social ont donné lieu à un nombre important de travaux de recherche, dans un laps de temps très réduit. En 2014, Andreas Jungherr dénombrait ainsi, dans un article de synthèse, pas moins de 115 études consacrées à cet objet¹⁰. Au-delà de leur nombre et de leur nouveauté, ces travaux se caractérisent également par la diversité disciplinaire de leurs auteurs. Ce nouveau terrain d'enquête « numérique » a en effet été investi par des chercheurs issus de champs disciplinaires variés, dialoguant souvent peu entre eux, tels que notamment les sciences de l'information et de la communication, les sciences informatiques et, dans une moindre mesure, la science politique

⁹ Pour une synthèse de ces travaux, voir notamment Andreas Jungherr, « Twitter in Politics: A Comprehensive Literature Review », Disponible en ligne : SSRN: <http://ssrn.com/abstract=2402443>.

¹⁰ *Ibidem*.

anglophone (les travaux français de science politique consacrés à Twitter, et plus largement au web 2.0, restent en revanche relativement rares¹¹).

L'origine disciplinaire de ces auteurs (pour une grande part issue des sciences informatiques), explique alors sans doute leur tendance à privilégier des dispositifs d'enquête informatisés de *web-tracking* et à proposer des approches très empiristes et descriptives du réseau social. Ces travaux comptabilisent ainsi le nombre de messages politiques publiés¹², analysent leur contenu et leur tonalité¹³, en les comparant parfois à des données de sondage d'opinion ou à des résultats électoraux (l'enjeu est alors de déterminer si les tendances observées sur Twitter sont plus prédictives que les sondages)¹⁴. Seule une minorité de travaux propose d'interpréter les résultats empiriques observés à partir d'un cadre théorique d'analyse¹⁵. Les références théoriques alors mobilisées s'inscrivent principalement dans le courant de recherche de communication politique s'intéressant aux effets des médias, notamment à travers les questions de mise sur agenda (*agenda-setting*)¹⁶ et de cadrage médiatique (*framing*)¹⁷. Twitter apparaît de fait comme un terrain d'étude particulièrement ajusté à ces questions de recherche. Le réseau social est en effet un espace très réactif à l'actualité et aux événements télévisés¹⁸. De plus, les réactions des internautes peuvent être situées dans un contexte médiatique et politique précis, chaque message publié étant, nous l'avons vu, précisément horodaté. En revanche, l'identité sociale des auteurs de ces messages demeure un point aveugle de cette littérature. Les « twittos » sont appréhendés comme une population unifiée et homogène, évoluant dans une sorte d'apesanteur sociale.

A *contrario*, les – rares – travaux académiques¹⁹ préférant une enquête par questionnaire à une analyse informatisée du réseau social identifient les propriétés sociologiques des internautes sans toutefois questionner le contexte politique et médiatique de production des messages. Ces auteurs entreprennent ainsi de mesurer le nombre d'utilisateurs de Twitter dans l'ensemble de la population²⁰ (et constatent généralement qu'ils sont relativement peu nombreux : les utilisateurs de Twitter représentaient seulement 8% de la

¹¹ À titre d'exemple, on ne compte à ce jour, sauf erreur de notre part, qu'un seul article paru dans la *Revue française de science politique* mobilisant une analyse de type *web-tracking*, cf. Sylvain Paraisie et Jean-Philippe Cointet, « La presse en ligne au service de la démocratie locale. Une analyse morphologique de forums politiques », *Revue Française de Science Politique*, 2012/1 (vol. 62).

¹² Voir par exemple Julian Ausserhofer et Axel Maireder, « National politics on Twitter: structures and topics of a networked public sphere », *Information, Communication & Society*, 16(3), 2013, p. 291-314.

¹³ Voir par exemple Dang-Xuan, Linh, Stefan Stieglitz, Jennifer Wladarsch et Christoph Neuberger, « An investigation of influentials and the role of sentiment in political communication on Twitter during election periods », *Information, Communication & Society*, 16(5), 2013, p.795-825.

¹⁴ Pour une synthèse des travaux interrogeant la prédictivité de Twitter, cf. Daniel Gayo-Avello, « A meta-analysis of state-of-the-art electoral prediction from Twitter data », *Social Science Computer Review*, 2013, vol. 31 n°6, p. 649-679

¹⁵ Andreas Jungherr, art. cit.

¹⁶ Max Mac Combs et Donald Shaw, « The agenda-setting function of mass-media », *Public Opinion Quarterly*, n°36, 1972, p. 176-187.

¹⁷ William, Gamson, « News as framing: Comments on Graber », *American Behavioral Scientist*, 33, 1989, p. 157-166.

¹⁸ Valérie Jeanne-Perrier, « Parler de la télévision sur Twitter : une réception oblique à partir d'une conversation médiatique », *Communication & langages*, n°166, 2010, p. 127-147.

¹⁹ Voir par exemple Marija Bekafigo et Allan McBride, « Who tweets about politics? Political participation of Twitter users during the 2011 gubernatorial elections », *Social Science Computer Review*, 31(5), 2013, p. 625-643 ou Daniela Dimitrova, Adam Shehata, Jesper Strömbäck et Lars Nord, 2014. « The effects of digital media on political knowledge and participation in election campaigns: Evidence from panel data », *Communication Research*, 41(1), 2014, p. 95-118.

²⁰ Lee Rainie, Aaron Smith, Kay Schlozman, Henry Brady et Sidney Verba, « Social media and political engagement », *Pew Internet & American Life Project*, 2012.

population internaute américaine en 2010)²¹ ainsi que les caractéristiques socio-démographiques de ces individus. Aux États-Unis, les producteurs de tweets politiques seraient ainsi majoritairement des hommes, blancs, démocrates et indépendants, très fortement diplômés²². Ces résultats, s'ils semblent plausibles (et relativement proches nous le verrons de nos propres données dans le cas français), sont cependant à interpréter avec précaution. Étant donné le faible taux d'internautes déclarant être inscrits sur Twitter, ces enquêtes par questionnaires (administrées à des échantillons représentatifs de 1000 à 2000 Américains) reposent souvent sur des échantillons très réduits (cent à trois cents individus tout au plus)²³. Plus fondamentalement, une autre limite de ces enquêtes est de ne pas mettre en relation les données socio-démographiques des utilisateurs avec leur activité de publication. Il est alors impossible d'évaluer dans quelle mesure le volume et la fréquence de publication de messages politiques d'un individu est déterminée par ses propriétés sociales et politiques, L'âge, le niveau d'étude ou l'orientation politique des utilisateurs déterminent-ils le nombre de messages politiques publiés ?

La littérature apporte peu d'éléments de réponse à ces questions. Seule une recherche combinant une analyse de *web-tracking* et une approche classique d'interrogation par questionnaire est à même, selon nous, de mettre au jour les facteurs dispositionnels et contextuels qui président à la production de messages politiques sur le réseau. C'est à cette fin que nous avons constitué notre propre dispositif méthodologique de panélisation du réseau social Twitter.

Panéliser Twitter : une méthode combinant enquête par questionnaire et web-tracking

Administrer un court questionnaire à un échantillon représentatif d'utilisateurs du réseau social Twitter, puis collecter, à l'aide d'un logiciel spécialisé, l'intégralité des messages publiés par ces comptes durant plusieurs mois permet de tirer profit des avantages associés aux analyses de *web-tracking* (verbatim non suscités par le chercheur, messages précisément horodatés, etc.), sans en reproduire la principale limite (des auteurs aux propriétés sociales mal identifiées). Cependant, constituer un panel réellement *représentatif* d'utilisateurs du réseau social susceptibles de publier des messages politiques soulève certaines difficultés méthodologiques. Décrire rapidement les différentes étapes de la constitution de notre panel d'usagers du réseau social permet d'en rendre compte.

La constitution des panels « répondants » et « non répondants »

Afin de déterminer les propriétés sociologiques des individus qui publient des messages politiques sur Twitter, nous avons choisi d'administrer un court questionnaire à un échantillon représentatif d'utilisateurs du réseau social. Contrairement aux approches de *web-tracking* qui permettent d'analyser de manière exhaustive l'ensemble de la population d'un réseau social, l'administration d'un questionnaire nécessite, pour des raisons pratiques, de recourir à une procédure d'échantillonnage. Cette opération suppose donc d'être en possession d'une liste exhaustive des utilisateurs du réseau social. Or une telle liste n'existe pas à notre connaissance, ou du moins n'est pas rendue publique par Twitter. Il est en

²¹ Jason Gainous et Kevin Wagner, *Tweeting to Power: The Social Media Revolution in American Politics*, Oxford, UK: Oxford University Press, 2014.

²² *Ibidem*, p. 46.

²³ Andreas Jungherr, art. cit.

revanche possible de recenser l'ensemble des utilisateurs susceptibles de publier des messages politiques, en collectant, sur une période de temps donnée, l'intégralité des tweets comprenant un ou plusieurs mots-clés politiques²⁴. Nous avons ainsi collecté la totalité des tweets publiés sur le réseau entre le 1^{er} et le 31 mars 2012 comprenant le nom (correctement ou incorrectement orthographié²⁵), le surnom, ou les principaux *hashtags* des candidats au premier tour de l'élection présidentielle française du 22 avril 2012. Durant cette période de forte politisation de l'actualité, plus de 2 800 000 tweets ont été ainsi collectés, provenant de 248 628 comptes uniques²⁶.

Nous avons ensuite extrait de cette base un échantillon aléatoire de 20 000 comptes que nous avons soumis à une analyse « manuelle » d'éligibilité. Pour être éligibles au panel, les comptes devaient avoir pour auteur un individu unique (et non une organisation, un média, ou une entreprise) et devaient être rédigés majoritairement en langue française. Nous avons ensuite « followé » les 10 229 comptes éligibles afin qu'ils consultent notre profil, qui les invitait à répondre à un questionnaire en ligne, comprenant une vingtaine de questions socio-démographiques (sexe, âge, PCS, niveau d'études, origine sociale, etc.) et politiques (orientation politique, trajectoire de vote, activités militantes, etc.). Après une vague de relances, ce sont 658 individus (6,4 % des comptes éligibles) qui ont répondu au questionnaire et 608 (6,3 %) qui ont accepté de faire partie du panel (le questionnaire spécifiait alors que leurs tweets seraient collectés et analysés – anonymement – par une équipe de chercheurs). Depuis lors, aucune autre interaction n'a eu lieu avec les enquêtés.

Afin d'évaluer dans quelle mesure les individus qui ont accepté de répondre à notre questionnaire étaient représentatifs de l'ensemble des comptes éligibles (et ainsi d'objectiver les possibles biais de sélection propres à toute enquête par questionnaire), nous avons constitué en parallèle de ce « panel répondants », un panel témoin, constitué d'un échantillon aléatoire de 620 individus qui n'avaient pas répondu au questionnaire. Nous avons alors cherché à qualifier socialement et politiquement les membres de ce « panel non répondants » à partir des informations parcellaires que les internautes indiquent parfois sur les réseaux sociaux ou leurs blogs. Ce panel « non répondants » se compose donc à la fois d'individus pour lesquels nous disposons d'une ou plusieurs informations socio-démographiques mais également d'individus pour lesquels nous n'en possédons aucune.

À l'aide de notre logiciel, nous avons enfin archivé et analysé la totalité des tweets publiés par les panels « répondants » et « non répondants » durant une période de 11 mois, du 1^{er} mars 2013 au 31 janvier 2014. Ainsi, sur cette période, près de deux millions de tweets (dont près de 125 000 messages contenant des occurrences politiques) ont été publiés par les membres de ces deux panels. Avant d'évoquer les différents résultats issus de l'analyse statistique des panels, nous souhaiterions ici évoquer les avantages de ce dispositif méthodologique.

²⁴ Selon notre définition de la politique donnée plus haut, des mots-clés se référant aux noms des principaux partis, acteurs et institutions politiques.

²⁵ Afin d'être le moins discriminant possible socialement.

²⁶ On peut alors supposer, sans toutefois être en mesure de le vérifier, que les utilisateurs qui n'ont publié aucun message politique durant cette période de très forte politisation des discussions ordinaires auront peu de chances d'en publier par la suite.

Les avantages de la panélisation

Notre dispositif de panélisation du réseau social permet selon nous de concilier les principaux apports des analyses de *web-tracking* et des enquêtes par questionnaires.

Comme dans les travaux reposant exclusivement sur des recherches de *web-tracking*, les messages publiés par les membres des panels ne sont pas suscités par l'intervention d'un chercheur. Si une interaction avec les enquêtés a bien eu lieu au début de l'enquête, afin d'administrer le questionnaire, plus aucun autre contact n'a été engagé par la suite. Il y a donc fort à parier qu'entre la date à laquelle les enquêtés ont répondu au questionnaire (en mars 2012) et la date à laquelle a débuté l'analyse informatisée de leurs messages (en mars 2013, soit un an plus tard), les enquêtés aient possiblement oublié notre enquête. D'autre part, tout comme les travaux reposant sur des analyses de *web-tracking*, les messages enregistrés par notre logiciel sont systématiquement horodatés : il est donc possible d'analyser la distribution quotidienne du nombre de tweets publiés, et de situer cette production dans un contexte politique précis.

Notre dispositif de panélisation du réseau social comporte néanmoins deux différences majeures par rapport aux recherches de *web-tracking*. D'une part, il ne permet pas d'étudier l'exhaustivité des messages publiés sur le réseau social. Néanmoins, notre panel constitue un échantillon statistiquement représentatif de l'ensemble des individus susceptibles de publier des tweets politiques sur le réseau social. La taille de l'échantillon ($n = 658$) et le caractère aléatoire de sa sélection assurent en effet sa représentativité et autorisent le traitement statistique. De plus, comme il l'a été dit plus haut, pour contrôler les possibles biais de sélection induits par le questionnaire, un panel témoin d'individus non répondants a été constitué. La comparaison des données issues de notre panel avec ce panel de contrôle permettra donc d'évaluer, le plus fréquemment possible, la représentativité des résultats observés.

D'autre part, contrairement aux recherches de *web-tracking*, notre dispositif de panélisation permet d'associer systématiquement les messages collectés par le logiciel aux données socio-démographiques et politiques de leurs auteurs. L'administration du questionnaire a en effet permis d'identifier précisément les caractéristiques sociologiques de chaque membre du panel (sexe, âge, PCS, intérêt pour la politique, etc.). Ces informations ont été programmées dans notre logiciel, de telle sorte que chaque message collecté soit associé aux caractéristiques sociales et politiques de son auteur. Ces données peuvent alors permettre, premièrement, d'établir une sociographie fine des individus qui s'expriment politiquement sur Twitter – et ainsi questionner l'hypothèse d'un espace de prise de parole « démocratique » –, deuxièmement, de comparer le volume et la fréquence de publication de messages entre différentes catégories d'utilisateurs (ouvriers et cadres, individus peu politisés et fortement politisés, électeurs de gauche et de droite, etc.), ou dit autrement, d'évaluer l'impact des variables sociologiques sur le nombre de tweets politiques publiés, et enfin, troisièmement, d'objectiver l'influence du contexte politique et médiatique sur ces différentes catégories d'utilisateurs.

L'analyse longitudinale des panels « répondants » et « non répondants » permet de vérifier ces hypothèses, et d'identifier ainsi les facteurs dispositionnels et contextuels qui président à la publication de messages politiques sur le réseau social. Cependant, avant-même d'évoquer ces différents facteurs, le résultat sans doute le plus manifeste qu'il convient d'évoquer ici concerne la très forte sélection sociale des membres de ces panels.

Un espace de prise de parole politique socialement très sélectif

Les individus qui publient, plus ou moins régulièrement, des messages politiques sur Twitter sont le produit d'une très forte sélection sociale. On observe en effet parmi cette population une nette surreprésentation d'hommes, diplômés, parisiens, occupant des positions de cadres. Fortement intéressés par la politique, votants assidus, militants : les producteurs de tweets politiques témoignent également d'un niveau très élevé de politisation. Cependant, contrairement aux populations d'ordinaires les plus politisées et les plus participationnistes, les producteurs de tweets politiques présentent une certaine spécificité sociologique : les catégories les plus jeunes de la population y sont très nettement surreprésentées.

Une jeunesse surreprésentée mais non représentative

La propriété *a priori* la plus remarquable des membres de notre panel est en effet leur jeune âge. Les catégories les plus jeunes de la population sont nettement surreprésentées au sein de notre échantillon. Alors qu'en 2012 les 18-25 ans représentaient 9,7 % de la population française, 39,8 % des individus qui ont répondu au questionnaire et 45,8 % de ceux qui n'ont pas répondu (et dont l'âge a pu être identifié) appartenaient à cette classe d'âge.

Tableau 2. Répartition par classe d'âge (en %) des populations « répondants » et « non répondants »

Classe d'âge	Panel Répondants (n = 658)	Panel Non Répondants (n = 240*)	Population française INSEE 2012
Moins de 18 ans	8,1	19,6	22,2
18 à 25 ans	39,8	45,8	9,7
26 à 34 ans	20,1	14,2	11,1
35 à 45 ans	18,8	14,6	14,7
46 à 60 ans	9,3	4,6	20,0
61 ans et plus	4,0	1,3	22,2
Total	100,0	100,0	100,0

*Individus dont l'âge a pu être identifié (soit 38,8% des non répondants tirés au sort)

Source : INSEE

Ce résultat peut sembler surprenant. En effet, les catégories les plus jeunes de la population sont précisément celles qui participent et qui votent le moins²⁷. C'est en effet parmi les 18-24 ans que l'on compte proportionnellement le plus d'individus qui ne sont pas inscrits sur les listes électorales et le plus d'abstentionnistes. Pourtant ce sont bien les 18-24 ans qui, à l'intérieur de notre panel, se sont le plus exprimés politiquement sur Twitter. Sur le réseau social, à l'inverse de ce que l'on observe dans le domaine électoral, les individus les plus jeunes sont donc ceux qui « participent » le plus. On pourrait alors supposer, à la lecture de ces résultats, que les « jeunes » ne seraient peut-être pas moins politisés, mais simplement moins attirés par les formes « classiques » de participation (voter, militer dans un parti, etc.) et

²⁷ Anne Muxel, *L'expérience politique des jeunes*, Paris, Presses de Science Po, 2001.

trouveraient dans les réseaux sociaux une nouvelle forme d'engagement qui leur conviendrait davantage. Les données issues de notre questionnaire semblent nuancer cette hypothèse. Les individus les plus jeunes qui ont publié des tweets politiques sont précisément ceux qui ont voté au premier tour des élections présidentielles. Alors que les Français âgés de 18 à 24 ans sont proportionnellement ceux qui ont le moins voté le 22 avril 2012 (seuls 72,4 % de cette classe d'âge a voté), les individus du même âge qui ont publié des tweets politiques ont *déclaré* voter à près de 92 % – soit le taux de participation par classe d'âge le plus élevé de notre panel « répondants ».

Les 18-24 ans qui composent notre panel ne sont donc pas représentatifs de l'ensemble des Français âgés de 18 à 24 ans. C'est en réalité seulement la fraction la plus politisée de cette classe d'âge qui publie des messages politiques sur Twitter. Comme le laisse suggérer leur haut niveau de participation électorale, les individus les plus jeunes de notre panel sont issus pour la plupart de milieux sociaux relativement aisés. Parmi les individus âgés de 20 à 24 ans qui ont répondu à notre questionnaire, 78,6 % sont étudiants, alors que les étudiants ne représentent que 21,7 % de cette classe d'âge²⁸. Mais les étudiants de notre panel ne sont pas pour autant représentatifs de l'ensemble des étudiants français. La répartition des étudiants par type d'établissement fréquenté en est une preuve. Alors que les étudiants engagés dans des formations universitaires courtes (type BTS, STS, etc.) représentaient 16 % des effectifs de l'enseignement supérieur en 2011-2012²⁹, seuls 6 % des étudiants de notre panel étaient inscrits dans ces filières. De manière encore plus remarquable, au sein de notre panel, le taux d'étudiants inscrits dans les grandes écoles et leurs classes préparatoires (31,9%) était près de 20 points supérieur à la moyenne nationale (12,9%)³⁰. Au sein de notre échantillon, près d'un étudiant sur trois est en effet inscrit dans une grande école (école de commerce, école d'ingénieur, Institut d'Études Politiques, école de journalisme, etc.).

Cette forte proportion d'étudiants issus des grandes écoles est selon toute vraisemblance révélatrice d'une origine sociale élevée et d'un haut niveau de capital culturel. Mesurée par la PCS du père, l'origine sociale des étudiants de notre panel confirme cette hypothèse. Près de 42 % des étudiants qui ont répondu au questionnaire sont des filles et fils de cadres et de professions intellectuelles supérieures. Au niveau national, ce taux était de 30,3 % en 2011-2012³¹. Ces résultats confirment donc l'origine sociale élevée des étudiants qui ont publié des tweets politiques durant la dernière séquence présidentielle. Issus majoritairement des classes moyennes supérieures, les étudiants de notre panel témoignent d'un haut niveau de capital culturel et de compétence politique. Les « jeunes » qui publient des tweets politiques ne sont donc représentatifs que de la fraction la plus politisée de la jeunesse.

Mise à part leur jeune âge – qui ne peut donc être interprété comme un signe de démocratisation de la participation politique – les individus qui publient des messages politiques sur Twitter présentent en réalité des propriétés sociales très similaires aux individus appartenant aux fractions les plus politisées de la population.

²⁸ Enquête de la DARES, « Emploi et chômage des 15-29 ans en 2012 », novembre 2013.

²⁹ Atlas régional, « Effectifs d'étudiants en 2011-2012 » : <http://www.enseignementsup-recherche.gouv.fr/cid67140/atlas-regional-effectifs-d-etudiants-en-2011-2012.html>

³⁰ *Ibidem*.

³¹ Repères et références statistiques sur les enseignements, la formation et la recherche. Statistiques - publications annuelles, Édition 2012 http://media.education.gouv.fr/file/2012/36/9/DEPP-RERS-2012_223369.pdf

Des propriétés sociales atypiques

Les producteurs de tweets politiques occupent ainsi des positions relativement élevées dans l'espace social. Les cadres et professions intellectuelles supérieures sont ainsi très nettement surreprésentés au sein de notre échantillon. En effet, alors que les cadres ne représentent que 14,8 % des Français en âge de travailler et n'étant plus en cours d'étude, 49,2 % des individus qui ont répondu à notre questionnaire et 71,1 % de ceux qui n'ont pas répondu (et dont la position sociale a pu être identifiée³²) appartiennent à cette catégorie socio-professionnelle. À l'inverse, on enregistre une sous-représentation des ouvriers et des employés. Seuls 2,4 % des répondants et 1,0 % des non répondants (dont la profession a pu être identifiée) sont des ouvriers, alors que cette catégorie socio-professionnelle représente 19,1 % de la population française en âge de travailler et n'étant plus en cours d'études.

Tableau 3. Répartition (en %) des panélisés selon la PCS

PCS (hors lycéens et étudiants)	Répondants (n = 378)	Non Rép. (n = 201)**	INSEE 2012
Agriculteurs exploitants	0,3*	0,0*	1,0
Artisans, commerçants et chefs d'entreprise	2,4*	4,0*	3,4
Cadres et professions intellectuelles sup.	49,2	71,1	9,6
<i>Professions libérales et assimilés</i>	4,0*	2,0*	1,0
<i>Cadres de la fonction publique, professions intell.</i>	21,7	33,8	3,2
<i>Cadres d'entreprise</i>	23,5	35,3	5,4
Professions intermédiaires	25,7	16,9	13,3
<i>Professions intermédiaires du public</i>	12,4	3,0	5,7
<i>Professions intermédiaires du privé</i>	8,2	9,5	4,1
<i>Techniciens, contremaîtres et agents de maîtrise</i>	5,1*	4,5*	3,5
Employés	10,6	3,5*	16,0
Ouvriers	2,4*	1,0*	12,4
Retraités	5,8	1,0*	26,5
Autres personnes sans activité pro.	3,7*	2,5*	17,7
Total	100	100,0	100,0

Source : INSEE, 2012 ; IFOP, 2010 *Effectifs faibles **Individus dont la profession a pu être renseignée***Uniquement des chômeurs et des femmes/hommes au foyer.

Les cadres qui publient des messages politiques sur le réseau s'apparentent à la « petite bourgeoisie nouvelle » décrite par Pierre Bourdieu dans *La Distinction*³³, ils occupent ainsi des « professions de présentation et de représentation » : attachés de presse, journalistes, chargés de communication, chargés d'études, chefs de projet, cadres commerciaux,

³² Concernant le panel non répondants, la PCS a pu être identifiée dans 32,4 % des cas. Il convient néanmoins de préciser que certaines PCS sont plus facilement identifiables que d'autres. La PCS des non répondants dont la profession a pu être identifiée ne se distribue pas de la même manière que la PCS des non répondants dont la profession n'a pas pu être identifiée. (Auto-) indiquer sa profession sur internet est en réalité une activité socialement située, qui ne se distribue pas de la même manière selon les différentes catégories socio-professionnelles. Les cadres ont ainsi plus tendance que la moyenne à indiquer leur profession sur leur profil Twitter ou sur d'autres réseaux sociaux. Leur présence sur des réseaux sociaux « professionnels » tels que LinkedIn ou Viadeo (davantage destinés aux emplois qualifiés), sur lesquels les comptes d'utilisateurs s'apparentent à des *curriculum vitae* en ligne, facilite ainsi l'identification de leur profession. En revanche, les professions intermédiaires, les ouvriers et les employés indiquent très rarement leur profession sur internet. À titre d'exemple, 74 % des cadres du panel « répondants » avaient rempli leur présentation de profil Twitter contre seulement 33 % des ouvriers.

³³ Pierre Bourdieu, *La Distinction*, op. cit.

responsables marketing, etc. À ces « professions nouvelles » des années 1960, s'ajoutent également de nouvelles professions apparues au cours des années 2000 avec l'émergence d'internet, orientées vers la production symbolique de nouveaux biens et services « numériques » : webmaster, *social community manager*, chargé de web marketing ou de marketing digital, etc. Notons enfin qu'à côté de ces professions du symbolique (numériques ou non), on retrouve parmi les classes sociales supérieures une part importante de professeurs du secondaire et du supérieur et, dans une moins mesure, de cadres administratifs d'entreprise.

Hormis leur profession, les membres de nos panels présentent d'autres propriétés sociales qui les rapprochent des populations d'ordinaires les plus politisées. La surreprésentation des hommes en est un exemple. Alors que la population internaute française se compose à part égale d'hommes et de femmes³⁴, les femmes sont en effet nettement sous-représentées au sein de notre échantillon. 59,4% des producteurs de tweets politiques qui ont répondu à notre questionnaire et 66,4% de ceux qui n'ont pas répondu sont des hommes.

Tableau 4. Distribution de la population selon le sexe (en %)

Sexe	Panel Répondants (n = 658)	Panel Non répondants (n = 590)	Pop. internet IFOP 2012 (n = 2005)	Pop. française INSEE 2012
Hommes	59,4	66,4	49,7	48,5
Femmes	40,6	33,6	50,3	51,6
Total	100,0	100,0	100,0	100,0

Sources : IFOP, INSEE

Au-delà des différences « historiques » de compétence politique entre hommes et femmes³⁵, cet écart genré de participation politique sur le réseau peut également s'expliquer par la prédominance des cadres au sein de notre échantillon. On sait en effet que chaque catégorie socio-professionnelle est définie par un sex-ratio qui lui est propre. En 2012, au sein de la population française âgée de 15 ans et plus, la catégorie des « employés » était ainsi composée de 76,6 % de femmes alors que ce taux, chez les cadres, n'était que de 40,2 %. La surreprésentation des cadres au sein de notre échantillon donne donc un élément d'explication supplémentaire à la sous-représentation statistique des femmes.

De manière analogue, la prédominance des cadres sur Twitter peut également permettre d'expliquer la surreprésentation des habitants de Paris et de son agglomération³⁶ au sein de nos panels. Alors que les Parisiens ne représentent que 3,4 % de la population française, 12,5 % des producteurs de tweets politiques qui ont répondu à notre questionnaire et 39,6 % de ceux qui n'ont pas répondu déclarent habiter à Paris *intra-muros*.

Tableau 5. Distribution des populations répondants et non répondants selon le lieu de résidence (en %)

³⁴ Sur Facebook, réseau social bénéficiant du meilleur taux de pénétration au sein de la population internaute, les femmes sont en revanche majoritaires : 53,1 % des inscrits seraient des femmes (source : IFOP, « Observatoire des réseaux sociaux », vague 2012).

³⁵ Pippa Norris « The Gender Generation Gap in British Elections », *British Elections and Parties Yearbook*, Vol. 3, Issue 1, 1993, p. 129-142.

³⁶ D'après l'INSEE, l'agglomération parisienne, ou l'aire urbaine de Paris, comprend quatre départements : Paris, les Hauts-de-Seine, la Seine-Saint-Denis et le Val-de-Marne, soit, en 2010, 412 communes et 10 460 118 habitants.

Lieu de résidence (en %)	Répondants (n = 658)	Pop. française INSEE 2012
Paris <i>intra-muros</i>	12,5	3,4
Agglo. parisienne hors Paris <i>intra-muros</i>	14,9	12,4
Province	72,6	84,2
Total	100,0	100,0

Source : INSEE

Ce résultat peut sembler de prime abord surprenant, voire paradoxale. À la différence de l'offre culturelle qui est géographiquement concentrée dans quelques grandes métropoles (dont principalement Paris), Internet et Twitter sont en effet accessibles depuis n'importe quel site du territoire géographique³⁷. La position sociale des usagers du réseau social peut alors permettre une nouvelle fois d'expliquer en partie ce paradoxe. La distribution des agents sociaux dans l'espace géographique n'est en effet jamais neutre socialement³⁸. Les agents issus de la petite bourgeoisie nouvelle et des classes intellectuelles supérieures résident ainsi plus que la moyenne des Français au sein de la capitale. Les entreprises de communication, d'études, de conseil, et plus généralement l'ensemble des « professions nouvelles » sont en effet essentiellement localisées à Paris. C'est donc la prédominance de ces professions sur Twitter qui expliquerait en grande partie cette surreprésentation parisienne.

Enfin, une dernière propriété sociologique caractérisant les individus qui publient des messages politiques sur Twitter concerne leur très haut niveau de diplôme.

Tableau 6. Distribution de la population répondants selon le niveau de diplôme (en %)

Dernier diplôme obtenu (hors lycéens et étudiants)	Panel Répondants (n = 378)
Diplôme inférieur au Baccalauréat	11,4
Niveau Baccalauréat	13,8
Deug, BTS ou équivalent Bac+2	14,8
Licence ou équivalent Bac+3	14,6
Maîtrise ou équivalent Bac+4	10,6
Master (ex. DEA ou DESS) ou équivalent Bac+5	31,2
Doctorat	3,7
Total général	100,0

Les individus qui ont accepté de répondre à notre questionnaire sont en effet nettement plus diplômés que la moyenne des Français³⁹. Cette surreprésentation ne concerne pas uniquement les individus en cours d'études, comme nous l'avons vu précédemment. Si l'on exclut cette catégorie de notre échantillon, trois quart des producteurs répondants sont diplômés de l'enseignement supérieur (74,9 %) et plus d'un tiers ont même un niveau de diplôme égal ou supérieur à Bac +5 (33,9 %).

³⁷ Même si les facilités d'accès aux infrastructures qui facilitent la navigation sur internet – le haut ou très haut débit, la fibre optique, la 4G, etc. – ne sont pas les mêmes partout sur le territoire géographique.

³⁸ Pierre Bourdieu, *La Distinction*, op. cit., p. 113.

³⁹ Notons néanmoins que la sous-représentation des catégories les plus âgées de la population dans notre échantillon explique mécaniquement cette élévation du niveau d'études, les chances d'être titulaire en 2015 d'un diplôme de Master n'étant pas les mêmes pour les plus de 65 ans et les 25-49 ans.

Les individus qui publient des messages politiques présentent donc des propriétés sociales atypiques : ils sont ainsi majoritairement des hommes, étudiants et cadres, fortement diplômés, Parisiens. Ces propriétés permettent alors d'expliquer leur très haut niveau de politisation de cette population.

Une population très fortement politisée

Les données issues de notre questionnaire permettent en effet de mesurer le niveau de politisation – particulièrement élevé – des producteurs de tweets politiques. 56 % des individus qui ont répondu à notre questionnaire déclarent ainsi « beaucoup » s'intéresser à la politique, un résultat très nettement supérieur aux résultats enregistrés par des enquêtes administrées auprès d'échantillons représentatifs de l'électorat français. Ainsi, d'après le sondage post-électoral du CEVIPOF⁴⁰, en juin 2012, seuls 21 % des électeurs français en âge de voter déclaraient s'intéresser « beaucoup » à la politique, un taux inférieur de 35 points à celui que nous avons enregistré.

Cette forte politisation des membres de notre panel s'observe également au niveau de leur participation électorale. Si, comme au niveau national, c'est lors des élections européennes de 2009 que les membres de notre panel se sont le plus abstenus, leur taux d'abstention était cependant près de 25 points inférieur à la moyenne nationale⁴¹. Alors que les électeurs français n'ont été en effet qu'une minorité à participer à ce scrutin (41%), deux tiers des producteurs de tweets politiques se sont au contraire exprimés (59,7 %).

Le niveau élevé de politisation des usagers du réseau social se manifeste également par une pratique importante d'activités politiques dites « non-conventionnelles ». Les individus qui ont répondu à notre questionnaire sont ainsi une grande majorité à avoir déjà signé une pétition (80%), manifesté (61%). Mais le résultat politiquement le plus remarquable de notre recherche réside dans la très forte proportion de militants politiques au sein de notre panel. 31 % des individus qui ont répondu à notre questionnaire déclarent en effet avoir déjà adhéré à un parti politique ou à un syndicat. Près d'un producteur de tweets politiques sur trois est, ou a été, un militant politique.

Tableau 7. Pratiques politiques des répondants

Avez-vous déjà...	Répondants (n = 658)
Signé une pétition	80,7%
Manifesté	61,1%
Fait grève	32,4%
Adhéré à un syndicat	15,8%
Adhéré à un parti politique	30,5%
Adhéré à une ou plusieurs association(s)	48,5%

⁴⁰ CEVIPOF, « Sondage post-électoral 2012 », 2012.

⁴¹ Précisons toutefois que contrairement au taux d'abstention qui est calculé au niveau national à partir d'indicateurs objectifs, le taux d'abstention des membres de nos panels a été obtenu à partir de données déclaratives formulées *a posteriori* (nous avons interrogé nos enquêtés en 2012 sur les votes de 2007 à 2012), ce qui n'est pas sans poser problème.

Au-delà de ce niveau très élevé de politisation, les producteurs de tweets politiques se caractérisent également par une identité politique relativement homogène : l'électorat de gauche est prédominant sur Twitter. 63 % des membres de notre panel qui ont voté lors du premier tour des élections présidentielles du 22 avril 2012 ont choisi un candidat de gauche ou d'extrême-gauche.

Tableau 8. Vote au premier tour des élections présidentielles du 22 avril 2012

Vote premier tour présidentielle 2012	Population Répondants et votants (n = 491)	Suffrages exprimés (Ministère de l'intérieur)
François Hollande (PS)	34,2 %	28,6 %
Jean-Luc Mélenchon (Front de gauche)	20,8 %	11,1 %
Nicolas Sarkozy (UMP)	19,1 %	27,2 %
François Bayrou (MODEM)	13,2 %	9,1 %
Éva Joly (EELV)	6,7 %	2,3 %
Marine Le Pen (FN)	3,1 %	17,9 %
Nicolas Dupont-Aignan (DLR)	1,8 %	1,8 %
Philippe Poutou (NPA)	0,8 %	1,1 %
Jacques Cheminade (Solidarité et Progrès)	0,2 %	0,3 %
Nathalie Arthaud (Lutte ouvrière)	0,0 %	0,6 %
Total	100,0 %	100,0 %

Source : Ministère de l'intérieur

La surreprésentation de l'électorat de Jean-Luc Mélenchon est un exemple particulièrement manifeste de cette prédominance de la gauche : le candidat du Front de Gauche enregistre en effet un score deux fois supérieur (20,8 %) à sa moyenne nationale (11,1 %). Et inversement, l'électorat de droite, et plus encore celui d'extrême-droite, sont nettement sous-représentés sur Twitter. Seuls 3 % des producteurs de tweets politiques qui ont répondu à notre questionnaire déclarent avoir voté pour Marine Le Pen. Moins de 5 % des répondants se situent d'ailleurs « très à droite » et seuls 6 % se définissent comme étant « ni de gauche, ni de droite »⁴². Les propriétés sociologiques spécifiques des membres de nos panels (issus des factions les plus dotées en capitaux culturels) permettent alors d'expliquer en partie cette prédominance de l'électorat de gauche sur Twitter.

Les individus qui publient des messages politiques sur Twitter sont donc le produit d'une très forte sélection sociale. Néanmoins, l'analyse longitudinale des panels laisse entrevoir qu'à l'intérieur même de cette population socialement et politiquement homogène, s'observent de fortes inégalités en termes de production de messages.

⁴² D'après le sondage post-électoral du CEVIPOF, ce sont 22 % des Français qui ne se déclarent « ni de gauche, ni de droite ».

L'impact des variables sociologiques et contextuelles sur la production de messages politiques

Du 1^{er} mars 2013 au 31 janvier 2014, les membres du panel « répondants » ont publié 840 251 tweets, dont 81 606 messages contenant des occurrences politiques⁴³, soit 9,7 % des tweets publiés. La distribution statistique de ces messages laisse clairement entrevoir de fortes inégalités de production. Ces inégalités s'observent dans le temps – le nombre de tweets politiques publiés par le panel est très inégal d'un jour à l'autre – mais également entre individus.

Les facteurs individuels de production de tweets politiques

Le nombre de messages politiques publiés par les usagers du réseau social est en effet très inégal d'un individu à l'autre. Sur l'ensemble de la période, 3% des membres du panel ont ainsi produit à eux seuls plus de la moitié des messages politiques publiés.

Tableau 9. Statistiques relatives à la production de messages du panel (n=558⁴⁴)

	Tweets politiques
Effectif	81 606
Minimum	0
Maximum	8 128
Moyenne	146
Médiane	14
Quartile 1	3
Quartile 3	74
Ecart type	569
Variance	323 282

Le nombre de messages politiques publiés s'avère donc être une variable très inégalement distribuée dans la population. L'analyse statistique permet alors d'isoler plusieurs variables permettant d'expliquer ces inégalités. Une première variable concerne ainsi l'âge des individus. On observe en effet une corrélation significative entre l'âge des usagers et leur nombre total de tweets politiques publiés.

⁴³ Nous avons arrêté une liste de 121 termes politiques, comprenant le nom des principaux partis, institutions, et personnalités politiques.

⁴⁴ Au 1er mars 2013, soit un an après qu'ils aient répondu à notre questionnaire, 8 % des répondants avaient supprimé ou désactivé leur compte Twitter. A cette date, le panel est alors composé de 558 individus.

Tableau 10. Nombres moyen de tweets politiques publiés par classes d'âge

	<i>n</i> =	Nombre de tweets moyen
Moins de 18 ans	43	25
18 à 25 ans	230	78
26 à 34 ans	108	75
35 à 45 ans	103	216
46 à 60 ans	52	229
61 ans et plus	22	921
Total	558	146

Le nombre de tweets politiques publiés croît ainsi en fonction de l'âge. Si nous avons vu que les catégories les plus âgées de la population étaient nettement sous-représentées au sein de notre population (les plus de 61 ans représentent 4 % des « répondants »), elles sont pourtant celles qui produisent en moyenne le plus de tweets politiques. Dit autrement, les plus de 61 ans sont proportionnellement les moins nombreux à produire des tweets politiques, mais la minorité qui en produit est celle qui en produit le plus en moyenne. C'est donc parmi les catégories les plus âgées que l'on retrouve les plus importants producteurs de messages. Ces individus semblent réunir simultanément quatre propriétés remarquables, qui expliquent leur importante production de messages : un âge relativement avancé (par rapport à l'âge moyen des internautes, qui était de 41 ans en 2012⁴⁵), une absence de contraintes professionnelles et familiales, et donc un temps libre plus important, un fort niveau de politisation et une maîtrise – souvent autodidacte – d'internet et de l'outil informatique.

Une deuxième variable significativement corrélée au nombre de messages politiques publiés concerne le niveau de politisation des enquêtés. On enregistre ainsi une corrélation entre le niveau d'intérêt déclaré pour la politique et le nombre de tweets politiques publiés. Les « répondants » qui déclarent « beaucoup » s'intéresser à la politique ont publié en moyenne huit fois plus de tweets politiques que ceux qui s'y intéressent « peu » ou « pas du tout » et quatre fois plus que ceux qui s'y intéressent « assez ».

Tableau 11. Intérêt déclaré pour la politique et nombre de tweets politiques publiés (Test de Spearman)

Variable	Groupe	<i>n</i> =	Nombre de tweets moyen	Test <i>t</i>	valeur de <i>p</i>	$p < 0,05 = *$
Nombre de tweets politiques publiés	Un peu ou pas du tout	72	29,14	0,272	0,000	*
	Assez	171	53,59			
	Beaucoup	315	223,31			
	Total	558	146,25			

Les individus qui s'intéressent le plus à la politique sont ceux qui publient le plus de messages politiques. Si ce résultat ne semble guère surprenant, il est néanmoins intéressant de

⁴⁵ Vincent Gombault, « L'internet de plus en plus prisé, l'internaute de plus en plus mobile », *INSEE Première*, n°1452, juin 2013.

constater que l'intérêt pour la politique, qui est une variable déclarative, détermine objectivement le nombre de messages politiques publiés. Dans ce cas de figure, les données issues de l'observation corroborent bien les données générées par le biais de l'interrogation. De manière analogue, le fait d'adhérer à un parti politique détermine également le nombre de tweets politiques produits. Les militants ont ainsi publié en moyenne sur la période deux fois plus de tweets politiques (225 tweets, soit 20 tweets par mois) que les non militants (109 tweets, soit 10 tweets par mois).

Tableau 12. Comparaison du nombre de tweets politiques publiés entre militants et non militants ((Test de de Spearman)

Variable	Groupe	n =	Nombre de tweets moyen	Test t	valeur de p	p < 0,05 = *
Nombre de tweets politiques publiés	Militants	178	225,30	2,256	0,024	*
	Non militants	380	109,22			
	Total	558	146,25			

Enfin, une troisième variable influençant significativement le nombre de messages politiques publiés concerne le niveau d'activité des usagers sur le réseau social. Ce niveau d'activité peut se mesurer à partir de plusieurs indices, comme notamment le nombre total de messages publiés. Le nombre total de tweets publiés comptabilise l'ensemble des messages, politiques ou non, publiés par les enquêtés. Un nombre élevé de messages peut alors laisser supposer que l'utilisateur est très actif sur le réseau et a ainsi l'habitude de publier régulièrement des messages et d'échanger avec d'autres usagers du réseau. Cependant, postuler que le nombre total de tweets publiés détermine le nombre de tweets politiques peut paraître tautologique de prime abord, les tweets politiques étant nécessairement inclus dans le nombre total de tweets publiés. Cependant, l'étude approfondie de notre panel nous permet d'affirmer que la réciproque n'est pas vraie : publier un grand nombre de tweets n'implique pas nécessairement de publier un grand nombre de tweets politiques.

Le nombre total de tweets publiés s'avère alors être la variable la plus fortement corrélée au nombre de tweets politiques produits. Tout se passe comme si publier régulièrement des tweets était la condition nécessaire (mais non suffisante) à la production de tweets politiques. Être « bavard » politiquement sur Twitter suppose donc avant tout d'être « bavard » de manière générale sur le réseau social.

Tableau 13. Nombre total de tweets et nombre de tweets politiques (Test de corrélation de Pearson)

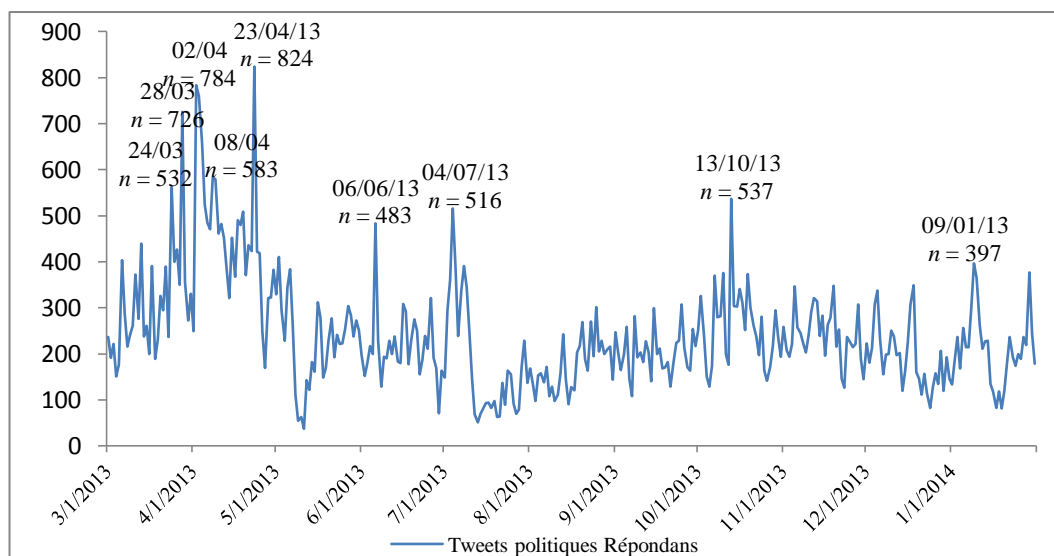
Population	Variable	n =	r	valeur de p	p < 0,05 = *
Répondants	Nombre de tweets politiques	558	0,437	0,000	*
	Nombre total de tweets				
Non répondants	Nombre de tweets politiques	517	0,309	0,000	*
	Nombre total de tweets				

L'âge, le niveau de politisation et le niveau d'activité des usagers sur le réseau social sont donc des variables qui permettent d'expliquer les fortes inégalités individuelles de publication de messages politiques.

Les facteurs conjoncturels de production de tweets politiques

Cependant, si le nombre de tweets politiques publiés est très inégal d'un individu à l'autre, il est également très inégal d'un jour à l'autre. Le nombre de tweets politiques, comme le nombre d'individus qui les publient, varie en effet significativement selon la date de publication des messages sur le réseau social.

Graphique 1. Distribution journalière des tweets politiques publiés par le panel « répondants » du 01/03/2013 au 31/01/2014



Comme dans les recherches anglophones évoquées plus haut, on peut alors observer à la lecture de ce graphique la présence de très nets pics de publication de messages politiques. On peut remarquer d'une part que ces pics, à l'exception de ceux parus au mois d'avril, sont relativement espacés dans le temps et qu'ils sont d'autre part très « brefs » : le volume de tweets chute dans la plupart des cas dès le lendemain de la formation du pic. Nous avons en particulier isolé huit pics de production et consulté des échantillons de tweets parus ces jours-là.

Tableau 14. Volume et taux de production quotidiens les plus élevés du panel « répondants »

Date	Nombre de tweets	% Du panel	Événement
23 Avril 2013	824	27	Adoption définitive de la loi sur le mariage pour tous par l'Assemblée nationale
2 Avril 2013	679	25	Jérôme Cahuzac avoue avoir fraudé le fisc
28 Mars 2013	726	19	Interview télévisée de François Hollande au journal de 20H de France 2
8 Avril 2013	583	19	Publication des déclarations de patrimoine de certains ministres
6 Juin 2013	483	19	Mort de Clément Méric, un jeune militant d'extrême gauche

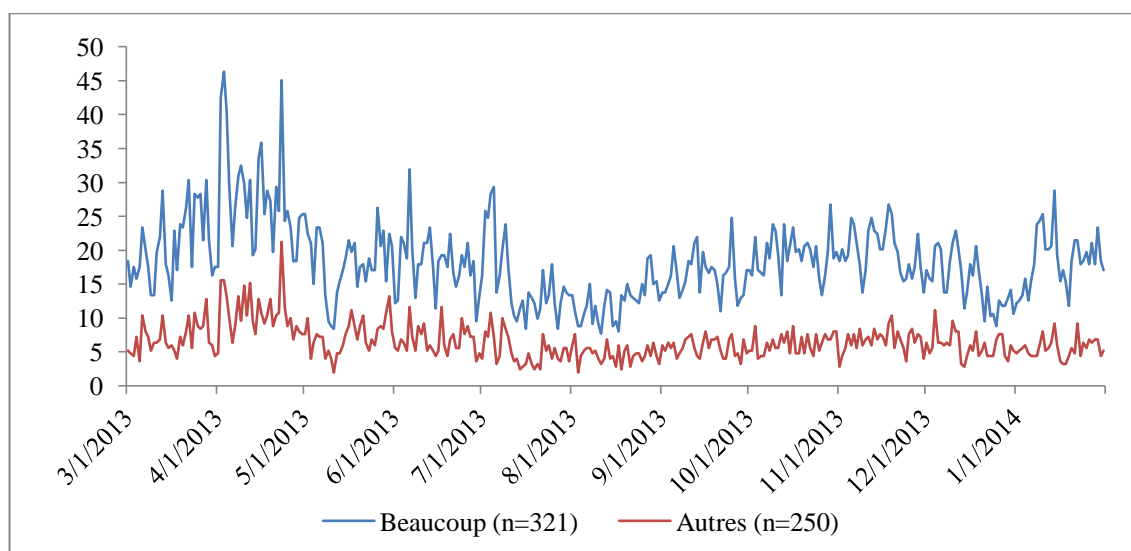
24 Mars 2013	562	17	« Manif pour tous » : 300 000 à 1,4 M de participants + Législative partielle UMP/FN Oise
4 Juillet 2013	516	17	Comptes de campagne de Nicolas Sarkozy invalidés
9 Janvier 2014	397	15	<i>Closer</i> révèle la liaison Hollande/Gayet

Chacun de ces pics correspond en fait à un « événement » politique qui a occupé « le haut de l’agenda médiatique » (scandales, révélations sur la vie privée du chef de l’Etat, loi sur le mariage homosexuel, manifestations en opposition à cette loi). En France également, les sujets de discussion politique des usagers du réseau social semblent en partie définis par les médias *mainstream*. Le nombre de tweets politiques publiés dépend ainsi fortement du contexte politique et médiatique.

On peut alors se demander quelles sont les catégories d’usagers les plus sensibles à ces « événements » politiques ? Est-ce les plus politisées, qui auraient une réelle appétence pour des « événements » politiques *extra-ordinaires* ? Ou au contraire les usagers les moins politisées, qui ne s’exprimeraient politiquement que lorsque le temps politique se synchronise avec le temps social⁴⁶ ?

Si l’on observe l’évolution comparée du taux de publication quotidien des « répondants » qui déclarent s’intéresser « beaucoup » à la politique avec celui des autres « répondants » qui déclarent s’y intéresser « assez, un peu ou pas du tout », on s’aperçoit que les deux courbes sont relativement symétriques. Autrement dit, à des échelles différentes, les inégalités journalières de publication de tweets politiques sont les mêmes parmi les catégories les plus politisées comme les moins politisées du panel (qui restent, rappelons-le, largement plus politisées que la moyenne des Français).

Graphique 4.2 Évolution quotidienne du taux de publication des « répondants » selon leur intérêt déclaré pour la politique



⁴⁶ Vincent Goulet, *Médias et classes populaires. Les usages ordinaires des informations*, INA Éditions, Paris, 2010.

Les événements « politico-médiatiques » semblent donc générer des niveaux inhabituels de publication de messages parmi les individus les plus politisés comme parmi les individus les moins politisés du panel. En résumé, le niveau de messages politiques publiés sur Twitter dépend à la fois de variables individuelles (l'âge, le niveau de politisation, le niveau d'activité sur le réseau) et bien plus sensiblement encore du contexte politique et médiatique.

Conclusion

Combiner une approche de *web-tracking* avec une approche classique d'interrogation par questionnaire permet donc de tirer profit des potentialités méthodologiques offertes par le Web tout en permettant au chercheur de situer très précisément les données collectées dans l'espace social « réel ». Ce dispositif méthodologique nous a ainsi permis de mettre en lumière la très forte sélection sociale dont les usagers de Twitter sont le produit, ainsi que l'influence des variables sociologiques sur les pratiques et habitudes de publication des usagers. De futures recherches questionneront l'impact des variables sociologiques sur la forme et le contenu des messages collectés.

Bibliographie

Anders Larsson et Hallvard Moe, « Twitter in politics and elections: Insights from Scandinavia », in Katrin Weller, Axel Bruns, Jean Burgess, Marja Mahrt et Cornelius Puschmann (dir.), *Twitter and Society*, New York, NY: Peter Lang Publishing, 2013, p. 139-330.

Andreas Jungherr, « The logic of political coverage on Twitter: Temporaldynamics and content », *Journal of Communication*, Vol. 64, n°2, 2014, p. 239-259.

Andreas Jungherr, « Tweets and votes, a special relationship: the 2009 federal election in Germany », in Ingmar Weber, Ana-Maria Popescu et Marco Pennacchiotti (dir.), *Proceedings of the 2nd workshop politics, elections and data*, New York, NY: ACM, p. 5-14.

Andreas Jungherr, « Twitter in Politics: A Comprehensive Literature Review », Disponible en ligne : SSRN: <http://ssrn.com/abstract=2402443>.

Anne Muxel, *L'expérience politique des jeunes*, Paris, Presses de Science Po, 2001.

Arnaud Mercier, « Twitter l'actualité : usages et réseautage chez les journalistes français », *Recherches en communication*, 39, 2013.

Dang-Xuan, Linh, Stefan Stieglitz, Jennifer Wladarsch et Christoph Neuberger, « An investigation of influentials and the role of sentiment in political communication on Twitter during election periods », *Information, Communication & Society*, 16(5), 2013, p.795-825.

Daniel Gaxie, « Retour sur les modes de production des opinions politiques », in Philippe Coulangeon et Julien Duval (dir.), *Trente ans après La Distinction de Pierre Bourdieu*, Paris, La Découverte, 2013.

Daniel Gaxie, *Le cens caché. Inégalités culturelles et ségrégation politique*, Paris, Édition du Seuil, 1978.

- Daniel Gayo-Avello, « A meta-analysis of state-of-the-art electoral prediction from Twitter data », *Social Science Computer Review*, 2013, vol. 31 n°6, p. 649-679
- Daniela Dimitrova, Adam Shehata, Jesper Strömbäck et Lars Nord, 2014. « The effects of digital media on political knowledge and participation in election campaigns: Evidence from panel data », *Communication Research*, 41(1), 2014, p. 95-118.
- Dominique Cardon *et alii*, « Présentation », *Réseaux*, 3/2006 (n° 137), p. 9-25.
- IFOP, « Observatoire des réseaux sociaux », 2013.
- Jason Gainous et Kevin Wagner, *Tweeting to Power: The Social Media Revolution in American Politics*, Oxford, Oxford University Press, 2014.
- Julian Ausserhofer et Axel Maireder, « National politics on Twitter: structures and topics of a networked public sphere », *Information, Communication & Society*, 16(3), 2013, p. 291-314.
- Jungherr, Andreas et Pascal Jürgens, « Through a glass, darkly: tactical support and symbolic association in Twitter messages commenting on Stuttgart 21 », *Social Science Computer Review*, 32(1), 2014, p.74-89.
- Lawrence Ampofo, Nick Anstead et Ben O'Loughlin, « Trust, confidence, and credibility: Citizen responses on Twitter to opinion polls during the 2010 UK General Election », *Information, Communication & Society*, 14(6), 2011, p. 850-871.
- Lee Rainie, Aaron Smith, Kay Schlozman, Henry Brady et Sidney Verba, « Social media and political engagement », *Pew Internet & American Life Project*, 2012.
- Marija Bekafigo et Allan McBride, « Who tweets about politics? Political participation of Twitter users during the 2011 gubernatorial elections », *Social Science Computer Review*, 31(5), 2013, p. 625-643.
- Max McCombs et Donald Shaw, « The agenda-setting function of mass-media », *Public Opinion Quarterly*, n°36, 1972, p. 176-187.
- Nick Anstead et Ben O'Loughlin, « The emerging Viewertariat and BBC Question Time: television debate and real-time commenting online », *The International Journal of Press/Politics*, 16(4), 2011, p. 440-462.
- Pierre Bourdieu, « L'opinion publique n'existe pas », *Les Temps modernes*, n°318, janvier 1973, p. 1292-1309.
- Pierre Bourdieu, Jean-Claude Chamboredon et Jean-Claude Passeron, *Le métier de sociologue. Préalables épistémologiques*, Paris, Éditions de l'EHESS, 1968.
- Pierre Bourdieu, *La Distinction. Critique sociale du jugement*, Paris, Les Éditions de Minuit, 1979
- Pippa Norris « The Gender Generation Gap in British Elections », *British Elections and Parties Yearbook*, Vol. 3, Issue 1, 1993, p. 129-142.
- Valérie Jeanne-Perrier, « Parler de la télévision sur Twitter : une réception oblique à partir d'une conversation médiatique », *Communication & langages*, n°166, 2010, p. 127-147.

Vincent Gombault, « L'internet de plus en plus prisé, l'internaute de plus en plus mobile », *INSEE Première*, n°1452, juin 2013.