*PA*rtecipazione e *CO*nflitto
** The Open Journal of Sociopolitical Studies*
http://siba-ese.unisalento.it/index.php/paco
ISSN: 1972-7623 (print version)
ISSN: 2035-6609 (electronic version)
***PACO*, Issue 11(2) 2018: 313-331**
DOI: 10.1285/i20356609v11i2p313

Published in July 15, 2018

EDITORIAL

POLITICS, POLITICAL PARTICIPATION AND BIG DATA
Introductory reflections on the ontological, epistemological, and methodological aspects of a complex relationship

Alice Mattoni
Scuola Normale Superiore

Elena Pavan
University of Trento

ABSTRACT: This editorial defines big data as an inherently political object and then briefly discusses its ontological, epistemological, and methodological implications in the social sciences. Furthermore, it addresses these issues in connections with the realm of politics, political participation and political mobilization. Finally, it addresses three main emergent themes related to big data in the broad realm of politics. First, big data as a methodological conundrum - something that can possibly empower or completely bias research activities and results. Second, big data as an object of study in its own right, a contested research and political terrain characterized by strong power dynamics between private and public actors and entwining with governance processes at all levels - from the national to the transnational one. Third, big data as research catalyser that can leverage our understanding of participation and contentious dynamics.

KEYWORDS: big data, ontology, epistemology, methodology, political participation, social movements

CORRESPONDING AUTHORS: Alice Mattoni, alice.mattoni@sns.it and Elena Pavan, elena.pavan@unitn.it

1. Introduction

This special issue aims to open a much-needed space for developing a critical and informed reflection on the multifaceted nexus between big data and the study of politics, political participation and political mobilization dynamics. It understands big data not solely in terms of large-scale datasets of textual or digital data that require us to tune our research practices. More radically, in this special issue we start from a conceptualization of big data as a complex set of cultural, political and scientific knowledge practices. In line with other authors (Elish and boyd 2018, Sætnan et al. 2018, Kitchin 2014a, boyd and Crawford 2012), we indeed define big data as heterogeneous aggregations of: 1) *technologies*, from large-scale data infrastructures to small-scale software and applications; 2) *imaginaries* that enter the public discourse, eliciting what big data might render possible and with what consequences for societies and 3) *people*, from data scientists to computer and social scientists, from policy makers to business professionals, from platform owners to media professionals and the end-users of such platforms. Because they are always-changing aggregations of technical, cultural and social processes, big data constitute an inherently "political object" with ontological, epistemological, and methodological implications - particularly, but not solely, when it comes to their connections with the realm of political participation and mobilization.

First, at the *ontological level*, big data mobilize different types of definitional issues, both on themselves and on the social realities that they come to constitute. Certainly, many definitions of big data have been pushed forward so far and mainly to distinguish them from "other" types of data that scholars, particularly social scientists, have typically leaned on in their research activities. To be sure, though, these definitions often limit themselves to underline big data material features, emphasizing in particular their large volumes, high velocity, and wide variety (Kitchin 2014b).

However, when recognizing the intertwining of technical, cultural and social processes in the making and the use of big data, it becomes evident that not only it is impossible to have just *one* definition to identify them but also that stressing exclusively their material features does constrain, not to say mystify, their very essence. First, big data emerge from and are analysed through many types of technologies that combine differently amongst themselves and also dynamically change across short spans of time. Moreover, increased availability and accessibility to unprecedented amounts of data stimulate many types of desires and desiderata, meanings and imaginaries within and beyond academia. Consistently, big data come from and, at the same time, refer to a myriad of behaviours and practices enacted by a huge variety of subjects - and they can be employed to reach a tremendous variety of goals.

Already on their own, each of these three features points to big data as something that is far than univocal and less so stable in its nature. Most notably, when the technical, the cultural, the social aspects of big data are considered in their entanglement, it becomes simply impossible to prioritize one facet over the others. Thus, big data are characterized by that same "variable ontology" that Taina Bucher (2018) recognizes as a property of algorithms and Adrian Mackenzie (2006, 97) ascribes to software projects like Java, pointing out that "the trademark of a variable ontology project is difficulty in saying what something is once and for all". The extent to which researchers (within and beyond social sciences) are able to acknowledge and deal in practice with this "variable ontology" of big data is simply key to mastering their potential to advance scientific knowledge.

Second, big data have consequences at the *epistemological level*, which are especially challenging when we recognize their variable ontological nature. In the past few years, choices to lean on big data to generate scientific knowledge have been made following more cautious approaches that reject a positivist understanding of big data as something that might put an end to centuries of theories on how societies work, as advocated in a famous op-ed by the director of *Wired USA*, Chris Anderson:

> This is a world where massive amounts of data and applied mathematics replace every other tool that might be brought to bear. Out with every theory of human behaviour, from linguistics to sociology. Forget taxonomy, ontology, and psychology. Who knows why people do what they do? The point is they do it, and we can track and measure it with unprecedented fidelity. With enough data, the numbers speak for themselves (2008).

Similarly, and yet moving one step beyond Anderson, Mayer-Schonberger and Cukier (2013, 6) underline that big data ultimately "challenge" our ways of understanding and organizing society not only by virtue of their volume, velocity and variety. According to the authors, big data are valuable because, they claim, "bigness" approaches "allness" - that is, the larger the amount of data available, the greater the possibility to look at everyone's behaviours in every context, surmounting the need for sampling and inferring, and guaranteeing the possibility of observing society universally. Thus, Mayer-Schonberger and Cukier argue that the continuous availability of large-scale datasets of various types pushes us towards a new mode of producing knowledge: the possibility to observe society universally through these data reduces the imperative of "exactitude", yields to consider "messiness" an affordable price to pay in exchange of a more complete information, and invites to be less concerned with causation and more with correlation - or, as they write "not knowing why but only what" (2013, 7).

Anderson's quote and Mayer-Schonberger and Cukier's account can be read as a relevant piece of imaginary that unfolded in the past decade about big data - one that has been often embraced by the business sector as much as by governmental institutions which have surrendered, way too quickly, to the "myth" of big data, as dana boyd and Kate Crawford (2012) labelled it. At the same time, they also convey a specific epistemological orientation not only towards the modes in which we can get to know the world around us but, more poignantly, on what *deserves to be known as part of the world around us*.

Against the evident perils spurring from the consolidation of this vision, scholars from different disciplinary fields but, particularly, within social sciences quickly engaged in the effort of developing and operating through a more careful reading of big data. In this context, attempts have been made to foster the encounter between social and data-driven research, to produce knowledge combining induction, deduction and abduction and, in this way, to master big data's potential while accounting for the constructed, and ever evolving, nature of such specific type of data (Kitchin 2014a). At the same time, caution has been called against invariantly associating big data with a "paradigm shift" particularly within social sciences. In this respect, Carl Lagoze underlines that big data are not revolutionary when they push upon us new ways of "doing" things - as it happens, for example, when we elaborate new techniques just to "handle" large amounts of data. Rather, big data revolutionize science when they "challenge existing epistemological norms, ways of knowing and framing the fundamental scientific questions of the field; institutional ecologies (Star and Griesemer 1989), agreements on scope, assumed knowledge, and boundaries of research work; reward structures, paths to tenure and promotion; and communication regimes, mechanisms, and norms for disseminating knowledge" (2014, 3).

As a matter of fact, however, scholars within but also beyond social sciences continue to embrace big data quite vigorously but, oftentimes, recklessly and failing to distinguish between "lots of data and big data" (Lagoze 2014, 3). In doing so, they overload large-scale datasets with unrealizable expectations about their capability of revealing the unknown while, in fact, reproducing established research patterns and giving in to the "normative pressure to pay homage to the objective authority of statistical significance" (Goldberg 2015, 1).

Third, while deepest epistemological issues still remain an open and yet urgent matter to face (Kitchin 2014a), it is undeniable that big data stimulate a swarm of new *methodological approaches* to investigate societies, social behaviours, systems of norms, values, and to find patterns of regularities that have hitherto remained invisible. Unprecedented levels of access to large amounts of digital and digitalized data

from a myriad of sources are clearly and undoubtedly allowing us "to do research" in different ways - skyrocketing techniques that were employed also in the past (network analysis, just to mention one) and continuously handing in new computational methods such as machine learning, neural networks, sentiment analysis, natural language processing, and many more.

Whether this possibility to change our research habitus is not necessarily tantamount to a radical paradigm shift, it is nonetheless an exciting opportunity that can and should be exploited and critically reflected upon. Certainly, as clearly stressed by Wagner-Pacifici, Mohr and Breiger (2015), methodological transformations within social sciences and the humanities are tightly linked with both ontological and epistemological questions. As they claim, big data do indeed challenge extant (competing) definitions of "entities, agents, acts, causes, meanings, temporalities, and contexts. [Thus,] many of these are being actively renegotiated in some way as scholars recalibrate to adjust to a new style of science where different scales of analysis are being used and new kinds of social situations are being measured" (Wagner-Pacifici, Mohr and Breiger 2015, 2). However, they note, it is precisely - and exclusively, we would add - in the context of what they call "a particular kind of duality relationship [between] ontology and methodology" (ibid.) that this renegotiation occurs insofar as big data, very much as any other type of data, hold value only when channelled within knowledge generation endeavours.

In this sense, it is of paramount importance to problematize even our very choice of leaning on big data in order to answer our research questions or to make of big data our very topic of interest, as these choices already signal a methodological preference (boyd and Crawford 2012) and, to be sure, are far from being our only research option. Moreover, it remains of greatest relevance to avoid assuming that, even when willingly chosen and recognized as non-neutral, big data can be analysed "objectively" particularly through computationally assisted methods. Coping with initial data abundance, dealing with big data "messiness", is far from being that affordable price that Mayer-Schonberger and Cukier suggested only five years ago. Fixing problems especially during the initial research phases, when data are stored somewhere awaiting to be processed, may be quicker and easier than ever before. However, the very ways in which we approach and decide to act upon data messiness is already transforming big data into "our" data thus taking away their alleged objectivity (Diesner 2015).

The actual modes in which big data are collected and analysed should also be constantly scrutinized. To begin with, big data pose a methodological challenge not just at the level of data analysis, but also with regard to the tools that scholars use to collect them. Who develop such instruments, their technical specificities, and the limits that

these tools might have are all aspects that often remain in the background while their discussion should be at the forefront - especially when scholars decide to lean on third-party collection and analysis instruments. Moreover, whatever method is chosen to analyse big data, it does entail a specific "way of reading" them - even when we resort to so-called "unsupervised" techniques in topic-modelling, for example. Lack of supervision during the analysis process does indeed exist, but it is not equivalent to the lack of an embedded logic of reading a text or understanding words' prominence (Wagner-Pacifici, Mohr and Breiger 2015).

Finally, choosing more or less sophisticated computationally-assisted analytical methods may have become a standard, but it does not constitute an imperative. Both in research that employs big data and in investigations that take big data as an object of study, different methodological traditions might be endorsed, often resulting in mixed-method approaches, and "small data" (particularly those collected qualitatively) continue to be simply crucial to add *sense* and *meaning* to patterns unveiled at a larger-scale (boyd and Crawford 2012).

Against this background, this special issue aims to set a contribution towards understanding more systematically how big data matter ontologically, epistemologically, and methodologically for the study of politics and, more specifically, of political participation dynamics. We designed and organized it as a twofold knowledge space. On the one side, we collected a set of original articles authored by researchers working at the crossroads between multiple disciplinary backgrounds and in connection with multiple research experiences. These articles either discuss in a critical manner the very political nature of big data or exploit them to inform ongoing studies and debates on political participation at different levels - from the national to the supranational - and in different political situations - from electoral campaigns to collective actions to public discussions on European topics. On the other side, we invited some colleagues to contribute a short commentary to the Symposium "A bird's eye view on big data and politics across the world". In this Symposium, we tried to compose a complex puzzle combining first of all multiple geographical perspectives but also different interpretations of politics and political processes, research concerns, and, of course, modes of understanding and contextualizing big data within the realm of politics.

The twofold structure of this special issue also mirrors our conviction that the potentials and the risks of introducing big data in the study of politics and, more specifically, of participation dynamics can be assessed only through a thorough combination between two elements. First, systematic empirical research practices that are, simultaneously, open towards the future and informed by past and current knowledges. Second, a deep awareness of the fact that the enthusiasms generated by the possibilities af-

forded by large-scale datasets and unprecedented analytical computational capabilities need to be balanced against their non-neutrality and their context-specificity.

Consistently, the original articles included in this special issue aim to provide concrete examples of how big data can be embedded within research practices with different research foci and can be empirically treated either as part of a composite toolbox to understand hybrid political scenarios or, alternatively, as proxies to investigate political and power dynamics taking place in different contexts and across the offline/online boundary. The commentaries gathered in the Symposium, instead, reveal the transversality of big data - a transversality that is not only disciplinary but also, and perhaps more importantly, substantive, insofar as big data are shown to be relevant, either as a research tool or as a topic in their own right, to advance our understanding of political dynamics and processes regardless of the actual geographic location or spatial dimension where these are occurring.

Together, the articles and the commentaries in the Symposium seek to promote a discussion that is, as much as possible, globally distributed, multilevel, and multidimensional. Also, they aim to testify the richness of ongoing big data-related research in the domain of politics and political participation - a richness that, we firmly argue, needs to be acknowledged and fiercely defended, before big data become just another exclusive and exclusionary space of knowledge production.

2. Dealing with the challenges, making themes emerge: the articles and the commentaries in this special issue

The original articles and the commentaries presented in this special issue speak, to various degrees, theoretically and/or empirically, to the ontological, epistemological, and methodological implications of big data with particular reference to the realm of politics, political participation and political mobilization. In this sense, they all seek to face, in their own ways, the challenges that we outlined in the previous section while, at the same time, they help us to understand more systematically what the specific relevance of big data for the study of the political realm is. Their joint reading suggests three main emergent themes in relation to which we can try to disentangle the complex relationship between big data and the study of politics, political participation and political mobilization. First, big data are a *methodological conundrum* - something that can possibly empower or completely bias research activities and results. Second, big data are an *object of study* in their own right, a contested research and political terrain characterized by strong power dynamics between private and public actors and en-

twining with governance processes at all levels - from the national to the transnational one. Third, big data are *research catalysers*, as they can leverage our understanding of participation and contentious dynamics and provide us with new methodological and informational resources to investigate structures, contents and mechanisms of citizens' public engagement.

2.1 Big data as a methodological conundrum

Similarly to what happens in the social sciences at large, the increased and increasing presence of big data in the realm of politics requires to openly discuss the potentialities and the limitations of using large-scale datasets and supervised and unsupervised analytic techniques for leveraging our understanding of political participation dynamics - from conventional participation processes, to political communications, to social movements and digital activism. To be sure, also research on politics, political participation and political mobilization seems to follow the path we outlined in the previous section: while often choosing to lean on big data, not much discussion has developed that reflects critically on the implication on their usage to study citizens' engagement. From this perspective, then, big data might be seen as a methodological conundrum: while they might certainly provide new ways to look at these political phenomena, scholars should also use caution when including them in their methodological toolkit. Even more importantly, their use needs to be discussed vis à vis broader matters of research design, research questions formulation and, when relevant, hypotheses testing strategies.

This is particularly true in the case of social movement studies, where big data, especially those coming from social media platforms, are increasingly employed yet without systematically reflecting on the place that media and online dynamics occupy within contemporary mobilizations (Mattoni 2017, Pavan 2017). Some debate has occurred within venues like *Mobilizing Ideas*, the blog of the journal *Mobilization*, based in the United States but recognized worldwide as one of the top journals of the discipline. The blog hosted a series of posts on big data in 2015 with the idea of pushing a debate on their relevance to understand social movements. The authors of these posts highlighted some relevant methodological points: from the need to ease the access to big data to any scholar, independently from her computational skills (Elliott 2015), to the limited capacity big data have to unveil causal patterns, being representative and free from selection biases (Rojas 2015, Schradie 2015).

Three commentaries in the Symposium provide valuable insights to this still emergent, and yet crucial, discussion within social movement studies. In her commentary,

Jennifer Earl points out that, often, scholars of this discipline do not pay attention to the constructed nature of big data. According to Earl, scholars studying political participation and mobilization forget "that digital and social media data, as with any other kind of data, has production, maintenance, and meaning patterns that we must understand before we can usefully learn from them" (p. 491). In her piece, she hence provides a set of principles that readers and reviewers of articles, chapters and monographs relying on big data should put into practice to improve the scientific standards of research on political participation and mobilization. These principles, Jennifer Earl claims, might become an antidote to the production of research that trusts "the veracity of big data and digital data" (p. 493) to the point of producing misleading understandings of what social movements and other forms of political participation are in the datafication age.

From different angle, Svetlana S. Bodrunova also suggests treating big data with wariness: the context in which big data are shaped, indeed, might be crucial for the understanding of datasets on political participation and mobilization. To sustain her arguments, she starts from her own research experience: when analysing a Russian-language Twitter dataset on a violent conflict in Moscow, many of the tweets seemed initially irrelevant and part of the usual noise that all datasets produced through social media platforms entails. However, a deep knowledge of the context in which such tweets were generated allowed Svetlana S. Bodrunova to reconsider her initial understanding of the datasets: what seemed a background confusion of voices in Twitter, revealed to be relevant data for producing grounded knowledge on the online/offline conflict dynamics. Indeed, the author pleas "for bigger attention to contextual knowledge in current studies of political conflicts, as contextual factors may turn online discussions into distorted mirrors of the respective societies and, hence, are relevant in relation to both the production and the interpretation of results" (p. 498).

Finally, in his commentary, Richard Rogers also concentrates on social media platforms as some of the most relevant loci where scholars interested in social movements and other forms of political participation can start from to produce their datasets. More specifically, he addresses five relevant points to be aware of when engaging in research that relies on social media platforms: the critique of social media for not producing "good data", as they lack stability; the understanding of social media as "human actors", hence putting forward ethical issues when conducting research through them; the relevance of "proprietary effects" of social media data, that creates a cleavage between those who have full access to such data and those who do not, usually scholars and researchers; the "repurposive" use of social media data, that are not originally in-

tended as data to conduct scientific research; and the search for "alternatives" to the use of social media data.

2.2. Big data as an object of study

Data increasingly regulate our lives. We are assessed, ranked, profiled and categorized according to data that we produce, more or less consciously, and that are collected "to know us". These assessments enable or limit our access to services and our ability to cross borders; they define us as potential "risks" to societies; and they affect core aspects of citizenship. New publics are created and receive differential treatments based on data analysis. While data regulate society, the regulation of data is becoming a pressing concern. Policies are required for data collection, sharing and analysis, and struggles over policy development have picked up, not least, since the Snowden revelations. From this viewpoint, big data are becoming a relevant object of study in themselves, one that is worth being analysed to understand the processes that lead to their existence, as well as the actors, infrastructures and premises involved in their construction. It is only through this type of research, indeed, that a grounded deconstruction of big data might take place, hence helping us to develop a full understanding of how they might restructure large societal processes as well as the micropolitics of citizens' daily lives.

In this respect, Biagio Aragona, Cristiano Felaco, and Marina Marino investigate in their original article big data as assemblages of three different domains: "things (infrastructures, devices, techniques, etc.), language (code, algorithms, etc.) and people (scientists, users, etc.)" (p. 455). While these three components are certainly difficult to be disentangled at the empirical level, the authors suggest that looking at the places in which big data are developed, stored and transformed is one way through which data assemblages can be grasped and analysed in all their multifaceted nature. Interviewing data scientists who work in three different research centres involved in data calculation, Biagio Aragona and his co-authors point to the political nature of big data, which are "socio-technical constructs that must be studied when they are in action [...] to follow the political use and the agency role of data" (p. 466).

Complementing the work of Biagio Aragona and his co-authors, in their commentary for the Symposium, Jyoti Panday and Jeremy Malcolm discuss data localization regulations and their consequences with specific attention to the Asia-Pacific region. In a world of international and transnational data flows, the two authors suggest that how data are created, controlled and moved across borders become increasingly im-

portant both for commercial purposes and with regard to government data. The emerging set of data localization policies aims at regulating to what extent and how data produced in one geographical space can be transferred into other regions, countries or continents. Through the adoption of a political economy perspective, Jyoty Panday and Jeremy Malcolm argue that "politics and economics interact to shape localization measures and other discriminatory practices to restrict information. While restricting data flows can be an economic strategy, such measures also have political and social implications because they affect public opinion and power. Human rights, too, can be impacted" (p. 521).

These two contributions remind us that big data are intertwined with a complex nexus of political, cultural, and economic forces not just when they are created and stored, but also when they are circulated and retreat for both commercial and government purposes. In short, they suggest that big data are not given, but constantly constructed through dynamic processes. This is certainly true when it comes to their entanglement of the social and the technical dimensions, with both social interactions and algorithmic computations co-evolving at a fast pace. But the constructed nature of big data also holds true concerning the meanings, discourses and imaginaries that sustain and surround them. In this sense, we must recognize that the symbolic level of big data is shaped in a wide range of spaces: from the individual data scientists who work on big data in a more direct way to the public discourse fostered through legacy media, social media and other media outlets. One of such spaces is, without any doubts, the scholarly debate on big data. While much research has been produced on big data themselves, we do not have many systematic insights on how scholars, overall, speaks about big data within and across different disciplines. Christina Neumayer and Luca Rossi (2016) made a valuable step in this direction: considering the relationship of novel media technologies and the parallel development of literature on political participation and social movements, they produced a sociotechnical timeline able to depict how the debate on digital media evolved at the scholarly level in the past 15 years.

Rose Marie Santini, Larissa Agostini, Carlos Eduardo Barros, Danilo Carvalho, Rafael Centeno de Rezende, Debora G. Salles, Kenzo Seto, Camyla Terra, and Giulia Tucci make a similar attempt in this special issue. By making use of the Systematic Literature Review method, Santini and her co-authors map the scholarly knowledge on how computational propaganda (made through the use of bots) is performed to manipulate public opinion on political issues. In so doing, they draw the interdisciplinary boundaries of a field of research that it still in its infancy: not only scholars are still struggling to develop big data tools to detect bots with a certain degree of certainty; they also tend to focus on Twitter as the main social media platform from which to extract big

data to conduct their researchers. At the more substantive level, then, Santini and her co-authors also point out that the debate still remains open with regard to interaction between non-human and human actors in the shaping of political propaganda and its effects.

2.3 Big data as research catalysers

As we mentioned above, the diffusion of big data in contemporary societies opens up novel areas of research and invites us to renegotiate the boundaries of the research domains that we inhabit. In this sense, big data work as research catalysers, stimulating innovative research that updates, integrates, specifies but, sometimes, also disrupts "core assumptions that undergird scholarly work in the social sciences and humanities" (Wagner-Pacifici, Mohr and Breiger 2015, 2).

The study of political participation makes no exception in this regard. First, big data and, more broadly, the progressive datafication of our societies are changing the landscape and the features of more conventional forms of political participation. Big data are already transforming core democratic institutions, like for instance general elections. This is certainly because they are today a fundamental component of electoral campaigns, governmental and legislative dynamics as well as of the interactions between political leaders with their constituencies through the development of microtargeting techniques applied to the realm of political communication. Whereas in the past the usage of big data to condition, not to say to bend, democratic participation did not receive much attention, in these days it has become a public issue. Privacy violations, information leaking, deceiving use of social media platforms are increasingly triggering public debates that, progressively, transform the frames through which big data are publicly perceived as they pass from being "a technical matter" to constitute a matter of "data justice" that integrates the broader social justice agenda (Dencik et al. 2016).

One of the most important and recent scandals in this respect has been the illegal and massive use of Facebook data by Cambridge Analytica to sustain both the Trump electoral campaign in the United States and the Brexit referendum campaign in the United Kingdom. As Annika Rictherich underlines in her contribution to this special issue, such scandal speaks volumes "on the nexus of big data, democratic elections and citizens' perceived freedom of choice as voters" (p. 529). But, more subtly, it also unveils how the somehow mainstream imaginary that depict big data as unbiased scientific tools to grasp social realities contributed to increase the credibility of Cambridge Analytica in the circles of politicians and their communication specialists. These two

aspects, Annika Richterich suggests, render clear that "data are never neutral … [on the contrary] data are normative and influential in that they are societally embedded and may be e.g. used to give credibility to claims and arguments, to advocate or undermine certain causes" (p. 530).

Conversely, in her commentary, Aimée Vega Montiel sheds light on how big data can potentially become tools for a more just and fair society. Focussing on the global level and on the case of Sustainable Development Goals, she discusses the potentialities and the challenges that characterize the use of big data to substantiate indicators that measure women's empowerment and gender equality. On the one hand, she argues, big data could be used as a "starting point to produce relevant information to monitor sustainable development in an open, participatory and transparent way, as well as to promote public debate and to improve public policies" (p. 545). However, this potential is severely limited by the situated nature of big data, which are produced and circulated within a patriarchal and gender-biased ICTs landscape, but also of institutional processes that remain poorly sensitive towards women's and girls' needs.

Big data, however, are not relevant just for the more conventional forms of political participation. In the past few years, indeed, emergent forms of activism have become increasingly common that reclaim citizens' agency towards the production, management and exploitation of (big) data, such as data activism (Milan 2018), statactivism (Bruno *et al.* 2014), and open data movements (Baack 2015). Citizens' engagement with big data from the grassroots might constitute a counter-balance to the challenges that they pose to democracy.

A relevant example in this direction is the one discussed in the contribution by Kelechi Okechukwu Amakoh, Babatunde Adeshina Faustino, Faith Aanu Oloruntoba, and Abigail Odozi Ogwezzy-Ndisika who present two relevant cases in which the use of big data fruitfully met the grassroots struggles to obtain governmental accountability and transparency in Nigeria. First, the authors present *BudgIT*, a civic organization which developed the platform *Tracka* to detect the misuse of public funding in local territories. Second, they discuss the grassroots campaign *#OpenNASS*, which was supported by a wide coalition of civic organizations to demand the transparency of the budget that the Nigerian National Assembly approves each year. Overall, the Nigerian case shows that the use of big data in the framework of civil society and in connection to social movement processes is certainly able to foster "new epistemic cultures" (Milan and van der Velden 2016) able to put into question the most common and commonsense interpretations of the world we inhabit. It also illustrates, though, how fragile a productive and progressive engagement with big data might be in the civil society sector: for instance, Kelechi Okechukwu Amakoh and his co-authors argue that while

skilled data scientists are key to the success of civil society efforts, there is the concrete risk that the interpretations of big data they provide results from the pressures of corrupted politicians.

Finally, big data allows us to read with a new glance those dynamics of political participation and mobilization we already know a great deal about. Three original articles in the special issue goes in this direction, illustrating how big data scraped from Twitter might shed light on topics like general election, grassroots mobilizations, and the Europeanization of the public debate. Roberta Bracciale, Antonio Martella and Chiara Visentin propose an innovative take on the Twitter political debate that developed during the last general elections in Italy, held in March 2018. More precisely, they consider the Twitter political debate as including different forms of actions in the social media platform and they focus on the connections between being active on Twitter and receiving attention on Twitter. Through their analysis, Roberta Bracciale and her co-authors unveil the patterns that characterized the public discourse surrounding the electoral campaign, pointing out that "several categories of users participated and received attention in different ways, and that the overall debate was not monopolized by traditional actors" (p. 386) like legacy media, political parties and politicians.

Elena Pavan and Arianna Mainardi consider the *Non Una di Meno* movement against gender-based violence in Italy. They do so by tracing and comparing online social and semantic networks that arose on Twitter during a national strike organized by the movement on March 8th, 2017 and during a march organized on November 25th of the same year. Overall, their analysis shows that it is possible to approach online networks as part of broader movement dynamics. Indeed, "not only these networks matter as they enrich the relational milieu of contemporary mobilizations. They nurture collective action dynamics insofar as they enable the continuous circulation of ideas, inputs, and frames which are integrated and provide and overall shared symbolic universe under which collective action can be undertaken" (p. 418). Even more relevantly, Elena Pavan and Arianna Mainardi argue that the intertwining of the structural and ideational level in such online networks has a mutual relationship with the moment in which digital media are embedded within specific protests. Specific episodes of contention within a long-lasting mobilization contribute to shape online networks, while at the same time being shaped by them.

Finally, Javier Ruiz-Soler focuses on Twitter looking at whether this platform can foster a Europeanized discourse on contentious issue that are of European relevance. With this purpose in mind, he scrapes and analyses data from the Twitter hashtags *#schengen*, which relates to the Schengen agreement, and *#ttip*, which refers to the Transatlantic Trade and Investment Partnership between the European Union and the

United States of America. Overall, the analysis shows that "Twitter has [...] the potential to boost European conversations between different national bubbles, acting as a bridge between different national spheres and allowing the participation of individual and organizational actors who did not previously have a place in which they could be sufficiently visible" (p. 440).

3. Acknowledgements

Before leaving the readers to the original articles and the Symposium commentaries in this special issue, we would like to spend a few words to situate this work in our broader research journey. This special issue, indeed, finds its roots in a series of events that we organized in the past few years to push further critical reflections on big data as they intertwine with the realm of politics, being it more conventional forms of participation or grassroots forms of collective action. We are grateful to the Centre On Social Movement Studies (COSMOS) at the Scuola Normale Superiore, which has provided us with an invaluable intellectual environment wherein to push forward our work on social movements and their relationship with legacy media, digital media and big data.

We are also particularly indebted to the colleagues and friends who joined us in three main events. First, a seminar on *"The subversion of big data. Cultures, discourses and practices of big data in social movement contexts"* organized (17-18 November 2016, at the Scuola Normale Superiore in Florence) convened and organized by Alice Mattoni and Veronica Barassi in the framework of their Seminar Series on Social Movements and Media Technologies funded by the Economic and Social Research Council in the United Kingdom. Then, the two sections "*Political Sciences and the Big Data Challenge. From Big Data in Politics to the Politics of Big Data*" that we organized in the General Conference of the European Consortium for Political Research in 2017 and 2018 (6-9 September 2017, at the University of Oslo; and 22-25 August 2018, at the University of Hamburg).

All the encounters, the discussions, the enthusiasms, the doubts, the comments, the ideas that populated these events were simply fundamental to motivate us to realize this special issue. They reinforced our conviction that big data matter to the study of political participation, but they also helped us to clarify, once and for all, that their importance is not absolute. Rather, it is always contextual, conferred, and bound by our research practices.

Finally, Alice Mattoni acknowledges that this special issue has also been produced in the framework of the research project PiCME – Political participation in Complex Media Environments, funded by the Italian Ministry of Education, University and Reseach's Scientific Independence for Young Researcher Programme (2015-2019, Grant Agreement Number RBSI14GUJE)

References

Anderson C. (2008), "The end of theory: The data deluge makes the scientific method obsolete", *Wired*, 23 June 2008, Available at: http://www.wired.com/science/discoveries/magazine/16-07/pb_theory

Baack S. (2015), "Datafication and empowerment: How the open data movement rearticulates notions of democracy, participation, and journalism", *Big Data & Society*, *2*(2), 2053951715594634.

boyd D., K. Crawford (2012), "Critical Questions for Big Data", *Information, Communication & Society*, *15*(5), 662–679.

Bruno I., E. Didier, T. Vitale (2014), "Statactivism: Forms of action between disclosure and affirmation", *Partecipazione e Conflitto*, *7*(2), 198–220.

Bucher T. (2018). *If... Then: Algorithmic Power and Politics*. Oxford: Oxford University Press.

Dencik, L., Hintz, A., & Cable, J. (2016). Towards data justice? The ambiguity of anti-surveillance resistance in political activism. *Big Data & Society*, *3*(2), 2053951716679678.

Diesner, J. (2015), "Small decisions with big impact on data analytics",*Big Data and Society*, 2(2): 1-6.

Elish, M. C., d. boyd (2018), "Situating methods in the magic of Big Data and AI", *Communication Monographs*, *85*(1), 57–80.

Elliott T. A. (2015, March 9), A Cloudy Future: The Possibilities and Perils of "Big Data" for Social Movement Research. Retrieved 7 September 2018, from https://mobilizingideas.wordpress.com/2015/03/09/a-cloudy-future-the-possibilities-and-perils-of-big-data-for-social-movement-research/

Goldberg A. (2015), "In defense of forensic social science", *Big Data and Society*, 2(2): 1-6.

Kitchin R. (2014a), "Big Data, new epistemologies and paradigm shifts", *Big Data & Society*, *1*(1), 2053951714528481.

Kitchin R. (2014b), *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences*, London, Sage.

Lagoze C. (2014), "Big Data, data integrity, and the fracturing of the control zone", *Big Data & Society*, 1(2), 1-11.

Mackenzie A. (2006), *Cutting Code: Software and Sociality*, New York, Peter Lang.

Mattoni A. (2017), "A situated understanding of digital technologies in social movements. Media ecology and media practice approaches", *Social Movement Studies*, *16*(4), 494–505.

Mayer-Schönberger V., K. Cukier (2013), *Big Data: A Revolution That Will Transform How We Live, Work, and Think*, New York: Houghton Mifflin Harcourt.

Milan S. (2018) "Data activism as the new frontier of media activism" In Pickard V. W., G. Yang (eds), *Media Activism in the Digital Age*. London: Routledge.

Milan S., L. van der Velden (2016), The Alternative Epistemologies of Data Activism. *Digital Culture and Society*, 2(2), 57-74

Neumayer C., L. Rossi (2016), "15 Years of Protest and Media Technologies Scholarship: A Sociotechnical Timeline", *Social Media + Society*, *2*(3), 2056305116662180.

Pavan E. (2017), "The integrative Power of Online Collective Action Networks Beyond Protest. Exploring Social Media Use in the Process of Institutionalization", *Social Movement Studies*, 16(4), 433-446.

Rojas F. (2015, April 2), Big Data and Social Movement Research. Retrieved 7 September 2018, from https://mobilizingideas.wordpress.com/2015/04/02/big-data-and-social-movement-research/

Sætnan A. R., Schneider I.,N. Green (2018), *The Politics and Policies of Big Data: Big Data, Big Brother?,* London, Routledge.

Schradie J. (2015, April 2), 5 reasons why online Big Data is Bad Data for researching social movements. Retrieved 7 September 2018, from https://mobilizingideas.wordpress.com/2015/04/02/5-reasons-why-online-big-data-is-bad-data-for-researching-social-movements/

Wagner-Pacifici R., J.W. Mohr, R.L. Breiger (2015), "Ontologies, methodologies, and new uses of Big Data in the social and cultural sciences". *Big Data and Society*, 2(2): 1-11.

Papers published in this Special Issue

Aragona B., C. Felaco, and M. Marino (2018), "The Politics of Big Data Assemblages", *Partecipazione e conflitto*, 11(2): 448-471.

Bracciale R., A. Martella, and C. Visentin (2018), "From Super-Participants to Super-Echoed. Participation in the 2018 Italian Electoral Twittersphere", *Partecipazione e conflitto*, 11(2): 361-393.

Marie Santini R., L. Agostini, C.E. Barros, D. Carvalho, R. Centeno de Rezende, D.G. Salles, K. Seto, C. Terra, and G. Tucci (2018), "Software Power as Soft Power. A Literature Review on Computational Propaganda Effects in Public Opinion and Political Process", *Partecipazione e conflitto*, 11(2): 332-360.

Pavan E. and A. Mainardi (2018), "Striking, Marching, Tweeting. Studying How Online Networks Change Together with Movements", *Partecipazione e conflitto*, 11(2): 394-422.

Ruiz-Soler J. (2018), "The Last Will Be the First. A Study of European Issue Publics on Twitter", *Partecipazione e conflitto*, 11(2): 423-447.

Papers published in the Symposium

Bodrunova S.S. (2018), "When Context Matters. Analyzing Conflicts with the Use of Big Textual Corpora from Russian and International Social Media", *Partecipazione e conflitto*, 11(2): 497-510.

Earl J. (2018), "The Promise and Pitfalls of Big Data and Computational Studies of Politics", *Partecipazione e conflitto*, 11(2): 484-496.

Okechukwu Amakoh K., B. Adeshina Faustino, F. Aanu Oloruntoba, and A. Odozi Ogwezzy-Ndisika (2018), "Big Data and Accountability in Nigeria. Insights from the *BudgIT* Organization and the *#OpenNASS* Campaign", *Partecipazione e conflitto*, 11(2): 472-483.

Panday j. and J. Malcolm (2018), "The Political Economy of Data Localization", *Partecipazione e conflitto*, 11(2): 511-527.

Richterich A. (2018), "How Data-Driven Research Fuelled the Cambridge Analytica Controversy", *Partecipazione e conflitto*, 11(2): 528-543.

Rogers R. (2018), "Social Media Research After the Fake News Debacle", *Partecipazione e conflitto*, 11(2): 557-570.

Vega Montiel A. (2018), "Gender Equality and Big Data in the Context of the Sustainable Development Goals", *Partecipazione e conflitto*, 11(2): 544-556.

AUTHORS' INFORMATION

Alice Mattoni is Assistant Professor at the Scuola Normale Superiore, where she is the Principal Investigator of the research project PiCME – Political participation in Complex Media Environments: A Multi-Level and Multi-Method Approach. Her research lays at the intersection of media and politics. She did research on the perceptions of the journalistic profession in local newsrooms; activist media practices in precarious workers' mobilizations; the use of social media during electoral campaigns; and the role of media technologies in pro-democracy and anti-austerity movements, with a special focus on protests in Italy, Spain and Greece. In the past years, Alice collaborated with several international universities, including the University of Toronto (Canada), the University of Technology, Sydney (Australia), the Robert Schuman Centre for Advanced Studies at the European University Institute (Italy), the University of Pittsburgh (USA), Goldsmiths University of London (UK), and Lakehead University (Canada). Amongst other publications, she wrote the monograph Media Practices and Protest Politics. How Precarious Workers Mobilise (Ashgate, 2012) and she co-edited the following volumes: Spreading Protests. Social Movements in Times of Crisis (ECPR Press, 2014), Mediation and Protest Movements (Intellect, 2013), Advances in the Visual Analysis of Social Movements (Emerald, 2013).

Elena Pavan is Senior Assistant Professor at the Department of Sociology and Social Research of the University of Trento. She holds a degree in Communication Sciences (University of Padova, Italy, 2004) and a PhD in Sociology (University of Trento, 2009). Her most recent research interests pertain to the relationships between collective action/political participation and digital media use. Within this area, she is working interdisciplinary combining technical and social knowledge as well as traditional qualitative and quantitative research methods with digital methods and big data approaches.