

# From data politics to the contentious politics of data

Big Data & Society  
July–December 2019: 1–11  
© The Author(s) 2019  
DOI: 10.1177/2053951719885967  
[journals.sagepub.com/home/bds](https://journals.sagepub.com/home/bds)



Davide Beraldo and Stefania Milan

## Abstract

This article approaches the paradigm shift of datafication from the perspective of civil society. Looking at how individuals and groups engage with datafication, it complements the notion of “data politics” by exploring what we call the “contentious politics of data”. By contentious politics of data we indicate the bottom-up, transformative initiatives interfering with and/or hijacking dominant processes of datafication, contesting existing power relations or re-appropriating data practices and infrastructure for purposes distinct from the intended. Said contentious politics of data is articulated in an array of practices of data activism taking a critical stance towards datafication. In data activism, *data as mediators* take a central role, both as part of an action repertoire or as objects of struggle in their own right. Leveraging social movement studies and science and technology studies, this theoretical essay argues that data activism can be mapped along two analytical dimensions: “data as stakes” (as issues and/or objects of political struggle in their own right) vs. “data as repertoires” (or modular tools for political struggle), and “individual practice vs. collective action”. Mapping action repertoires and tactics along these axes allows us to chart the potential emergence of a *political (contentious) data subject* at the intersection of these two dimensions. This furthers our understanding of people’s engagement with data in relation to other forms of activism and existing work in social movement studies. It also helps us interpreting potential trajectories of contemporary social movements, as they increasingly interface with data, devices and platforms.

## Keywords

Data politics, contentious politics of data, data activism, agency, data practices, social movement studies, science and technology studies

## Introduction

Datafication, in virtue of which both system and life-world are transmuted into data and/or mediated by data, constitutes a fundamental paradigm shift for contemporary society (Kitchin, 2014). This paradigm shift has both ontological and epistemological consequences. By “refram[ing] key questions about the constitution of knowledge” (Boyd and Crawford, 2012: 665), datafication alters what we may call the social epistemology of modernity. Attributing to technology a key role in the constitution of the social, it promotes a novel social ontology that changes the very nature of liberal democracy. This article approaches this sweeping paradigm shift from the perspective of the so-called civil society, that is to say the realm of human activity outside the state and the market. Looking at how variably skilled individuals and groups engage with datafication, it proposes to change the focus from “data

politics” (Ruppert et al., 2017) to a “contentious politics of data”. In particular, it looks at how emerging practices of data activism, taking a critical stance towards datafication and massive data collection, articulate specific contentious politics of data.

Ruppert et al. (2017) understand data politics as “concerned with not only political struggles around data collection and its deployments, but how data is generative of new forms of power relations and politics at different and inter-connected scales” (2).

---

Department of Media Studies, University of Amsterdam, Amsterdam, Netherlands

### Corresponding author:

Stefania Milan, Universiteit van Amsterdam, Turfdraasterpad 9, Amsterdam 1012XT, Netherlands.  
Email: [s.milan@uva.nl](mailto:s.milan@uva.nl)



The authors develop their notion of data politics in response to a number of shortcomings they identify in the current recollections of the relation between people and data, which, in their opinion, over-emphasize the inertness and atomism of social actors, as well as the immediacy of their reactions. Within this frame, we would like to shed light on a fundamental component of today's datafied society, namely grassroots contentious processes expressed by laypersons, nongovernmental organizations and social movement networks alike. By contentious politics of data we mean the multiplicity of bottom-up, transformative initiatives interfering with and/or hijacking dominant, top-down processes of datafication, by contesting existing power relations and narratives and/or by re-appropriating data practices and infrastructure for purposes distinct from the intended. Those initiatives vary for scale, organizational forms, tactics, political values and sociotechnical imaginaries. What they all share, however, is the central role of *data as mediators*: from tactical tools deployed as part of an action repertoire (e.g. citizen-led data collection to map a social problem), to the very same *raison d'être* of a mobilization (e.g. the campaign to stop police "spying on" the residents of a certain area).

As "a field of power and knowledge" data politics "is yet to find its subjects", claimed Ruppert et al. (2017: 2–3). We argue that a complex "field of shared action" (Melucci, 1996: 16) is emerging within the civil society realm. It is composed of loosely connected initiatives which variably intervene in the politics of data from a critical standpoint. Thus the question, is the constitution of a collective political subject under way? Grounded on the disciplines of social movement studies (SMS) and science and technology studies (STS), this article offers an analytical grid to, on the one hand, read the contemporary activist landscape and, on the other, start imagining how such a *political data subject* might look like.<sup>1</sup>

We argue that data activism can be mapped along two analytical dimensions. As we shall explain, the edges of the first dimension consist of data-as-stakes (that is, as issues/objects of political struggle in their own right) and data-as-repertoires (in other words, as modular tools for political struggle). The second dimension sees engagement with data as individual practice vs. an instance of collective action. We argue that laying out this analytical map can help investigating further the emergence of a political data subject in the civil society realm.

The article is organized as follows. In the next section, we present the conceptual tenets of the paper. We then review the burgeoning literature on political reactions and/or uses of Big Data by civil society and offer a first, broad-strokes distinction of the

consequences of datafication on people. We proceed to offer a preliminary working definition of data activism, breaking it down in its bare constituent parts, namely data, agency and technology, and bring this distinction to bear on the definition of data politics by Ruppert et al. (2017). Afterwards, we present a more refined typology of data activism that makes sense of the distinction between data-as-stakes and data-as-repertoires. We conclude by reflecting on the consequences of our analytical model for the present and future of digital activism, and for our understanding of the articulation between worlds, subjects and rights evoked by Ruppert et al.

### Where data re-mediate activism: An interdisciplinary perspective

Inspired by other notable examples in this direction (e.g. Hess, 2005, 2016; McFarlane, 2009), we bring SMS and STS into a fruitful dialogue. This combination empowers us to foreground the relevance of the materiality of datafication and its entanglement with the meaning-making practices of contemporary social struggles. While SMS enable us to scrutinize struggles around meaning-making of social actors (e.g. della Porta and Tarrow, 2005; McDonald, 2004; Melucci, 1996), STS adds a sensibility towards material-semiotic processes (e.g. Bijker and Law, 1992; Bowker and Star, 2000; Latour, 1994). Furthermore, SMS enable us to acknowledge the creative and transformative function of grassroots initiatives and "counterpublics" (Hess, 2016) in resisting or redirecting datafication processes, for data enable new ways of knowing the world and new ontologies of participation (Milan and van der Velden, 2016). Meanwhile, STS empower us to consider the active role that material elements—technological infrastructure, software tools, data ecosystems—play in facilitating and/or constraining such bottom-up initiatives.

Following STS, which disallow the existence of pre-defined objects of analysis, we trace the complex articulation of assemblages *in fieri* by unfolding the underlying socio-technical relations (Law and Hassard, 1999). We thus explore how activism re-shapes data and how data re-shapes activism, moving away from seeing data as mere intermediaries of socio-technical processes to understanding data as proper mediators (Latour, 2005). As a result, we conceptualize data activism as the field *where data re-mediate activism*. One of the main assumptions of this essay is that data is increasingly contentious, both in the sense that it becomes a locus of struggle and in the sense that it turns into tools (or agents) of conflict. In addition, the perspective of social movement

scholarship allows us to see data activism through the collective action lens, foregrounding meaning-making and organizational activities (Gamson, 1992; Polletta, 2002). We thus understand data activism as a critical relation *with and towards* data.

If we understand communication as a process centered around meaning-making and central to organizational definition (Kavada, 2016; Schoeneborn, 2011), we can appreciate the key role of the affordances of the technological dimension of communication in the co-constitution of a field of action and a web of meanings—a critical analytical tool to understand social movement formation and impact (Melucci, 1996). This calls for a careful consideration of the infrastructural layer of activism, that is to say the often invisible, embedded socio-technical arrangements that make social practices possible (Edwards et al., 2009; Star and Bowker, 2002). The infrastructural layer of activism can be seen as a constitutive part of what Ruppert et al. (2017) call the “conditions of possibility of data” (1). Infrastructure is key to the contentious politics of data especially because the backbone of contemporary datafication processes is currently a much contentious domain. In this domain, composite interests and values are negotiated and often clash with each other, getting entangled with a dominant “master narrative” (Star, 1999; see also Milan and ten Oever, 2017). The recent mobilization against the repeal of net neutrality is a living proof of the growing importance of the infrastructural layer today (cf. Stier et al., 2017).

Our definition of contentious data politics builds on the notion of “contentious politics” by social movement scholars Tilly and Tarrow. They described those “interactions in which actors make claims bearing on someone else’s interests, leading to coordinated efforts on behalf of shared interest or programs” (2015: 4). More specifically, contentious politics occurs “when ordinary people—often in alliance with more influential citizens and with changes in public mood—join forces in confrontation with elites, authorities, and opponents” (Tarrow, 2011: 7). Furthermore, it “is used by people who lack regular access to institutions, who act in the name of new or unaccepted claims and who behave in ways that fundamentally challenge others or authorities” (2011: 4). Departing from Tilly and Tarrow, who claimed that contentious politics is by definition collective (2015), we include in the rubric of contentious data politics *also* individual acts of rebellion and practices as they are made possible by the infrastructure of datafication. “Making claims” in the age of Big Data and machine learning is still a dynamic and relational affair (cf. Tilly, 2008). But because collectivity is an objective to work towards, rather than a given, as both Melucci (1996) and Latour (2005) made clear, it is worth investigating a

spectrum of empirical phenomena that might or might not be involved in a process of group-building. In other words, the contentious politics of data embraces also the array of everyday practices of resistance, subversion and creative appropriation embodied by individuals. Think for instance of the adoption of encryption or the act of covering the laptop camera with a sticker as mechanisms of self-protection from the interference of states, corporations or malicious actors.

Our conceptual work at the intersection of SMS and STS seeks to overcome what we may call a “historical” divide in the study of social movements. SMS have been traditionally split along an analytical cleavage: a largely North American tradition emphasizing “how” questions, that is to say by which means, resources and forms social movements mobilize (e.g. McAdam et al., 1996; Tilly and Tarrow, 2015), and a mainly European tradition focusing on the “why”, e.g. for what goals and with what implications people protest (e.g. Melucci, 1989; Touraine, 1981). In other words, the former has been more concerned with investigating a movement’s repertoire, while the second has emphasized the agendas, including also culture and values, movements put forward. STS’ emphasis on flat ontologies (DeLanda, 2006) and symmetrical epistemologies (Latour, 2005) encourage us to explore all the directions of the nexus data-activism. We might explore, for example, data as assembling new forms of contention as well as data being (re-)assembled in an open-ended contentious process. Hence, injecting an STS sensibility in the SMS reasoning allows us to bridge existing traditions within the discipline, recognizing that investigating how activists can creatively employ data and which power struggles surround the control of data are to be treated in a comprehensive analytical framework—one exploring how data remediates activism and how activism remediates data.

## Data politics and the civil society

Today, data is an “object whose production interests those who exercise power” (Ruppert et al., 2017: 3). It is increasingly contentious in two ways: it is strategically mobilized within more generic instances of contentious politics (e.g. counter-mapping and critical cartography), or it is at stake in crucial contemporary struggles (e.g. against algorithmic discrimination). We subsume the various empirical manifestations of the contentious politics of data in the category of data activism. Data activism, in its variable manifestations, has recently come under the spotlight of critical social science scholarship, which recognizes the originality and relevance of practices of engagement with datafication emerging within the citizenry at large.

Scholars have encouraged us to investigate how agency is re-defined in the age of Big Data (Kennedy et al., 2015; Milan, 2018), paying attention to “bottom up” initiatives (Couldry and Powell, 2014; Gabrys et al., 2016) and “everyday” practices of “living with data” (Kennedy, 2018). They have stressed how nongovernmental organizations and human rights defenders, among others, encounter difficulties in making sense of widespread surveillance (Dencik and Cable, 2017; Gangadharan and Jędrzej, 2018), and how mechanized processing of data amplifies existing forms of discrimination (Arora, 2016; Hoffmann, 2019). They have shown how novel alliances between previously discrete realms of activism emerge (Baack, 2018; Gutierrez, 2018), and how people creatively make sense of the complexity of algorithmic processes (Bucher, 2017; Lehtiniemi and Ruckenstein, 2018).

Zooming in on emerging forms of grassroots engagement with datafication, we can distinguish between two “types” of consequences of datafication on people: its top-down effects on groups and individuals (which we will term “institutional politics of data”) and the bottom-up practices embodied and promoted by individuals and groups (the “contentious politics of data”) (cf. Hintz and Milan, 2018). The first refers to the numerous and diverse consequences on people’s lives of datafication as a systemic process, such as government surveillance, corporate profiling, algorithmic discrimination and platform capitalism, to name but a few. These rather heterogeneous processes occur both in authoritarian and in liberal regimes and make visible the reassessment of power relations promoted by datafication. Scholars have referred to “surveillance capitalism” (Zuboff, 2019), “surveillance culture” (Lyon, 2018) and “data colonialism” (Couldry and Mejiias, 2018) to capture this difficult relation, which invariably positions the individual at the losing end of the equation. Conversely, the contentious politics of data indicates the sociotechnical practices of individual users and civil society entities, who are increasingly aware of the issues at stake in datafication as well as of the opportunities for democratic empowerment embedded in data practices. Despite the internal heterogeneity of these “categories” of after-effects of datafication, both strands can be seen as structural transformations of civil society in response to datafication. We focus on the contentious politics of data to offer some analytical tools to decipher the complexity of the contemporary landscape in relation to, on the one hand, an attention towards inequality, social change and its enabling factors embodied by SMS, and on the other hand, a sensibility towards materiality and relationality of technology and information proper of STS.

## From data to data activism

Contrary to the Latin etymology of the term—a past participle meaning “given”—data does not exist as *datum*. Rather, it is the outcome of historical and social processes of (co-)construction, (re-)definition, and appropriation (cf. Gitelman, 2013). This constitutes data as a potential site of struggle in its own right—i.e. issues of socio-cultural and political significance around which contention might emerge. Data nonetheless embodies also a dimension of “givenness”, a functional rather than an ontological one: it is a ready-made input for standardized practices, the result of pre-determined processes of acquisition and classification (Mosco, 2014; cf. Espeland and Stevens, 2008). This aspect turns data into an increasingly important element of the action repertoire through which political contention of various nature is expressed today.

But turning data into data activism can be a complex affair in the hands of people, regardless of their expertise. To be sure, here the so-called “Big Data” are not evoked in relation to the scale and complexity of information in absolute terms. Instead, data is defined in function of *what people do with it*, with accent on human and political agency rather than size. More often than not, grassroots activists and laypersons deal in fact with “good enough data”—that is to say, “citizen data” generated for example through measuring and quantifying practices sufficient to provide “ways of realizing environmental and social justice” (Gabrys et al., 2016: 14).

As argued elsewhere, contemporary examples of data activism include the range of sociotechnical practices positioned along a continuum between two approaches, namely “re-active” and “pro-active” (Milan, 2017). Reactive forms of data activism primarily concern individuals and groups resisting the perceived threats of massive data collection, often by way of technical fixes. We refer to, for instance, the organization of security training for non-experts (Daskal, 2018) and the obfuscation of industrial data collection techniques (Brunton and Nissenbaum, 2011). In contrast, proactive data activism characterizes projects taking advantage of the possibilities for civic engagement, advocacy, and campaigning that the datafication of social life brings about. Examples include the development of “civic technologies” (Wissenbach, 2019) and of forensic devices for activism (van der Velden, 2015) as well as open data and transparency initiatives (Schrock, 2016; Torres, 2019). The distinction between data-as-stakes and data-as-repertoire is explained in more detail below. It represents a step forward in analytical terms in that it combines the “how” and the “why” of social action.



Engagement with data, in this context, necessarily means both technological engagement (e.g. software development, hacks and fixes) and political engagement (e.g. standard setting, redefinition of cultural valence, questioning of power relations). The two are indeed entangled in the practice of the contentious politics of data, where technology has always a political dimension (better illustrated by data-as-stakes) and politics concerns also the technological infrastructure (better illustrated by data-as-repertoires).

It is worth noting that the notion of data activism weaves together reactions and interactions with datafication (and technology in its broadest sense) that have been typically threatened as antithetical, for their accent on either the opportunities of datafication or its threats. By taking data activism as a whole, we acknowledge that data-as-repertoire and data-as-stakes adhere to the same phenomenon, whose components cannot be fully understood if taken in isolation. To put it simply, they represent two sides of the same coin as they are both effects of the expeditious evolution of information and technology. In the same vein, the notion of data activism embraces discrete but complementary means to achieve political goals in the age of datafication.

All things considered, our definition of data activism embodies three innovative aspects. Firstly, it emphasizes the aspects of practice and agency in approaching data and data infrastructure. Secondly, it brings together two approaches to data—as opportunity or threat—that might at a first sight seem poles apart, but which cannot be fully understood in isolation. Thirdly, it foregrounds the productive role of infrastructure as a key component of (digital) activism today. But how does data activism contribute to furthering our understanding of data politics?

### **Contentious data politics vs. data politics: Two complementary perspectives**

Like the concept of data politics by Ruppert et al., our notion of contentious data politics supports a “shift[t] to social practices and agents” (2017: 3). It also complements our understanding of data politics in four main ways. Firstly, it emphasizes the productive role of the material infrastructure of datafication and of its incorporation in material and semiotic struggles. It shows how subjects can make sense of the invisible yet filled with consequences, or the “worlds” evoked by Ruppert et al.—that is to say, the “critical infrastructures that are generative of politics and struggles” (p. 4). Secondly, by postulating the possibility of a political data subject, and even of a collective subject, it moves beyond the atomism that worries Ruppert et al.

Thirdly, by underscoring the novel temporalities of data activism, it partially offsets the critique of immediacy rightly moved by the proponents of data politics. Finally, by contributing to redefine how citizenship can be enacted in the age of datafication, it provides a concrete example of how data politics from the ground-up “reconfigures relationships between states and citizens” (p. 1). In what follows, we explore these four claims.

We are reminded that “[d]ata has a performative power that is resignifying political life” (Ruppert et al., 2017: 2). Yet, we are told, thus far this composite field of action and power has not found its political agents. If data professionals are the custodians and the gatekeepers of the age of Big Data and artificial intelligence, then data activists are the mediators between the data layer and ordinary users. They are the interpreters of the “transversal relations and legal and political tensions that make up data politics” (p. 6). To put it simply, in virtue of their skills and/or awareness, data activists are able to “translate” data and data infrastructure into a site or means of struggle. In this sense, data activism, especially when it becomes a collective matter, constitutes an antidote to the atomism of many contemporary approaches to and narratives of datafication (see, for example, the flamboyant promises of Internet of Things technology). It also contributes to rationalize and re-articulate collectivities, as the grand challenges of data and data infrastructure are capable of rallying disparate subjects with distinct expertise, agendas and motivations. The various “civic tech” groupings emerging locally worldwide, for instance, are illustrative of the novel collectives reordering around data challenges, as they bring together community revitalization activists, data wizards, and, e.g., parents or environmental advocates in view of ameliorating the state output.

But data activism also contributes to challenge the monopoly of the state over information and infrastructure. We agree that “data and politics are inseparable. Data is not only shaping our social relations, preferences, and life chances but our very democracies” (p. 2). To be sure, the battle for the control of flows and infrastructure of information and related policies predates the rise of automated massive data collection (see, e.g. Braman, 2009; Castells, 1996; DeNardis, 2009; Nye, 2011). Today, however, resistance and re-appropriation of data represent even more so central mo(ve)ments of tension between civil society and the state and/or corporate power. What is new is that data can be repurposed as a sort of “new currency” in the relationship between the state and the people. For instance, activists can analyze the open data made available by public administrations to find evidence to support their claims towards the state. They can use apps and platforms to report failures in public

services to their municipality. In so doing, individuals and groups can enact their citizenship—intended as political agency<sup>2</sup>—through data. They can use digital media to facilitate large-scale protest to hijack agenda-setting mechanisms. Drawing on radical democracy and feminist theories, we can see this type of engagement with data as a recurrent exercise of intervention in and transformation of the state, oriented to contesting consolidated ways of “doing things”, institutionalized social relations and legitimized sources of authority (cf. Rodriguez, 2001). Data activism can thus be seen as an effort of (re-)activating and (re-)defining citizenship as a *lived experience in the first person*—one which allows individuals to move beyond the sporadic activation of citizens as voters, and to do so in the framework of a collective experience.

This continuous activation of citizenship via data, apps and platforms points to the ability of data activism to promote novel temporalities, offsetting the trend towards the disempowering immediacy rightly criticized by Ruppert et al. (2017). By combining the present and the past of automatized data collection with sense-making, data activism can uphold original temporalities of material preservation, shared memory and collective mobilization. The Syrian Archive is a working example of how these novel temporalities are forged. Human rights activists and programmers engage in the systematic collection, verification and systematization of images and video footage of human rights violations in the Syrian conflict gathered from social media, and thus condemned to oblivion (Deutch and Habal, 2018). At the same time, however, one must notice how the temporality of contention over and through data might work the other way around: whereas the transformation of digital communication into persistent traces is at the core of the logic of surveillance capitalism, end-to-end encryption or right-to-be-forgotten advocacy are at the center of the resistance towards it.

We now proceed to presenting a more sophisticated typology of data activism, in view of offering an analytical grid with which to attempt to reduce the complexity of the data activism landscape. This analytical grid can help us better understand how a political data subject might be underway from the composite field of contention around and with data.

### **A new typology of data activism: Towards a political data subject**

As an expression of the contentious politics of data, data activism manifests itself both as contention *over* the control of data, and as contention manifested *through* data practices. Moreover, the various instances

of data activism can be positioned along the spectrum of purely individual acts of resistance (however within a collective imaginary) and organized collective action proper. Both dimensions are to be conceived not in a descriptive but rather in an ideal-typical sense; their heuristic value lies in their capacity to analytically bring together a set of phenomena which might look different if taken at face value. This allows us to identify (potential) lines of convergence towards one or more cohesive political data subjects.

#### ***Data-as-stakes vs. data-as-repertoire***

Looking at the relation between activism and data from a bidirectional perspective, we can distinguish analytically between two “flows”, namely how activism shapes data and how data shapes activism. The first refers to initiatives that identify data as objects of intervention in their own right, where data is quite directly involved in the emergence of the rationales and goals of collective action. The second identifies cases where data is largely instrumental and/or a strategic variable, reshaping the means and conditions of activism more or less besides its specific ends. In the first instance, thus (collective/individual) action is *oriented* towards data. In the second (collective/individual), action is *enabled* by data. In other words, data in relation to activism can be seen both as the main *stake* in a hypothetical claim-making agenda and/or as part of the *repertoire* of action, alongside other more traditional forms of protest and civic engagement.

Data-oriented activism—that is, activism that identifies data-related processes as the main stake of action—emerges as a grassroots response towards critical aspect of datafication. It adopts a variety of action repertoires, including advocacy, literacy promotion, software development, campaigning. Examples include the activities promoted by digital rights organizations, security trainers, and algorithmic transparency initiatives. Data-enabled activism—in other words activism that exploits data collection, analysis and visualization as part of a repertoire for promoting its goals—arises instead from the opportunities that datafication opens up for mobilization, campaigning and evidence-production for claim-making. It mobilizes data practices towards a variety of social, political or personal causes. Data journalism oriented to advocacy, (critical) citizen science, and open data initiatives are illustrative of this ideal type.

The distinction is analytically evident when comparing two examples: signing a petition to protest the repeal of net neutrality vs. tracking air pollution data to advocate for better air quality policies. In the first case, the means of the action do not have much to do with the effects of datafication, whereas the issue of

data is strictly involved in the orientation of such action. In the second case, what is at stake is largely unrelated to data as such. What we see here is a “data practice” entering the repertoire of environmental activism. Besides this analytical distinction, data-oriented and data-enabled practices are often conflated to various degrees in the empirical manifestations of data activism. While advocating for the right to resist massive data collection, privacy activists might engage with data obfuscation practices in order to tactically affirm this right. Whereas open data is in some cases primarily a tool to promote government accountability, open data activists generally also make a point about data policy as such.

In SMS, means and ends are increasingly understood as a continuum and as affecting each other (Dekker and Duyvendak, 2013; Jasper, 2011). The idea of a purely instrumental movement selecting the most appropriate tactics in order to reach its exogenous goals is often an untenable simplification. This is increasingly visible in the practice of data activism. Whereas certain initiatives adopt a mix of digital and non-digital tactics, relying on data-enabled action repertoires is gradually becoming unavoidable. Similarly, activist networks adopting data practices to fight for a great variety of goals, tend to increasingly embed a critical attitude towards the premises and consequences of these same practices.

### *Individual vs. collective action*

Data activism is expressed through a variety of organizational and identity forms. Another relevant analytical dimension is indeed the individual vs. collective one. Some of the existing data activism initiatives take the shape of individual acts of resistance, while others consist of large-scale collective mobilizations. It is evident that an individual embracing tool for encrypting communications represents a rather different phenomenon than a global advocacy network established to demand greater algorithmic accountability. Once more, this distinction is largely analytical rather than empirical. A seemingly individual practice like adopting anonymous browsing is not only rooted on the existence of a collective project like the anonymity network Tor (<https://www.torproject.org>) but also routed through several Tor relays volunteered by thousands of users. Similarly, deleting a Facebook account might be an individual act, but it might as well be a manifestation of the #DeleteFacebook campaign. More often than not, these individual acts of resistance and/or rebellion feed themselves on widespread imaginaries and popular slogans.

It is worth noting how the appropriate “scale” at which contentious politics is expressed nowadays

represents one of the most heated debates in contemporary social movement literature. Some scholars argue that, no matter how redefined by technology, digitally mediated movements still manifest a proper collective character (Gerbaudo and Treré, 2015). Other contends that contemporary forms of activism generally take place in more individualized forms (McDonald, 2004), or novel hybrid “connective” formats (Bennett and Segerberg, 2013; Milan, 2015). Data activism, we argue, points to yet another “hybrid” form (cf. Treré, 2019), with specific temporalities and increasingly also a collective identity of its own.

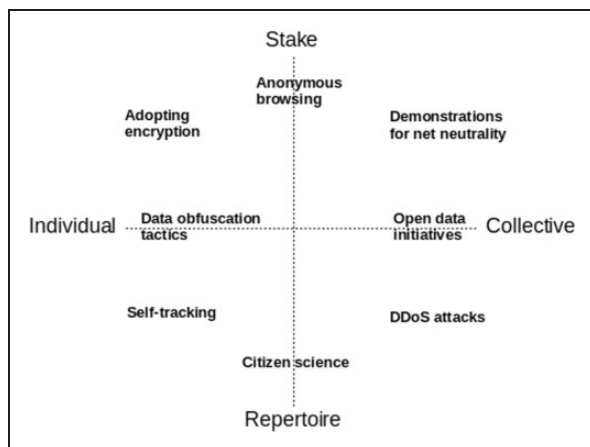
### *Heuristic value of the typology*

We are aware that both the distinctions introduced above are simplifications of a much complex and entangled reality, where the categories “data-as-stakes” and “data-as-repertoires”, as well as “individual practice” and “collective action”, are often not mutually exclusive. The classes identified by these distinctions typically overlap in the empirical reality and are often simultaneously present in the same empirical case. In other words, they are ideal types useful to inform a systematic exploration of such a broad field. Figure 1 plots these dimensions along two axes, in order to propose a typological grid through which to observe a variety of empirical cases. Illustrative examples are tentatively positioned on the grid in relation to their relative distance from the selected ideal types.

The first analytical category is that of individual practices that take data as their stake. An example of this is “hiding from the Matrix”, that is to say the adoption of a number of technological habits to avoid to the best of possibilities being a target of surveillance, such as using email encryption or avoiding data intensive applications. Secondly, data is at stake of contention also from a collective action point of view. This might also not heavily involve data-specific practices. A case in point is the adoption of traditional protest tactics, such as organizing a demonstration or submitting a petition, in order to influence government decisions infringing on “Internet freedoms” (e.g. German Working Group on Data Retention, 2008).

At the intersection of individual action and the adoption of a datafied repertoire stands, for example, the practice of self-tracking. Whereas not a paradigmatic expression of contentious politics, it does sometimes repurpose existing tools for self-improvement and “body hacking” which have not (necessarily) to do with data issues as such. Although sometimes framed within the so-called Quantified Self movement (Lupton, 2016), it generally relies on individual practices.





**Figure 1.** Typology of data activism. The vertical axis distinguishes between the ideal types of “data-as-stake of contention” and “data-as-repertoire of contention”. The horizontal axis distinguishes between the ideal types of “individual practices” and “collective action”. Examples are scattered on the grid based on their proximity to the ideal types. Examples are located in the map merely for illustrative purposes.

Datafied repertoires of action also manifest a collective dimension. This is the case for Distributed Denial of Service (DDoS) attacks or other forms of collaborative “hacktivism” (Jordan, 2002). Whereas the goals of those tactics may vary, they exploit the disruptive power of data to make a political claim. A paradigmatic case of data-oriented practices in which the boundaries between the individual and the collective are explicitly blurred is that of anonymous browsing. As mentioned above, browsing the web with Tor can be hardly understood as a solitary act, as it critically relies on the network effects made possible by a community of volunteers cooperating to mask the routing paths of information packets.

Certain strands of citizen science are illustrative of data practices mobilized as tactics for activist goals, in which individual and collective elements are combined. Think of initiatives monitoring noise or air quality in order to pressure the local governments for better environmental policies (Berti Suman, 2018). These generally rely on individual data collection practices within the frame of a collaboration with scientific organizations and/or civic associations. But data sometimes intervene as both issues in their own and as component of a repertoire of action. A largely individual expression of this convergent trend are data obfuscation practices, in which users aware of issues of profiling deliberately “mess up” with their own data to interfere with the classification logic of artificial intelligence systems. An equivalent example of a “mixed” logic between data-oriented and data-enabled activism, in which the collective element is more evident, is that of open data

initiatives. Claims for a more transparent government by collecting, arranging and publicizing datasets are in this case often merged with demands for the regulation of data itself.

Thus far, we have located some examples on the analytical grid identified by the dimensions “data-as-stake/data-as-repertoire” and “individual/collective action” recognizing the overlap of these categories in the empirical world. This analytical map allows to formulate a working question: Does a political data subject—expression of the contentious politics of data beyond the scattered landscape depicted above—emerge at the intersection between these dimensions?

We might argue that the convergence of data activism along these axes represents a *condition of possibility* for the coalescing of the existing field of action around a subject capable of interpreting the crucial tensions of contemporary datafied societies. Mobilizing around data issues without incorporating datafied tactics within the action repertoire is arguably a limited approach, because it misses out on the historical potential of empowerment through data practices showcased by proactive data activism. Conversely, exploiting data as mere tools for action, without intervening on existing data-related dynamics, lacks the critical attitude towards the power relations embedded in and expressed through data flows, which is a precondition for “redirecting” processes of datafication. As for the second dimension, it is self-evident that individual action alone lacks the scale and network effects required to question structural processes, while collective action which is not grounded on individual, everyday technological practices, can hardly fulfill the requirements for proper social change to happen.

The proposed analytical map and working question can inform a sociologically oriented empirical investigation of the diverse forms of data activism. Moreover, it provides a question for critical reflection in relation to the broader envisioning of an emerging contentious data subject. It is important to stress that the idea of a (more or less) collective subject evoked here is not to be understood in monolithic terms, but as the coagulation of a multiplicitous field of action—a process that can follow many empirical trajectories. We argue that an attention to the ways in which relations to data-as-stakes and data-as-repertoire are folded onto each other offers an excellent vantage point to observe these trajectories.

## Conclusions

This article proposed to explore data infrastructure and data flows as a critical site of contestation. To do so, it investigated the central features of data activism, understood as a critical relation with and towards



data and as action enabled by data practices and structures. It illustrated how data activism intercepts emerging forms of action that interpret data as both repertoires and stakes of contention. It showed that data activism underpins and illustrates an emerging “contentious politics of data”, to be seen as a complement to the notion of “data politics” proposed by Ruppert et al. (2017).

The key contribution of our program lies in the articulation of a social movement perspective on data activism that goes beyond the “means vs. ends” divide, and that avoids both the techno-enthusiasm and the techno-skepticism typical of many contemporary narratives of data activism (including our earlier work). We offer an analytical map and a working question to orient the empirical exploration of the composite assemblage of contentious practices articulated around and through data. Consequently, we ask whether contentious data politics might express the coagulation of a field of action into a political (and/or historical) subject at the intersection of four elements: data understood both as fields of and tools for political contention, and data practices expressed through the entanglement of individual and collective practices.

Our analysis intends to intervene also in the debate around the analytical status of contemporary, digitally mediated social movements (e.g. Bennett and Segerberg, 2013; Beraldo, 2017; Karatzogianni, 2015; Postill, 2018). We believe that looking at the contentious politics of data, and at data activism as its primary manifestation, might help us to update the taxonomy of contemporary digital activism and to give adequate consideration to its infrastructural dimension. It remains an open question to what extent these variably structured, variably ephemeral forms of resistance and political agency emerging around data, data infrastructure and data flows display a cohesive collective identity of their own and the capacity of action sustained over time, and whether it provides the much-needed counter-weight to the impending surveillance capitalism.

### Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was made possible by funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program

(grant agreement Nos. 639379-DATACTIVE 825974-ALEX, with Stefania Milan as Principal Investigator).

### ORCID iD

Stefania Milan  <https://orcid.org/0000-0002-9314-2889>

### Notes

1. The chapter illustrated with examples from a five-year project investigating the emergence and “coming of age” of data activism after Snowden (see <https://data-activism.net>).
2. In this paper, citizenship is intended as exercise of political agency in the feminist tradition (see, e.g. Fraser, 1992; Mouffe, 1992) rather than in its functional meaning of holder of civil rights in a given polity.

### References

- Arora P (2016) Bottom of the data pyramid: Big data and the global South. *International Journal of Communication* 10: 1681–1699.
- Baack S (2018) Practically engaged. The entanglements between data journalism and civic tech. *Digital Journalism* 6: 673–692.
- Bennett LW and Segerberg A (2013) *The Logic of Connective Action Digital Media and the Personalization of Contentious Politics*. Cambridge, UK: Cambridge University Press.
- Beraldo D (2017) *Contentious Branding. Reassembling Social Movements through Digital Mediators*. Amsterdam: University of Amsterdam.
- Berti Suman A (2018) Challenging risk governance patterns through citizen sensing: The Schiphol airport case. *International Review of Law Computers & Technology* 32: 155–173.
- Bijker WE and Law J (eds) (1992) *Shaping Technology/Building Society*. Cambridge, MA: MIT Press.
- Bowker GC and Star SL (2000) *Sorting Things Out*. Cambridge, MA: MIT Press.
- Boyd D and Crawford K (2012) Critical questions for big data. Provocations for a cultural, technological, and scholarly phenomenon. *Information Communication and Society* 15: 662–679.
- Braman S (2009) *Change of State: Information, Policy, and Power*. Cambridge, MA: MIT Press.
- Brunton F and Nissenbaum H (2011) Vernacular resistance to data collection and analysis: A political theory of obfuscation. *First Monday* 16. <https://firstmonday.org/ojs/index.php/fm/article/view/3493/2955>
- Bucher T (2017) The algorithmic imaginary: Exploring the ordinary affects of facebook algorithms. *Information Communication and Society* 20: 30–44.
- Castells M (1996) *The Rise of the Network Society*. Oxford: Blackwell.
- Couldry N and Mejias UA (2018) Data colonialism: Rethinking big data's relation to the contemporary subject. *Television & New Media* 20: 336–349.
- Couldry N and Powell A (2014) Big data from the bottom up. *Big Data & Society* 1: 1–5. <https://doi.org/10.1177/2053951714539277>

- Daskal E (2018) Let's be careful out there...: How digital rights advocates educate citizens in the digital age. *Information Communication and Society* 21: 241–256.
- Dekker M and Duyvendak JW (2013) Reclaiming urban public spaces protest form as means and end in occupy. In: *International RC21 conference*, Berlin. 29–31 August.
- DeLanda M (2006) *A New Philosophy of Society: Assemblage Theory and Social Complexity*. London/New York: Continuum.
- della Porta D and Tarrow S (2005) Transnational processes and social activism: an introduction. In: Porta D and Tarrow S (eds) *Transnational Protest and Global Activism*. Lanham, MD: Rowman & Littlefield, pp.1–17.
- DeNardis L (2009) *Protocol Politics*. Cambridge, MA: MIT Press.
- Dencik L and Cable J (2017) The advent of surveillance realism: Public opinion and activist responses to the snowden leaks. *International Journal of Communication* 11: 763–781.
- Deutch J and Habal H (2018) The Syrian archive: A methodological case study of open-source investigation of state crime using video evidence from social media platforms. *State Crime* 7: 46–76.
- Edwards PN, Bowker GC, Jackson SJ, et al. (2009) Introduction: An agenda for infrastructure studies. *Journal of the Association for Information Systems* 10: 364–374.
- Espeland WN and Stevens ML (2008) A sociology of quantification. *European Journal of Sociology* 49: 401–436.
- Fraser N (1992) Rethinking the public sphere: A contribution to the critique of actually existing democracy. In: Calhoun C (ed.) *Habermas and the Public Sphere*. Cambridge, MA: MIT Press, pp.109–142.
- Gabrys J, Pritchard H and Barratt B (2016) Just good enough data: Figuring data citizenships through air pollution sensing and data stories. *Big Data & Society*. <https://doi.org/10.1177/2053951716679677>
- Gamson WA (1992) *Talking Politics*. Cambridge, MA: Cambridge University Press.
- Gangadharan SP and Jędrzej N (2018) Between antidiscrimination and data: Understanding human rights discourse on automated discrimination in Europe. Department of Media and Communications, LSE, London, UK. <http://eprints.lse.ac.uk/88053/>
- Gerbaudo P and Treré E (2015) In search of the 'we' of social media activism: Introduction to the special issue on social media and protest identities. *Information Communication and Society* 18: 865–871.
- German Working Group on Data Retention (2008) Report of the 'Freedom not Fear' worldwide protests against surveillance. <http://www.vorratsdatenspeicherung.de/content/view/267/79/lang/en/>
- Gitelman L (2013) *Raw Data Is an Oxymoron*. Cambridge, MA: MIT Press.
- Gutierrez M (2018) *Data Activism and Social Change*. Cham: Palgrave MacMillan.
- Hess DJ (2016) *Undone Science. Social Movements, Mobilized Publics, and Industrial Transitions*. Cambridge, MA: MIT Press.
- Hess DJ (2005) Technology-and product-oriented movements: Approximating social movement studies and science and technology studies. *Science, Technology, & Human Values* 30: 515–535.
- Hintz A and Milan S (2018) Through a glass, darkly: Everyday acts of authoritarianism in the liberal west. *International Journal of Communication* 12: 3939–3959.
- Hoffmann AL (2019) Where fairness fails: Data, algorithms, and the limits of antidiscrimination discourse. *Information Communication and Society* 22: 900–915.
- Jasper JM (2011) Emotions and social movements: Twenty years of theory and research. *Annual Review of Sociology* 37: 285–393.
- Jordan T (2002) *Activism! Direct Action, Hacktivism and the Future of Society*. London: Reaktion Books.
- Karatzogianni A (2015) *Firebrand Waves of Digital Activism 1994–2014: The Rise and Spread of Hacktivism and Cyberconflict*. London: Palgrave MacMillan.
- Kavada A (2016) Social movements and political agency in the digital age: A communication approach. *Media and Communication* 4: 8–12.
- Kennedy H (2018) Living with data: Aligning data studies and data activism through a focus on everyday experiences of datafication. *Krisis Journal for Contemporary Philosophy* 1. 1/2018, <http://krisis.eu/living-with-data/>
- Kennedy H, Poell T and van Dijck J (2015) Data and agency. *Big Data & Society* 1–10. July–December 2015, 10.1177/2053951715621569
- Kitchin R (2014) Big data, new epistemologies and paradigm shifts. *Big Data & Society*. (January 2014). doi:10.1177/2053951714528481.
- Latour B (1994) On technical mediation – Philosophy, sociology, genealogy. *Common Knowledge* 3: 29–64.
- Latour B (2005) *Reassembling the Social: An Introduction to Actor-network-theory*. Oxford: Oxford University Press.
- Law J and Hassard J (eds) (1999) *Actor Network Theory and After*. Malden, MA: Blackwell.
- Lehtiniemi T and Ruckenstein M (2018) The social imaginaries of data activism. *Big Data & Society*. <https://doi.org/10.1177/2053951718821146>
- Lupton D (2016) *The Quantified Self*. New York: Wiley.
- Lyon D (2018) *The Culture of Surveillance: Watching as a Way of Life*. Cambridge, UK: Polity Press.
- McAdam D, McCarthy JD, Zald MN (eds) (1996) *Comparative Perspectives on Social Movements: Political Opportunities, Mobilizing Structures, and Cultural Framings*. Cambridge: Cambridge University Press.
- McDonald K (2004) One as another: From social movement to experience movement. *Current Sociology* 52: 575–593.
- McFarlane C (2009) Translocal assemblages: Space, power and social movements. *Geoforum* 40: 561–567.
- Melucci A (1989) *Nomads of the Present. Social Movements and Individual Needs in Contemporary Society*. London: Hutchinson Radius.
- Melucci A (1996) *Challenging Codes. Collective Action in the Information Age*. Cambridge, UK: Cambridge University Press.
- Milan S (2015) From social movements to cloud protesting: The evolution of collective identity. *Information Communication and Society* 18: 887–900.

- Milan S (2017) Data activism as the new frontier of media activism. In: Pickard V and Yang G (eds) *Media Activism in the Digital Age*. New York: Routledge, pp.151–163.
- Milan S (2018) Political agency, digital traces and bottom-up data practices. *International Journal of Communication* 12: 507–527.
- Milan S and ten Oever N (2017) Coding and encoding rights in internet infrastructure. *Internet Policy Rev* 6. 10.14763/2017.1.442
- Milan S and van der Velden L (2016) The alternative epistemologies of data activism. *Digital Culture & Society* 2: 57–74.
- Mosco V (2014) *To the Cloud: Big Data in a Turbulent World*. New York: Paradigm Publishers.
- Mouffe C (ed.) (1992) *Dimensions of Radical Democracy*. London: Verso.
- Nye JS Jr (2011) *The Future of Power*. New York: PublicAffairs.
- Polletta F (2002) *Freedom Is an Endless Meeting: Democracy in American Social Movements*. Chicago: University of Chicago Press.
- Postill J (2018) *The Rise of Nerd Politics. Digital Activism and Political Change*. London: Pluto Press.
- Rodriguez C (2001) *Fissures in the Mediascape. An International Study of Citizens' Media*. Cresskill, NJ: Hampton Press.
- Ruppert E, Isin E and Bigo D (2017) Data politics. *Big Data & Society* 1–7. <https://doi.org/10.1177/2053951717717749>
- Schoeneborn D (2011) Organization as communication. A Luhmannian perspective. *Management Communication Quarterly* 25: 663–689.
- Schrock AR (2016) Civic hacking as data activism and advocacy: A history from publicity to open government data. *New Media & Society* 18: 581–599.
- Star SL (1999) The ethnography of infrastructure. *American Behavioral Scientist* 43: 377–391.
- Star SL and Bowker GC (2002) How to Infrastructure. In: Lievrouw LA and Livingstone S (eds) *Handbook of New Media: Social Shaping and Consequences of ICTs*. Thousand Oaks, CA: Sage, pp.151–162.
- Stier S, Schünemann WJ and Steiger S (2017) Of activists and gatekeepers: Temporal and structural properties of policy networks on Twitter. *New Media & Society* 20: 1910–1930.
- Tarrow S (2011) *Power in Movement: Social Movements and Contentious Politics*. 3rd ed. Cambridge, UK: Cambridge University Press.
- Tilly C (2008) *Contentious Performances*. Cambridge: Cambridge University Press.
- Tilly C and Tarrow S (2015) *Contentious Politics*. 2nd ed. New York: Oxford University Press.
- Torres G (2019) Auditing the state. Everyday forms of institutional resistance in the Mexican freedom of information process. DATACTIVE Working Paper no. 1/2019. <https://data-activism.net/wordpress/wp-content/uploads/2019/05/TorresG.-2019-22Auditing-the-State22-DATACTIVE-WP-Series.pdf>
- Touraine A (1981) *The Voice and the Eye: An Analysis of Social Movements*. New York: Cambridge University Press.
- Treré E (2019) *Hybrid Media Activism*. London: Routledge.
- van der Velden L (2015) Leaky apps and data shots: Technologies of leakage and insertion in NSA-surveillance. *Surveillance & Society* 13: 182–196.
- Wissenbach KR (2019) Accounting for power in transnational civic tech activism: A communication-based analytical framework for media practice. *International Communication Gazette*. (March 2019). doi:10.1177/1748048519832779.
- Zuboff S (2019) *The Age of Surveillance Capitalism*. New York: Profile Books.