Mining Political Opinion on Twitter: Challenges and Opportunities of Multiscale Approaches

Marta Severo, Robin Lamarche-Perrin

Mining Political Opinion on Twitter
Challenges and Opportunities of Multiscale Approaches

Marta Severo

EA Dicen – Université Paris Nanterre-IUF

msevero@parisnanterre.fr

Robin Lamarche-Perrin

ISC-PIF – LIP6 – CNRS

Robin.Lamarche-Perrin@lip6.fr

Acknowledgements

A French version of this paper has been published under the title of "L'analyse des opinions politiques sur Twitter. Défis et opportunités d'une approche multi-échelle" in *Revue française de sociologie*, 2018/3 (Vol. 59), p. 507-532. DOI 10.3917/rfs.593.0507

This work has been partially funded by the European Commission H2020 FETPROACT 2016–2017 program under grant 732942 (ODYCCEUS).

Abstract

Social research on public opinion has been affected by the recent deluge of new digital data on the Web, from blogs and forums to Facebook pages and Twitter accounts. This fresh type of information useful for mining opinions is emerging as an alternative to traditional techniques, such as opinion polls. Firstly, by building the state of the art of studies of political opinion based on Twitter data, this paper aims at identifying the relationship between the chosen data analysis method and the definition of political opinion implied in these studies. Secondly, it aims at investigating the feasibility of performing multiscale analysis in digital social research on political opinion by addressing the merits of several methodological techniques, from content-based to interaction-based methods, from statistical to semantic analysis, from supervised to unsupervised approaches. The end result of such an approach is to identify future trends in social science research on political opinion.

1. Introduction

The study of political opinion is a traditional field in social sciences. Researches related to elections, political parties and representatives have been filling the pages of social sciences journals for the last century. They often rely on empirical analysis based on traditional methods such as opinion polls, surveys, focus groups, or interviews. More recently, the development of the World Wide Web has been responsible for important changes in this field by offering new arenas for expression for both politicians and citizens. From the exemplary case of Barack Obama until the more recent use of big data in the Trump-Clinton competition, the Internet has become a key tool to convince voters, organise supporters and spread political messages. If ten years ago, websites, blogs and forums were the main spaces of political exchanges, today Facebook, Twitter and YouTube seem to be the strategic places for expressing and disseminating political opinions and also for debating with candidates, among friends or with strangers about political topics. All these online interactions between politicians and citizens, but also between journalists, party representatives and other kinds of influencers, generate a large amount of digital data[1], commonly called big data[2], that have been recently introduced in social research in order to studying political opinions of individuals and groups. According to numerous scholars, techniques for mining opinions on social media can be used as an alternative to traditional methods.

Facing such situation, this paper aims at studying the impact of the use of social media data on the social research on political opinion, by considering the specific case of studies based on Twitter data. When compared to other social media, Twitter has received a lot of attention from scientists both because of how easier it seems to make the collection of data[3], and because of the wide scope of topics covered by online exchanges. Those studies vary according to the size of corpora, from small to big data, to the type of data, from content- to interaction-based approaches, and of methods. As regards to methods, it is generally assumed that research in this

---

1 If Web data are surely an important novelty for social research, their use is not so straightforward and calls for a critical approach (Hogan, 2014). In particular, the technical and commercial issues related to data access, the legal issues related to privacy and copyright and, last but not least, the issues of veracity and representativeness of these data don't have to be overlooked (Severo et al., 2016).

2 In this paper, big data generally refers to "datasets that are large in both size and complexity, with which new algorithmic techniques are required in order to extract useful information from them" (Holmes, 2017, p. 7). This is why we later refer to the challenge of multiscale analyses of big data, which is the ability in computer science to apply classical algorithms to larger problems. Hence, we are not interested here in limitations regarding databases management, that relate more to storage constraints than to analysis constraints.

3 Twitter provides APIs for collecting tweets. These APIs have several limitations about the size, the duration and the type of corpora that can be collected. Nevertheless, they, being freely available, have drawn a lot of attention of scholars.

field needs to go beyond the opposition between quantitative and qualitative techniques[4] and find new solutions for the "redistribution of methods" (Marres, 2012). While big data analyses have encountered several difficulties in the assessment of their validity, evidence based on small data[5] has appeared hard to be applied to larger scales. This is why scientists are currently assessing new mixed methodological solutions in order to develop "quali-quantitative" approaches (Venturini and Latour, 2010).

Considering the variety of these studies, our first goal is to outline the state of the art in the research concerned with this field. The merits of completing a state of the art of such recent and vast literature not only lie in the possibility to compare case studies and methodological solutions but foremost in the exploration of the different definitions of political opinion used in these studies. Eventually, it will be interesting to see whether the use of new data coming from social media, and especially the exploitation of large corpora considered by the authors of the studied papers as big data, makes possible to build a genuine new approach for studying political opinions able to investigate new dimensions of this object or whether the approach to political opinion of these studies have nothing different from researches obtained with traditional techniques. So, our ultimate goal is not to monitor a comparative study of empirical solutions but to disclose the (implicit or explicit) theory that these studies entail.

Then, based on the state of the art, our second goal is to verify if a new approach in social research on political opinion could be based on multiscale techniques. Indeed, in the analysed corpus, multiscale approaches have emerged as the main novelty compared to traditional statistical techniques. Differently from statisticians, some data scientists[6] promise to social

4 This opposition has been defined in several ways. A very simple definition can be that: "quantitative research is empirical research where the data are in the form of numbers… qualitative research is empirical research where the data are not in the form of numbers" (Punch, 1998, p. 4). Here, with these terms, we try to identify to two different approaches to social research. Quantitative methods are used to quantify the research object by way of generating numerical data or data that can be transformed into usable statistics. Qualitative methods, that are generally based on unstructured or semi-structured techniques (focus groups, interviews, observation, etc.), are mainly exploratory and are meant to study reasons, opinions or motivations.

5 With small data, we refer to data that is "small" enough for human comprehension. It is data in a volume and format that makes it accessible, informative and actionable. This kind of data has proved to be very useful for empirical research. Yet, several methodological issues are raised when trying to apply methods and verify evidence obtained on small data through the analysis of big data corpora.

6 Data science is used as a "concept to unify statistics, data analysis, machine learning and their related methods" in order to "understand and analyse actual phenomena" with data (Hayashi, 1998). However, several scholars observed that there is no difference between data scientists and statisticians. Here, we use the term in order to identify a more interdisciplinary approach to data

scientists to be able to reproduce similar analyses at different scales, from small to big and vice versa, without loosing in quality. Considering that, our purpose it to investigate the feasibility of performing multiscale digital social research on political opinion, in order to apply qualitative approaches to larger corpora or, even more challenging, to exploit big data approaches while simultaneously meeting the epistemological requirements of qualitative analysis. At the individuals' level, the automatic measurement and quantification of political opinion from social media data are challenged by the limited content of micro-blogging exchanges, by its often ambiguous, ironic, or paradoxical nature, and by the multimedia nature of a message that can include images, URLs, and emoticons. At the system's level, the analysis of activity traces of several million individuals requires algorithmic breakthroughs so that traditional methods are efficient at larger scales. This paper thus proposes to address the merits of general methodological options for the multi-scale analysis of political opinions on Twitter, from content-based to interaction-based methods, from statistical to semantic analysis, from supervised to unsupervised[7] approaches. Regarding the latter opposition, we will discuss in detail the possible statuses of qualitative knowledge within quantitative analysis, whether they are responsible for *a priori* formalising the problem (supervised approaches) or *a posteriori* interpreting the results (unsupervised approaches), thus constituting a practical instance of the classical epistemological opposition between deduction and induction. More generally, current challenges of machine learning, that is the use of artificial intelligence to get computer to "learn" from data, to infer new models, and to act without being explicitly programmed, will be addressed in this paper as practical re-actualisations of classical issues within the realm of digital data. If these questions are not new, big data makes them actual.

As a final result, we intend to identify future trends in social sciences research on political opinion and to support the adoption of a solid sociological methodology, which will preserve empirical sociology from the "oncoming crisis" (Savage and Burrows, 2007).

2. Political opinion on Twitter: a state of the art

2.1. Political Opinion and Opinion Polls

Before moving on to Twitter, it may be useful to ponder the various literature on opinion polls since this approach could prove relevant in several studies involving Twitter. If the use of polls dates back to the end of the 19th century (Cayrol, 2011), their popularity really kicked off in 1936

---

including statistics but also mathematics, information science, and computer science and using automated algorithms.

7 Supervised approaches are based on the *a priori* formalisation of expert knowledge, for example through the manual annotation of data, while unsupervised approaches focus on the *a posteriori* interpretation of statistical results, typically performed on unlabelled data.

when George Gallup was able to predict Roosevelt's victory. In France, Jean Stoetzel established the first opinion polls in 1938 when he founded the *Institut francais d'opinion publique*. Ever since polls were created, measuring public opinion has gone hand in hand with organizing opinion polls. Polls are based on the principle that the opinion of a sample chosen among carefully selected likely voters can be representative of the opinion of the greater crowd. As noted by Loïc Blondiaux (1998), for many years, the success of polls has been responsible for overlooking the value of opinions from a theoretical point of view.

In the last fifty years, scholars have developed controversial positions about the validity of polls. Particularly famous is the essay written by Pierre Bourdieu (1973) in which he challenges views on opinions by stating that opinions do not exist so as to underline that actual opinions differ from what is measured by opinion polls. Bourdieu claims that opinion, as measured in polls, is an artefact created by researchers, who therefore input their own questions, and possibly their own answers. Conversely, Page and Shapiro (1992) consider opinions as real, measurable, and rational. In their view, it is possible to rely solely on survey data as referents for public opinion. Roland Cayrol (2011) underlines that the purpose of polls is not to know the opinions of individuals but the aggregate public opinion. Loïc Blondiaux (1998) proposes an intermediary position. According to him, even if polls are useful for studying public opinion, it does not mean that opinion and polls are the same thing. Similarly, Ginsberg (1986) underlines that polls' answers are the result of the interaction between an opinion and a query tool: therefore polls do change opinions.

This condensed review of different positions on opinion polls is instrumental in highlighting two decisive points that will be useful for our subsequent analysis. First, there is confusion between the concept of aggregate public opinion (where "opinion" is always singular) and the one of individual opinions (in the plural). Taking one approach or the other has important consequences on the empirical techniques that come useful for analysing political opinion. Therefore, if we look at it from the viewpoint of social statistics (Reynié, 1989) and consider that one opinion equals another and that, consequently, public opinion is an aggregate phenomenon, it means that opinion polls are to be considered as an effective technique of measurement. Conversely, in the case where we consider opinions as consistently related to the individual, we will need methods that are able to measure the contribution of each person and the interaction between each other. Yet, if there is almost general agreement that opinion polls only allow for investigating the aggregate dimension of the opinion, literature on the subject is not so straightforward in proposing alternatives to polls when investigating individual opinions. This brings us to our second point: the importance played by the researcher's choices and the adopted method. It appears quite clear that studying the definition of political opinion should start precisely from the analysis of empirical choices, that is to say which data are selected and how they are treated in order to understand how the researcher is framing public opinion. This is precisely our purpose in the next subsection.

2.2. A Classification of Public Opinion

Our research calls for the identification of theoretical views on political opinion in the literature based on Twitter data. Yet, this is not an easy task, especially because scholars focus on empirical investigation and rarely explicitly define how they proceeded when framing the concept. For such a reason, as a preliminary step, it is useful to summarise the classification proposed by Robert Entman and Susan Herbst (2001) that identify four forms of framing public opinion based on methods used by scholars in the analysis of public opinion.

The first form is "mass opinion". According to this form, undoubtedly the most popular, public opinion can be defined as "the aggregation or summation of individual preferences as tabulated through opinion polls, referenda or elections". Within this form, opinions are not "reflective of thoughtful, informed citizens". They are rather artefacts of the tool used to collect them. The second form of public opinion is the "latent public opinion", which "underlies more fleeting and superficial opinions we find when conducting polls of the mass public". This deeper preference depends on individual considerations and political predispositions, that is to say "stable, individual-level traits that regulate the acceptance or non-acceptance of the political communications that people receive" (John R. Zaller, 1992, p. 22). The third form, called the "activated public opinion", refers to opinion "of engaged, informed, and organized citizens – those who are mobilizable during campaign periods and between elections as well". Within this form, we can clearly include all studies that insist on the role of opinion leaders and "influentials" (Katz and Lazarsfeld, 1955; Lazarsfeld et al., 1948; Merton, 1968), but also of the media (Watts and Dodds, 2007) in the formation of political opinions. The fourth form of public opinion, called "perceived majorities" [8], refers to the situations where the term opinion indicates "the perceptions held by most observers, including journalists, politicians, and members of the public themselves, of where the majority of the public stands on an issue". As opposed to previous forms, opinions here do not correspond to the preferences of people whether considered aggregately or individually, but rather to the representations of opinion produced by the media. They are not the actual sentiment but what the media reports. Consequently, in this case, studies do not try to evaluate people's opinion through polls or other techniques, but focus on the analysis of mediated opinion. This form of public opinion mainly includes studies focusing on agenda-setting.

8 If we apply the distinction between 'opinion' and 'opinions' to the four forms, we can easily see that the "mass opinion" and "predictive majorities" correspond to a collective view of the opinion, while the "latent opinion" and the "activated opinion" are based on the idea of multiple individual opinions.

2.2. Mining Political Opinion on Twitter

In hundreds of papers[9] that try to analyse political opinion on Twitter, most of them focus on the use of Twitter for predicting voting result and that consider tweets as an alternative to traditional polls. It may be worthwhile to establish whether studies with similar research questions also share a similar theoretical view on political opinion.

As a first observation, we note that papers can be divided into two groups. The first group is constituted by scholars who offer a positive vision by validating the use of Twitter as a vote predictor. This is the case in the famous paper by Tumasjan et al. (2010, cited 1487 times according to Google Scholar, but absent in Scopus), which demonstrates that with the help of a sentiment analysis[10] of tweets, it would have been possible to predict the result of German federal elections in 2009 (based on ca. 100,000 tweets). The authors state that "the mere number of tweets mentioning a political party can be considered a plausible reflection of the vote share and its predictive power even comes close to traditional election polls". It has to be noted that the authors define political opinions as political sentiments, yet, even if they use sentiment analysis, sentiment is reduced to a very basic computation of mentions and conclusions are based on the numerical equivalence between votes and tweets. Similarly, Livne et al. (2011) elaborate a proof of concept for predicting election results with an accuracy of 88% (based on 460,000 tweets). Their prediction method is based on the evaluation of several indicators derived not only from text mining but foremost from structural data based on network analysis. The position defended by O'Connor et al. (2010) seems more cautious, since they argue that Twitter can be used for pulsing real-time opinion but that such a tool cannot have a prediction power comparable to that of traditional polls. In their paper, we find – for the first time in this field – the use of sentiment analysis techniques, even if they remain basic at that point since they simply relay on lexicons to determine the polarity of tweets (based on 1 billion of tweets).

The second group is constituted by scholars that reject or relativise the use of Twitter as a vote predictor. Some of them concentrate their efforts on rejecting previous studies. The authors' criticism generally focuses on the demographic difference between Twitter users and likely voters, and therefore on the consequent non-representativeness of samples induced by Twitter data. Metaxas et al. (2011) propose the formula "predicting the present" by underlying that researchers who predict election results are doing so after the elections, when they already know the result (based on ca. 235,000 tweets). Daniel Gayo-Avello (2013) draws a state of the art of the question by highlighting weaknesses in previous studies (especially those related to the reproducibility of results). Jungherr et al. (2011) invalidate the study by Tumasjan et al. (2010) by showing that the selection of parties and the chosen particular timeframe have a strong influence on such empirical

9 The state of the art is based on the analysis of the 70 most cited papers according to Scopus and Google Scholar corresponding to the query "political opinion AND twitter" (Annexe 2).

10 The term « sentiment analysis » is used to identify a group of varied techniques aiming at extracting and analysing affective states in texts.

results. Skoric et al. (2012) show that there is certain correlation between Twitter chatter and votes, but not enough to make accurate predictions. A slightly different critique can be found in Boyadjian (2014) who does not focus on the limit of empirical analysis but rather investigates the more general question of the representativeness of Twitter compared to traditional polls.

Some other scholars focus on the limits of the chosen methods by offering more advanced solutions. Some papers suggest the use of machine learning for performing sentiment analysis. Investigating the 2011 Irish General Election (32,578 tweets), Bermingham and Smeaton (2011) demonstrate the validity of using Twitter as a predictor through an approach combining sentiment analysis based on supervised learning and volume-based measures. Regarding supervised learning, the authors declare: « We instructed annotators not to consider reporting of positive or negative fact as sentiment but that sentiment be one of emotion, opinion, evaluation or speculation towards the target topic » (p. 5).

When considering this whole range of papers, we observe that no one investigates the equivalence between tweets and opinions. The fact that these short messages allow us to know people's opinions is somehow legitimate. Even scholars in political science, who are very critical about the predictive power of the platform, focus on choices made for empirical analysis without questioning the relation between the data and the studied object. Yet, by focusing on the selected data and methods, it is possible to distinguish four conceptual models of political opinion, which are not exclusive and can also be combined in the same study (Table 1).

Conceptual model	**Eldam and Herbst's form**	**Data**	**Methods**
Preference	Mass opinion	Tweet as a unit	- Statistics - Basic sentiment analysis (lexicon)
Sentiment	Latent opinion	Tweet as a content	- Advanced sentiment analysis (unsupervised and supervised learning)
Interaction	Activated opinion	Tweet as an interaction	- Network analysis
Agenda	Perceived majorities	Tweet as a medium (agenda setter)	- Discourse analysis - Text-mining

Table 1. The four conceptual models of political opinion.

3.3.1. Opinion as a Preference

In the first group of studies (*preference*), political opinion is considered as an aggregate preference in relation to a determined object selected by the researcher. Researchers adopting this model are building corpora of tweets containing specific keywords or hashtags, such as the name of a candidate or a political party. They are interested in the tweet as a unit, as a whole (without considering co-occurrences inside it or the context of usage) and they are observing mainly the variation of volume of tweets according to different parameters (time, space, user, topic, etc.). Some of them use very basic sentiment analysis techniques (manually built lexicons) that produce a simple word count. In this group, we mainly find studies trying to predict election results (Tumasjan et al. 2010; Livne et al, 2011; Skoric et al. 2012). Using different kinds of quantitative techniques, their aim is to verify whether there is a correlation between the number of tweets mentioning a candidate and the number of votes he or she receives. In this type of study, political opinion is clearly framed as mass opinion. Opinion is treated as quantifiable, measurable and countable. Tweets are used to study it as an aggregate phenomenon. Similarities to traditional polls can be easily identified. The researcher is forcing a question onto a pre-existing sample of data, assuming that this sample contains the answer.

3.3.2. Opinion as a Sentiment

Taking into account the limits of an approach simply based on the count of preferences, more recent studies have focused on individual attitudes. Researchers who adopted this definition may be also building corpora of tweets containing specific words or hashtags, yet they are interested in the tweet as content rather than as a mere countable unit. They may study co-occurrences of words or more advanced textual structures while trying to interpret the sentiment expressed in the text. The final goal is to obtain a complex view that takes into account the individual positions related to the object of study in accordance with the "latent opinion" form.

A tricky issue is how to define sentiment. In the last few years, the so-called sentiment analysis has become very popular. According to the definition of Wilson et al. (2005), sentiment is a question of contextual polarity: "Sentiment analysis is the task of identifying positive and negative opinions, emotions, and evaluations". The authors offer a sentiment lexicon enriched through supervised learning. Within our corpus, O'Connor et al. (2010), cited 635 times according to Scopus and 1487 times according to Google Scholar, base their analysis of sentiments on the lexicon in OpinionFinder. Their approach has been reproduced dozen of times in the following years. For example, Conover et al. (2011, the third more cited paper in Scopus) develop a content-based method on manual annotation (labelled data) for analysis of 355 millions tweets. More recently, scholars (Bermingham and Smeaton, 2011) have proposed supervised approaches for building sentiment classifiers. This particular approach also includes studies that qualify sentiments in a more qualitative way based on small corpora of data.

3.3.3. Opinion as an Interaction

Some researchers have broaden the scope of their study, shifting their attention from the tweet to its context in order to identify the network of interactions related to the formation and circulation of opinions. According to this conceptual model, opinions are individual sentiments generated not only by the predispositions of a person, but also and foremost influenced by his or her role in the society. Most of the studies in this field focus on the role of influentials coherently with the form of the "activated opinion". Here we can mention quantitative studies trying to identify opinion leaders based on network metrics. Thanks to the analysis of a corpus related to the political hashtags #FreeIran, #FreeVenezuela and #Jan25, Bastos et al. (2013) studied the structure of gatekeeping in Twitter by analysing retweet, mention and followers-following networks for each hashtag. They rejected the idea of the existence of hubs acting as gatekeepers by underlining the importance of committed minorities. Slightly different is the view developed by Park (2013) in his paper, in which he carried out a survey highlighting the difference between traditional opinion leadership based on the two-step flow theory (Katz & Lazarsfeld, 1955) and opinion leadership on Twitter.

Some researchers combine content analysis with network analysis. Xu et al. (2012), focusing on activism networks, explored both opinion leadership through network statistics measures and political involvement through the analysis of the information profile and the content of tweets. Their results were the opposite of those obtained by Bastos et al. (2013), showing the connection between centrality and leadership. In order to predict political affiliation of Twitter users, Conover et al. (2011) combined content-based methods with structure analysis of political information in diffusion networks (retweet and mention networks), and actually validated network analysis as a more efficient solution for identifying political alignments of users. Similarly, Stieglitz and Dang-Xuan (2012) combined sentiment analysis (using Linguistic Inquiry and Word Count LIWC software[11]) with network analysis on a corpus of 64,000 tweets. Their purpose is to study whether articulated sentiment in political tweets has an effect on their retweetability. More qualitative studies analysed tweets of specific classes of users considered as influential, such as journalists (Molyneux, 2015).

3.3.4. Opinion as an Agenda

Other studies focus on the role of Twitter as a medium responsible for setting the opinion's agenda. In this case, the relation between the tweet and its author becomes irrelevant. Tweets are not equivalent to aggregate or individual opinions of people, yet they convey, to the people, social representations directly generated by the platform.

As an example, the study by Papacharissi and de Fatima Oliveira (2012, the second more cited in Scopus) traced the rhythm of news storytelling on Twitter via the #Egypt hashtag. The authors

11 With built-in dictionaries.

intended to identify the evolution of news values that determine the selection of news on Twitter. Methods here are content computer-mediated text analysis combined with discourse analysis. The attention is not really focused on opinion-making but on Twitter as a medium for sharing news.

This last approach is profoundly different from the others because research in this field rarely uses tweets as data but rather investigate Twitter as a social actor. Moreover, it raises the subjacent but essential issue of the validity of tweets as bottom-up data representative of people's opinions. Indeed, according to this model, tweets are the product of the platform rather than the product of the people (Marres, 2017).

4. The Methodological Challenges of Political Opinion

With categories now defined, it is interesting to match up conceptual views of political opinion with the question of methods. In particular, this paper intends to observe cases where studies have been able to scale up from small to big data or, even better, to combine different scales of analysis in order to fill the gap between qualitative and quantitative research.

In order to embrace this question, it is however important to dedicate a few lines to the scientific context. If they sometimes disagree on the best position to adopt, many researchers do agree on the fact that digital data, and in particular the advent of big data, is currently disrupting many fields in social sciences and should hence be cautiously considered as a genuine paradigmatic shift. Regarding the development of digital platforms such as Twitter, T. Venturini and B. Latour (2010) claimed that "[they] offer much more than just another field to apply existing methods: they offer the possibility of restructuring the study of social existence". The impact of such scientific revolution is likely to not only affect social sciences *from within* by invalidating traditional methods and promulgating new ones, but also *from the outside* by modifying institutions, practices, and even some of the epistemic objectives of the field. Among the external effects of big data on the practice of social sciences, the most frequent are: ethical challenges regarding access and privacy, political challenges regarding the dependence of scientific research on the production and the access conditions of this data (Driscoll and Walker, 2014), and even institutional challenges that require a complete reworking of research and education communities in this particular area (Lazer et al., 2009).

However, besides raising the numerous issues about the practices of social sciences, we would rather focus on the impact that such paradigmatic shift has on methods, that is on the actual change *from within*. In order to do this, our analysis will be articulated around three main methodological oppositions related to the use of computational research in social sciences: only-micro and only-macro vs. multiscale approaches; exploratory vs. predictive approaches; supervised vs. unsupervised approaches.

4.1. About Multiscale vs. Macro or Micro Approaches

According to Venturini and Latour (2010), one of the major epistemological challenges that social sciences must face because of big data concerns the reconsideration of a deeply-grounded dualistic vision of data. The classical use of statistical tools – such as aggregative methods – to unravel macroscopic social structures introduces "a fictive distinction between micro-interactions and macro-structures". As a loose analogy, statistical physics early developed inter-level models to distinguish classical macro-measurements from micro-measurements and thereby filled the causal gap between individuals and aggregates. Social sciences are currently working on analogous bottom-up models of social phenomena. Even though we can ascribe it to the tremendous complexity of social objects, this dualistic stance nevertheless leads traditional approaches to the mere "juxtaposition of statistical analysis with ethnographic observation". Following this disturbing statement, Venturini and Latour hence advocate in favour of a genuine "quali-quantitative" approach, building on digital traces to combine the precision of ethnographic surveys and the large-scale scope of statistical analysis in an integrative and unified framework. This epistemological position was already advocated by complexity sciences in the 1970s, as in *The Macroscope* imagined by J. de Rosnay (1975), to observe societies in all their complexity through computer simulation, thus anticipating the growing field of Computational Social Sciences through some tentative essays of using agent models in order to grasp the causal structure of emergent phenomena (Lazer et al., 2009; Cioffi-Revilla, 2016; Casini & Manzo, 2016).

In our corpus, many studies are only interested in the question of aggregates. This is the case in most of the articles dealing with the prediction of election results as their research objective is *per se* driven by a collective framing of political opinion. This objective is enough justification for using methods that generate erroneous measurements when it comes to the analysis of individual opinions. For example, O'Connor et al. (2010) first stated that they were "only interested in aggregate sentiment". Hence, "a high error rate [at the individual level] merely implies [that] the sentiment detector is a noisy measurement instrument. With a fairly large number of measurements, these errors will cancel out relative to the quantity we are interested in estimating, aggregate public opinion." With this line of thinking, the only criterion for method validation is the ability to predict aggregate values, without any guarantee for their potential coupling with individual-oriented researches.

Yet, according to other scholars, the impact of some highly influential individuals is crucial to the explanation of global dynamics. Stieglitz and Dang-Xuan (2012) stressed that "political discussion on Twitter is led by a few highly active users only representing about one percent of all users". They hence performed a detailed analysis of this "one percent" and proposed consistent measurements to simultaneously deal with individuals and aggregates. Similarly, Bermingham and Smeaton (2011), after having performed a large-scale sentiment analysis of 32,578 tweets, tried to "identify terms that provide a path to qualitatively exploring the dataset". They went from a macroscopic observation of opinion to a microscopic explanation by investigating the individual

factors that most significantly impact the aggregate results. The work of J. Boyadjian (2014) can also be considered as a multiscale approach – as regards the building of a corpus – mixing precise data collection through traditional survey with large-scale aggregative collection through computational methods.

4.2. About Explanatory vs. Predictive Approaches

The fast development of big data and its combination with classical algorithmic tools from machine learning have led to another epistemological crisis. As expressed by "neo-positivist" researchers, the paradigmatic shift could equally lead to a practice of social research where social scientists are no longer required. This is the provocative yet sincere claim made ten years ago by Anderson (2008) when he wrote that "correlation [now] supersedes causation, and [that] science can advance even without coherent models, unified theories, or really any mechanistic explanation at all." While the epistemological opposition between correlation and causation is neither new nor solved, the tremendous amount of information made available by digital data on the Web seems to have strongly favoured the application of purely-statistical methods during the last decade. As can be observed in many papers from our corpus, one major objective of today's research is rather the prediction of social phenomena than their explanation. Yet, Metaxas et al. (2011) state that "in the past, some research efforts have treated social media as a black box: it may give you the right answer, though you may not know why. We believe that there is an opportunity for intellectual contribution if research methods are accompanied with at least a basic reasonable model on why they would predict correctly."

Against the idea that sufficiently big data, along with machine learning techniques, can lead to the automatic discovery of significant correlations, some researchers argue, using information-theoretical arguments, that approaches exclusively based on decontextualised statistical analysis are doomed to fail (Calude and Longo, 2016), while others point out that, even if computer-discovered correlation is achievable, it does not fulfil the basic objectives of social sciences. As stated by Pigliucci (2009), science "is not about finding patterns—although that is certainly part of the process—it is about finding explanations for those patterns." The difference between a statistical law, which can be efficient for prediction, and a model, which is able to provide an additional mechanistic explanation, is nicely illustrated by Masad (2014) in a rather simple experiment on Schelling's model of spatial segregation.

This lively distinction between black-boxed learning algorithms and explanatory approaches is another crucial issue for the field we are considering. As a first example, Stieglitz and Dang-Xuan (2012) tried to correlate sentiment expressions in political tweets with their retweetability. They built to do so a "predictive retweet model" using sentiment categories as informative features. While mentioning theoretical background in communication science about "the role of sentiment in the communication in newsgroups, discussion forums or in other contexts", the authors never make explicit the possibility of psychological behaviours underlying the correlation revealed by

their findings. So, it is possible that sentiment expressions are in fact not the cause of retweetability, but only correlated to it through some hidden variable (such as the presence of a URL in the tweet). Similarly, Ceron et al. (2014) made "quite speculative" (*sic*) interpretations about the correlations they find between surveys and Twitter data regarding political leaders' popularity, by discussing the dynamics of the elections in the study. However, a deeper qualitative analysis of these elections would be necessary to go beyond mere statistical correlations, to better understand the causal relations that might exist between online and offline popularity, that is to say how one could explain the other.

Conversely, according to Bermingham and Smeaton (2011), "in opinion measurement and social media analytics it is limiting to simply measure without providing means to explain measurements". In this sense, in order to transfer statistically-discovered correlations into mechanistic models that genuinely integrate a qualitative sociological understanding, Colleoni et al. (2014) proposed an informed interpretation of the discovered correlations between structural features of the retweet network and political leaning found in the Twitter accounts. Exploiting the classical model of political homophily in online and offline discussion, they built their argumentation on typologies developed by political sociology (e.g., "political thinkers", "political activists", "general public") and by communication science (e.g., the distinct modalities of political communication in digital networks, seeing Twitter as a "social medium" or as a "news medium"). This pairing allows qualitative models to provide existent correlations with an adequate sociological explanation. In Jürgens et al. (2011), such pairing is even performed at the formal level. Key concepts inherited from mass communication research (such as "news gatekeepers") are expressed in the framework of graph theory, building on the "Key Player Problem" of Borgatti (2006). Here, the mathematical formalisation of qualitative sociological models are then used for hypothesis testing in order to detect "users who had the strongest possibility to block or disrupt the flow of information".

4.3. About Supervised vs. Unsupervised Approaches

The opposition between predictive and explanatory approaches lies, as we claimed, in the way sociological knowledge is integrated into the quantitative analysis process: either as a way to *a posteriori* interpret correlation results based on qualitative typologies, or as an *a priori* model that needs to be first formalised for hypothesis testing. This distinction hence appears to be a practical instance of the classical epistemological opposition between inductive approaches, where qualitative interpretation comes after statistical analysis, and hypothetico-deductive approaches. In particular, in machine learning methods, the distinction between supervised and unsupervised learning constitutes such a practical case of the opposition. Supervised learning indeed consists in the deduction of categories from *a priori* annotated data, while unsupervised learning proceeds from unlabelled data to infer a classification which receives an *a posteriori* interpretation. In the literature about the use of sentiment analysis in Twitter data for predicting the political orientation

of individuals, three coarse categories of approaches can be distinguished: opinion lexicons, unsupervised learning, and supervised learning.

First, traditional methods for sentiment analysis are based on opinion lexicons such as those provided by OpinionFinder (O'Connor et al., 2010), by an opinion corpus (He et al., 2012), or by a linguistic software using a "psychometrically validated internal dictionary" (Stieglitz and Dang-Xuan, 2012). Such methods are often called "unsupervised methods" (Bermingham and Smeaton, 2011) in the sense that, once the lexicon is set, the classification of tweets is automatic: no further human interaction is required. We however find this name misleading since, on the contrary, much supervision is actually required beforehand: expert knowledge is requested at the very beginning of the classification task. For this reason, O'Connor et al. (2010) claim that opinion lexicons are more "transparent" than other approaches in the sense that the result of classification can easily be explained by the linguistic models that were used to build the lexicon. However, we also note that these so-called "unsupervised methods" are not "learning methods" since the sentiment categories are not learned from the data, but provided *a priori*. As a result, the process is quite rigid, not adaptive to a particular context or to a particular corpus.

To the extreme opposite, unsupervised learning aims at organising the tweets according to their content without requesting any linguistic knowledge. Such approaches therefore build consistent groups of tweets by finding statistically-significant similarities between them, using very generic features such as the words they contain. The resulting categories can then be *a posteriori* labelled by looking at the tweets within, and then used to classify new tweets. Since no prior knowledge is required, such approaches are very cheap in terms of implementation and might result in quite efficient predictions within the discovered categories. However, it might be very difficult to actually make sense of these categories with respect to a particular psychological or sociological framework since they have been built in a purely inductive fashion. Few studies in our corpus rely on this second kind of approach, but we can at least cite the use of topic models (Mei et al., 2007; He et al., 2012) and more generally of latent space models (Barberá et al., 2015).

A third category of highly used approaches lies in between the two categories mentioned above, and that is supervised learning. In this case, a set of sentiment categories is first defined, either a quite simple unidimensional typology ("positive", "negative", "neutral") or a more complex one related to classical human emotions (e.g., "happiness", "excitation", "sadness", "fear"). Then, a set of features extracted from the tweets' content is defined as potentially informative. It can simply consist in counting sequences of words (Bermingham and Smeaton, 2011), in term frequencies (Colleoni et al., 2014), or in more complex grammatical decompositions into part-of-speech elements (Pak and Paroubek, 2010). The objective of machine learning is then to build a "classifier", that is a function using the content-based features of tweets to identify the sentiment category they belong to. To do so, a list of examples obtained by hand annotating a significant corpus of tweets (the training set) is used to initialise the classifier. This is where, for the most part, human knowledge is convoked. Then, a statistical technique is applied to find correlations between the features and the provided annotations. Many such techniques are exploited in the

literature, such as Support Vector Machines (SVM) as in Conover et al. (2011), iterative learning (Bermingham and Smeaton, 2011), passive-aggressive classification (Colleoni et al., 2014), or naïve Bayes classifier (Pak and Paroubek, 2010). Lastly, correlations that have been found are tested and validated by comparing them to another corpus in order to measure its sensitivity (number of true positives) and its precision (number of false positives).

The value of opinion lexicons and supervised learning, compared to unsupervised learning, is that sentiment categories are defined beforehand by the experts, building for example on successful psychological models of human communication. However, in the case of opinion lexicons, the links between features and categories are expressed *a priori* by linguistic expertise whereas, in the case of supervised learning, these links are inductively discovered by learned correlations. According to Ceron et al. (2014), "human coders are, of course, more effective and careful than ontological dictionaries," suggesting the best solution is a more flexible supervised classifier built on real practical examples. Yet, it is important to note that when the selected features are very generic (*e.g.*, any word can be a feature), it might be difficult to disentangle indirect correlations (a word is correlated with a sentiment through a third hidden variable that explains both) from causal relations (a word is actually a direct expression of the sentiment). The resulting classifier, though quite predictive, does not allow an explanation for the prediction result in a clear linguistic manner. Thus, researchers have to choose between explanatory-but-rigid approaches, hence not suitable for the diverse and highly ambiguous use of language in Twitter, and flexible-but-only-predictive approaches that might not fit in with a causal explicative model.

5. Future Trends with the Advent of Big Data

After examining the current state of the art and identifying the main theoretical and methodological approaches developed by digital research about political opinion, it is time to move forward and reach our second objective, that is to propose some guidelines for future trends in the field. In practice, we are interested in computational methods that would allow for a multiscale description of the data, that is methods able to simultaneously describe long-terms dynamics, macroscopic community structures as well as crucial details and micro-events that slip through the control of aggregative methods. If we try to cross conceptual approaches of political opinion with the above mentioned methodological oppositions, a number of interesting insights emerge. In each cell of Table 2, we describe the present situation and future perspectives in research that embraces the different opinion definitions in relation to the three methodological challenges. This table may be a simple abstract representation, since most papers mix more than one conceptual model, yet this abstraction aims at showing the connection between conceptual choices and methodological solutions.

	Multiscale	Explanatory	Supervised
Opinion as a preference	No interest for multiscale: Research focuses only on aggregate analysis	No interest for explanatory methods: Research focuses only on prediction (no need for a model)	No interest for supervised learning: Basic lexicons are satisfying
Opinion as a sentiment	Toward multiscale: Looking at aggregates, but also at crucial users that have the strongest activity (or the strongest influence) on the platform	Toward explanatory: Studying the causal interactions between online and offline opinions (and not only their correlations)	Toward supervised: Developing more flexible sentiment categories that can be partially induced by data, yet supported by psychological models
Opinion as an interaction	Toward multiscale: Using network as a formal tool to measure multiscale structural properties of interactions and their formal inter-level relations (how micro influences macro, and vice-versa)	Toward explanatory: Developing mechanistic models of opinion diffusion to explain how the network structure is responsible for the observed (individual and collective) opinions	Toward supervised: Using structural measures of networks motivated by communication models, yet sufficiently general to discover new patterns of interaction
Opinion as an agenda	Toward multiscale: Studying the contribution of social media's specific users in setting mass media agenda	Limited interest in explanatory methods: Research focuses mainly on the analysis of the existent agenda and on the prediction of future agendas	No interest for supervised learning: Traditional techniques of text mining suffice

Table 2. Intersection between conceptual approaches of political opinion and methodological oppositions related to big data challenges.

Concerning the first conceptual model, where opinion is considered as a collective preference, macro-structures are alone sufficient for the analysis. Studies in this field propose statistical techniques suitable for aggregates and almost exclusively result in predictive analyses. However, the merit of such research group is to raise the question of the representativeness of activity traces obtained through digital media, and notably that of the comparability between Twitter data and traditional opinion polls. Although it has been shown that there is no bijection between Twitter

accounts and individuals (Boyd and Crawford, 2012) and that strong socio-economical biases in the current use of digital media might invalidate their representativeness, many scholars still have "some doubts about whether such bias could affect the *predictive skills* of social media analysis compared to traditional offline surveys" (Ceron et al., 2014).

Papers exclusively based on the preference conceptual model are however very rare. This model is often mixed with the sentiment model. Yet, in both cases, when opinion is considered either as a sentiment or as a preference, methods for the prediction of election results or political polls could improve their efficiency by also focusing on crucial users with the strongest activity (or the strongest influence) on the platform, as suggested by Stieglitz and Dang-Xuan (2012). Such individuals could indeed constitute quite powerful predictors of global trends at a micro-level. The combination of such qualitative analysis with quantitative prediction would however require the use of network-based approaches in order to automatically identify such crucial individuals. Accordingly, going from the mere prediction of election results to their explanation requires a better understanding of causal interactions between online and offline opinion making (and not only of their statistical correlations). As expressed in Ceron et al. (2014), one has to address "the question of the direction of causality", that is: "is the social media opinion becoming more similar to the general public opinion, or, on the contrary, are social media driving (or anticipating) the general public opinion?" In order to do so, the development of relevant sentiment categories is crucial, yet difficult in the context of ambiguous communication on Twitter. Developing more flexible sentiment categories that can be partially induced by data, yet supported by linguistic and psychological models, might help bridge the gap between the too-rigid opinion lexicons and the weakly-supervised approaches of machine learning.

Mixing analysis scales is especially relevant for scholars that investigate the network structure of political opinion. Indeed, interaction-based approaches often imply the combination of micro- and macro-measurements. In general, these approaches propose many formal tools to help measure structural properties of interactions on Twitter at different levels, from micro-structure, such as hubs and bridges, to macro-structures such as communities and other connectivity patterns. What is currently missing is a clear theoretical and empirical understanding of interconnections between such micro- and macro-measurements. For example, how is the presence of bridges and hubs in the network, corresponding to potential influencers from the perspective of communication sciences, correlated with the global connectivity or polarisation of opinions in the network? Consequently, mechanistic agent-based models of opinion diffusion, in particular the ones developed by Computational Social Sciences (see Section 4.1), constitute a promising line of research to explain how such network structures (both micro and macro) are responsible for the observed opinions (both individual and collective). To do so, the chosen structural measures need to be genuinely motivated by communication models, as described in Jürgens et al. (2011). They also need to remain sufficiently generic to discover new modes of interaction, thus achieving a trade-off between purely hypothetico-deductive approaches, formalising and testing communication models with graph-theoretic tools, and inductive

approaches, able to adapt such structural analysis to the very diverse uses of digital media.

Studies that focus on tweet-generated agenda are rarely interested in mixing analysis scales. They either focus on the macro-level scale by identifying general factors capable to influence an agenda (such as studies on news value) or, conversely, on specific individuals, such as journalists or politicians, in order to study how the representations they produce can influence citizen or media agendas. These approaches do not currently study the two levels simultaneously, for example by providing a clear model of how the micro-agendas followed by journalists and politicians impact the collective macro-agendas and, conversely, how the macro-agendas might also produce top-down feedback on these micro-agendas. Yet, it is worthwhile to note that multiscaling is technically possible. Indeed, studies in this field might benefit from a better understanding of informational interactions between mass media (expressing a top-down political agenda) and digital media (building a bottom-up political agenda), thus showing how individual and collective agendas are tangled through the interaction of structurally-different areas of discussion. In order to do so, it is necessary to develop methods able to consistently address political opinion in their different production and consumption contexts, and to interpret them as different levels of opinion making.

Conclusion

This paper took up the considerably hard task of verifying the possibility of performing multiscale analysis of political opinions based on digital data, more precisely on Twitter data. As a first step, we adopted a conceptual viewpoint that allowed us to identify four major approaches in considering political opinions, that is as a preference, a sentiment, an interaction and as an agenda. We then explored methods that have been proposed in the past to fill the gap between the qualitative and quantitative approaches.

First, whether relative to data collection (from individual self-reported surveys to large-scale collection of digital traces) or to data treatment (from the contextualised analysis of particular individuals or events, to the interpretation of macroscopic trends or long-term dynamics), we argued that the concept of "multiscale analysis" can be of great benefit in thinking this methodological challenge by encouraging the simultaneous consideration of crucial details and significant aggregates. We saw that adopting the "sentiment" or the "interaction" conceptual approaches calls for multiscale techniques.

Second, we tried to alleviate one of the main criticisms of the current computational methods when it comes to big data, that is the lack of explanatory power of unsupervised machine learning and other purely-statistical methods. It is indeed essential that the result of data treatment is not limited to the automatic discovery of correlations between explored variables, but that it also provides a causal model allowing to accurately interpret such correlations within the realms of social sciences. We saw that this was all the more made possible when interdisciplinary teams,

constituted by social scientists and computer scientists[12], were able to relate their empirical results with pre-existing models in communication studies and/or political sciences.

Third, these methodological issues are related to the role given to expert knowledge within computational approaches. In this regard, we focused our analysis on the specialised field of sentiment analysis where several proposals have been made to integrate (or to discount) linguistic knowledge within the algorithmic analysis of tweets, leading to different advantages and drawbacks when it comes to the qualitative interpretation of the results.

Needless to say, the good practices here identified are time- and resource-consuming. This implies a long-term project and long delays in the production of scientific publications. Moreover, as previously mentioned, these studies are bound to face important ethical issues. All this explains why, even when hundreds of papers investigate political opinions through tweets, very few offer multiscale methods that can be valid both for computer science and social science standards. Yet, our hope is that this study will help to make progress in this particular field by highlighting the importance of the connection between conceptual and methodological choices, and its consequences.

References

ALLPORT F. H., 1937, « Toward a science of public opinion », *Public Opinion Quarterly*, 1, 1, p. 7-23.

ANDERSON C., 2008, « The End of Theory: The Data Deluge Makes the Scientific Method Obsolete », *Wired*, 16:7.

BARBERÁ P., JOST J. T., NAGLER J., TUCKER J. A., and BONNEAU R., 2015, « Tweeting From Left to Right: Is Online Political Communication More Than an Echo Chamber? » *Psychological Science*, 26:10, p. 1531-1542.

BASTOS, M. T., RAIMUNDO, R. L. G., and TRAVITZKI, R., 2013, « Gatekeeping Twitter: message diffusion in political hashtags », *Media, Culture & Society*, 35, 2, p. 260-270.

BERMINGHAM A. and SMEATON A. F., 2011, « On Using Twitter to Monitor Political Sentiment and Predict Election Results », *Workshop Sentiment Analysis where AI meets Psychology (SAAIP) at the International Joint Conference for Natural Language Processing (IJCNLP)*.

BLONDIAUX, L., 1998, *La fabrique de l'opinion. Une histoire sociale des sondages*, Paris, Le Seuil.

12 Indeed, big data and data science do not change methods of social science, statistics and computer science. The main change concerns the awareness of the necessity and interest of working in interdisciplinary frameworks without neglecting one of these disciplines.

BORGATTI, S. P., 2006, « Identifying sets of key players in a network », *Computational, Mathematical and Organizational Theory*, 12, 1, p. 21-34.

BOURDIEU P., 1973, « L'opinion publique n'existe pas », *Les Temps Modernes*, janvier, 318.

BOYADJIAN, J., 2014, *Analyser les opinions politiques sur Internet: Enjeux théoriques et défis méthodologiques,* Doctoral dissertation, Montpellier 1, France.

BOYD D. and CRAWFORD K., 2012, « Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon», *Information, Communication & Society*, 15, 5, p. 662-679.

CALUDE C. and LONGO G., 2016, « The Deluge of Spurious Correlations in Big Data », *Foundations of Science*, p. 1-18.

CASINI L., and MANZO G., 2016, « Agent-based models and causality: a methodological appraisal", working paper, Linköping University. Available at http://www.diva-portal.org/smash/get/diva2:1058813/FULLTEXT01.pdf.

CAYROL R., 2011, *Opinion, sondages et démocratie*, Paris, Sciences Po Presses.

CERON A., CURINI L., IACUS S. M., and PORRO G., 2013, « Every tweet counts? How sentiment analysis of social media can improve our knowledge of citizens' political preferences with an application to Italy and France », *New Media & Society*, 16, 2, p. 340-358.

CIOFFI-REVILLA C., 2016, « Bigger Computational Social Science: Data, Theories, Models, and Simulations—Not Just Big Data », *8th International ACM Web Science Conference (WebSci'16)*.

COLLEONI E., ROZZA A., and ARVIDSSON A., 2014, « Echo Chamber or Public Sphere? Predicting Political Orientation and Measuring Political Homophily in Twitter using Big Data », *Journal of Communication*, 64, p. 317-332.

CONOVER M. D., GONÇALVES B., RATKIEWICZ J., FLAMMINI A., and MENCZER F., 2011, « Predicting the Political Alignment of Twitter Users », *Proceedings of the IEEE Third International Conference on Privacy, Security, Risk and Trust (PASSAT 2011) and the IEEE Third International Conference on Social Computing (SocialCom 2011)*, p. 192-199.

DE ROSNAY J., 1975, *Le Macroscope, vers une vision globale*, Paris, Le Seuil.

DRISCOLL K. and WALKER S., 2014, « Working Within a Black Box: Transparency in the Collection and Production of Big Twitter Data », *International Journal of Communication,* 8, p. 1745-1764.

ENTMAN, R. M. and HERBST, S., 2001, « Reframing public opinion as we have known it », in W. BENNETT, L. and ENTMAN R. M., *Mediated politics: Communication in the future of democracy*, Cambridge, Cambridge University Press, p. 203-225.

GAYO-AVELLO, D., 2013, « A meta-analysis of state-of-the-art electoral prediction from Twitter data », *Social Science Computer Review*, 31, 6, p. 649-679.

GINSBERG, B. 1986, *The Captive Public: How Mass Opinion Promotes State Power*, New York, Basic Books.

HABERMAS, J., 1962, *L'espace public*, Paris, Edition Payot.

HAYASHI C., 1998, « What is Data Science? Fundamental Concepts and a Heuristic Example», in C. Hayashi et al., *Data Science, Classification, and Related Methods. Studies in Classification, Data Analysis, and Knowledge Organization*. Springer Japan. p. 40–51.

HE Y., SAIF H., WEI Z., and WONG K.-F., 2012, « Quantising Opinions for Political Tweets Analysis », *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC 2012)*, p. 3901-3906.

HOGAN, B., 2014, "From invisible algorithms to interactive affordances: Data after the ideology of Machine Learning" in E. Bertino, S. A. Matei (eds.), *Roles, Trust, and Reputation in Social Media Knowledge 1 Markets, Computational Social Sciences*, DOI 10.1007/978-3-319-05467-4_7.

HOLMES, D.E., 2017, *Big Data. A very Short Introduction*, Oxford, Oxford University Press.

JUNGHERR A., JÜRGENS P., and SCHOEN H., 2011, « Why the Pirate Party Won the German Election of 2009 or the Trouble With Predictions: A Response to Tumasjan, A., Sprenger, T. O., Sander, P. G., & Welpe, I. M. "Predicting Elections With 140 Characters Reveal About Political Sentiment" », *Social Science Computer Review,* p 1-6.

JÜRGENS P., JUNGHERR A., and SCHOEN H., 2011, « Small Worlds with a Difference: New Gatekeepers and the Filtering of Political Information on Twitter », *Proceedings of the Third International Web Science Conference (WebSci 2011)*.

KATZ, E. and LAZARSFELD, P. E., 1955, *Personal influence: The part played by people in the flow of mass communication*, New York, Free Press.

KEY, V. O., 1966, *The Responsible Electorate: Rationality in Presidential Voting, 1936-1960,* with the Assistance of Milton C. Cummings, Cambridge, Harvard University Press.

LAMARCHE-PERRIN R., DEMAZEAU Y., and VINCENT J.-M., 2014, « Building Optimal Macroscopic Representations of Complex Multi-agent Systems », *Transactions on Computational Collective Intelligence*, vol. XV, LNCS 8670, p. 1-27.

LAZARSFELD, P. F., BERELSON, B., and GAUDET, H., 1948, *The peoples choice: how the voter makes up his mind in a presidential campaign*, New York, Duell, Sloan and Pearce.

LAZER D., PENTLAND A., ADAMIC L., ARAL S., BARABÁSI A.-L., BREWER D., CHRISTAKIS N., CONTRACTOR N., FOWLER J., GUTMANN M., JEBARA T., KING G., MACY M., ROY D., and VAN ALSTYNE M., 2009, « Computational Social Science », *Science*, 323, 5915, p. 721-723.

LIPPMANN, W., 1925, *The Phantom Public*, New York, Harcourt Brace.

LIVNE A., SIMMONS M. P., ADAR E., and ADAMIC L. A., 2011, « The Party is Over Here: Structure and Content in the 2010 Election », *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*, p. 201-208.

MASAD D., 2014, « Computational social science, data, and complexity », *Bad Networking,* available at http://davidmasad.com/blog/css-data-complexity/

MARRES, N., 2017, *Digital Sociology. The Reinvention of Social Research*, Wiley.

MARRES, N., 2012, « The redistribution of methods: on intervention in digital social research », *The sociological review*, *60*, S1, p. 139-165.

MEI Q., LING X., WONDRA M., SU H., and ZHAI C., 2007, « Topic Sentiment Mixture: Modeling Facets and Opinions in Weblogs », *Proceedings of the 16th international conference on World Wide Web (WWW'07)*, p. 171-180.

MERCIER, A., 2012, *Médias et opinion publique*, Les Essentiels d'Hermès, Paris, CNRS éditions.

MERTON, R. K., 1968, « Patterns of Influence: Local and Cosmopolitan Influentials », in R. K. MERTON (eds.), *Social Theory and Social Structure*, New York, Free Press, p. 441-474.

METAXAS, P. T., MUSTAFARAJ, E. and GAYO-AVELLO, D., 2011, « How (not) to predict elections », *Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third International Conference on Social Computing (SocialCom)*.

MOLYNEUX, L., 2015, « What journalists retweet: Opinion, humor, and brand development on Twitter », *Journalism*, 16, 7, p. 920-935.

NOELLE-NEUMANN, E., 1984, T*he Spiral of Silence. Public Opinion – Our Social Skin*, Chicago and London, University of Chicago Press.

O'CONNOR B., BALASUBRAMANYAN R., ROUTLEDGE B. R., and SMITH N. A., 2010, « From Tweets to Polls: Linking Text Sentiment to Public Opinion Time Series », *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media*, p. 122-129.

PAGE, B. I., and SHAPIRO R.Y., 1992, *The Rational Public: fifty years of Trends of Americans' Policy Preferences*, Chicago, University of Chicago Press.

PAK A. and PAROUBEK P., 2010, « Twitter as a Corpus for Sentiment Analysis and Opinion Mining », *Proceedings of the International Conference on Language Resources and Evaluation (LREC 2010)*, p. 17-23.

PAPACHARISSI, Z., and DE FATIMA OLIVEIRA, M., 2012, « Affective news and networked publics: The rhythms of news storytelling on# Egypt», *Journal of Communication*, 62, 2, p. 266-282.

PARK C. S., 2013, « Does Twitter motivate involvement in politics? Tweeting, opinion leadership, and political engagement », *Computers in Human Behavior*, 29, 4, p. 1641-1648.

PIGLIUCCI M., 2009, « The end of theory in science? », *EMBO reports*, 10:6.

PUNCH K., 1998, *Introduction to Social Research: Quantitative and Qualitative Approaches*, London, Sage.

REYNIÉ D. 1989, « Le nombre dans la politique moderne », *Hermès*, 4, Paris, CNRS edition, p.159-164.

SAVAGE M. and BURROWS R., 2007, « The coming crisis of empirical sociology », *Sociology,* 41, 5, p. 885–899.

SEVERO M., FEREDJ A., and ROMELE A., 2016, « Soft Data and Public Policy: Can Social Media Offer Alternatives to Official Statistics in Urban Policymaking? », *Policy & Internet*, 8, 3, p. 354-372.

SKORIC M., POOR N., ACHANANUPARP P., LIM E. P., and JIANG J. 2012, « Tweets and Votes: A Study of the 2011 Singapore General Election », *Proceedings of 45th Hawaii International Conference on Systems Science (HICSS-45 2012), IEEE Computer Society*, Los Alamitos, CA, USA, p. 2583-2591.

STIEGLITZ S. and DANG-XUAN L., 2012, « Political Communication and Influence through Microblogging – An Empirical Analysis of Sentiment in Twitter Messages and Retweet Behavior », *Proceedings of the Forty-fifth Hawaii International Conference on System Sciences (HICSS 2012)*, p. 3500-3509.

TARDE G., 1989 (1901), *L'Opinion et la foule*, Paris, Les Presses Universitaires de France.

TUMASJAN A., SPRENGER T. O., SANDNER P. G., and WELPE I. M., 2010, « Predicting elections with Twitter: What 140 characters reveal about political sentiment », *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media*, p.178-185.

VENTURINI, T., and LATOUR, B., 2010, « The social fabric: Digital traces and quali-quantitative methods », *Proceedings of Futur en Seine 2009*, p. 87-101.

WATTS, D. J., and DODDS, P. S., 2007, « Influentials, networks, and public opinion formation », *Journal of consumer research*, *34,* 4, p. 441-458.

WILSON, T., WIEBE, J. and HOFFMANN, P., 2005, « Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis », *Proceedings of the 2005 Conference on Empirical Methods in Natural Language Processing*, p. 347–354.

XU, W. W., SANG, Y., BLASIOLA, S., and PARK, H. W, 2014, « Predicting opinion leaders in Twitter activism networks: The case of the Wisconsin recall election», *American Behavioral Scientist*, 58, 10, p.1278-1293.

ZALLER, J. R., 1992, *The Nature and Origins of Mass Opinion*, Cambridge, Cambridge University Press.