

The Alternative Epistemologies of Data Activism

Stefania Milan, Lonneke van der Velden

Abstract

As *datafication* progressively invades all spheres of contemporary society, citizens grow increasingly aware of the critical role of information as the new fabric of social life. This awareness triggers new forms of civic engagement and political action that we term “*data activism*”. *Data activism* indicates the range of sociotechnical practices that interrogate the fundamental paradigm shift brought about by *datafication*. Combining Science and Technology Studies with Social Movement Studies, this theoretical article offers a foretaste of a research agenda on *data activism*. It foregrounds democratic agency vis-à-vis *datafication*, and unites under the same label ways of affirmative engagement with data (“*proactive data activism*”, e.g. data-based advocacy) and tactics of resistance to massive data collection (“*reactive data activism*”, e.g. encryption practices), understood as a continuum along which activists position and reposition themselves and their tactics. The article argues that *data activism* supports the emergence of novel epistemic cultures within the realm of civil society, making sense of data as a way of knowing the world and turning it into a point of intervention and generation of data countercultures. It offers the notion of *data activism* as a heuristic tool for the study of new forms of political participation and civil engagement in the age of *datafication*, and explores *data activism* as an evolving theoretical construct susceptible to contestation and revision.

Keywords: *datafication*; *data activism*; democratic agency; Big Data epistemologies; Social Movement Studies; Science and Technology Studies.

Acknowledgements

This article was researched with the support of a Starting Grant of the European Research Council, awarded to Stefania Milan as Principal Investigator (ERC-StG-2014-639379; <https://data-activism.net>). The authors wish to thank the editors of this special issue; the anonymous reviewers; Prof. Sandra Braman

(Texas A&M); the attendees to the *Contentious Data* workshop (Amsterdam, 15–16 September 2016); the participants to the data activism sessions at the 2016 conferences of the International Communication Association (Fukuoka, 10 June) and of the Association of Internet Researchers (Berlin, 7 October) for the invaluable critical questions and insights. Our gratitude extends also to the DATACTIVE team at the University of Amsterdam for the inspiring conversations.

Introduction

It was the summer of 2012 when the Italian hacker and artist Salvatore Iaconesi was diagnosed with brain cancer. Still hospitalised, he “hacked” his medical data to make them accessible to non-experts, published them online, and launched *La Cura*, a “participatory performance aimed at redefining the word ‘cure’, bringing it out of hospitals [...] back into society” (Iaconesi 2016). He invited anyone in the world to send him a cure. Over half a million people contributed advice and emotional support.

In 2012, British unemployed finance worker Eliot Higgins, also known as Brown Moses, turned into a weapon analyst by systematically monitoring videos of the Syrian conflict on YouTube, experimenting with a new form of “social media weapon tracking”. In 2014, he founded Bellingcat.com, a platform for self-taught open-source intelligence (OSINT) analysts and investigative journalists, which gained the praise of Amnesty International, among others (O’Brien 2013).

In order to support citizen witnessing and improving grassroots documentation efforts, the Guardian Project has developed the mobile application CameraV. It captures and archives image metadata as an extra layer of contextual verification, making possible the potential use of the footage as evidential proof in court. The footage is password-protected, encrypted and stored on the user’s device (The Guardian Project: n.d.).

With the progressive datafication of an ever-expanding range of human activities – from personal health to interpersonal connections, from public administration to security – people become increasingly aware of the critical role of information in contemporary societies. The anecdotes opening this article belong to the realm of the transformative experiments that see citizens putting data to new uses, developing “new rationalities and alternative social imaginaries around datafication” (Baack 2015: 8). These “moments where meaningful change can occur, even if those changes are [...] tinged with technocracy” (Schrock 2016: 583), speak to the unprecedented possibility for larger publics to foster social change by engaging in data politics. But today’s massive data collection is also employed for monitoring people, as we know from the classified information leaked by whistleblower Edward Snowden (Lyon 2015). Users’ online activities are regularly “sucked up as data, quantified and classified, making possible real-time tracking and monitoring” (Lyon 2014: 4), which generates an unparalleled power asym-

metry between the state and its citizens (Brunton/Nissenbaum 2015). What for some observers is the hopeful “industrial revolution of data” (Hellerstein 2008), represents a new form of “surveillance capitalism” for others (Zuboff 2015).

While the industry as well as the state apparatus have long acknowledged the value of massive data collection for their activities, the so-called civil society – the realm of human activity outside the state and the market – is now slowly but steadily catching up and turning “big data” to its own ends. Civic hackers “requesting, digesting, contributing, modelling and contesting data” (Schrock 2016: 584) and civic tech activists (Russon Gilman 2016) seek to improve institutional output and democratic governance by means of software and data. Tech-savvy people experiment with arguably “radical” (Birchall 2014) transparency devices for the analysis of previously closed or hidden data (Heemsbergen 2014). In the footsteps of WikiLeaks, a range of whistle-blower platforms such as GlobalLeaks and PubLeaks has emerged, allowing for the tech-mediated protected transfer of data to journalists and/or the citizenry at large. Activists engage in “information activism” by monitoring powerful actors and curating datasets (Ganesh/Hankey 2015). Data visualisation is entering the activists’ skillset; capacity building projects and knowledge sharing manuals increasingly target low-skill users (Tactical Tech Collective 2013; 2016). Activists become progressively aware of how the use of digital tools plays out technically, politically, and economically, rethinking what the “politics of data” means for their own practice, especially in the context of surveillance (Ganesh/Hankey 2015). They increasingly seek to counter massive data collection by means of resistance and obfuscation. Although the circumvention of surveillance is a long-standing practice amongst social movements and certainly predates datafication (see, e.g., della Porta 1995), dedicated events, trainings, and off-the-shelves tools to secure digital communications have mushroomed over the last few years (Aouragh et al. 2015), and encryption features prominently in counter-surveillance initiatives (Gürses/Kundnani/van Hoboken 2016). We subsume these diverse manifestations of an emerging “critical” attitude towards datafication under the rubric of “data activism”.

This theoretical article explores the notion of data activism as a heuristic, polysemic tool to think politically about big data from the perspective of users and citizens, analysing in particular the contemporary evolution of activism vis-à-vis datafication. In what follows, we offer a conceptual map to approach grassroots engagement with data by combining insights from Social Movement Studies (SMS) and Science and Technology Studies (STS). First, we situate the notion of data activism in the domain of digital activism. Second, we show that data activism practices point to the emergence of novel epistemic cultures arising from within the civil society realm. Third, we explore the potential of data activism as a heuristic tool for empirical analysis and theoretical development. Finally, we sketch a future research agenda on data activism. In so doing, we intend to contribute to the study of “the variable ways in which power and participation are constructed and enacted” (Couldry/Powell 2014: 1) in bottom up data practices.

Defining data activism

Big data evokes a broad set of socio-technical phenomena enveloped in quasi-mythological narratives that univocally emphasise possibility and magnitude. Kitchin, however, showed that big data goes well beyond a matter of data volume, and “is characterised by being generated continuously, seeking to be exhaustive and fine-grained in scope, and flexible and scalable in its production” (2014: 2). For the purposes of our analysis, which underscores human agency in relation to data and technology, we understand big data as information-related tasks whose complexity requires individuals to take action with the support of software (cf. Cukier/Mayer-Schoenberger 2013). We take big data to refer also to “human subjects data, since it mainly (though not universally) consists of data about, and produced by, people” (Dalton/Taylor/Thatcher 2016: 2). While deliberately glossing over the nature of the data in question, these definitions stress the transformative and empowering potential of (any) data, focusing on the complexity of the tasks that can be performed rather than data magnitude, and on the socio-cultural meaning of said tasks. We consider data as a technology in itself, looking back to the semantic roots of the latter: the ancient Greek noun *techne*, “making”, referred to what both art and engineering have in common (Braman 2004: 4). Seeing data as a technology allows us to call attention to what people do and can do with it.

The notion of data activism is not entirely new. Earlier characterisations have focused on particular instances of data-driven contestation, emphasizing aspects as different as the affective potential of the engagement with data, the contribution to institutional reform, and the relation with governance. For instance, Renzi and Langlois explored how data partake in generating individual and collective action, which they describe as “new modes of being and acting together through a direct engagement with data and the means of its mobilisation” (2015: 203). Their notion of data activism highlights how alternative media are re-appropriations but also sites for experimentation that allow for the generation of new affective bonds. Focusing instead on the relation between the state and its citizens, Schrock coined the heading of “data activism and advocacy” to identify contemporary civic hackers who “operate through a range of data-driven political modes [...] to bring about systemic change [...] [and] participate in civic data politics” (2016: 591), with the aim to ameliorate the governance process and output.

To be sure, data activism comes close to other variations of the wide-ranging category of cyber- or digital activism (McCaughey/Ayers 2003), such as hacktivism (Jordan/Taylor 2004), statactivism (Bruno/Didier/Prévieux 2014), and information politics (Jordan 2015). Hacktivism indicates “collective action in cyberspace that addresses network infrastructure or exploits the infrastructure’s technical and ontological features for political or social change” (Milan 2015a: 550). Similarly, data activists might engage with the infrastructure of platforms and code. They, too, “engage in politically motivated use of technical expertise in view of fixing society through software and online action” (*ibid.*). Yet, our notion of data

activism outgrows the engagement with infrastructure to embrace information and knowledge as a broader category of intervention.

Looking at the prescriptive potential of numbers, Bruno and colleagues coined the notion of statactivism, a portmanteau of statistics and activism identifying the bottom-up experiments aimed at the re-appropriation of statistics' power of denunciation and emancipation (Bruno/Didier/Prévieux 2014). The role of statistics in statactivism is two-fold: to criticise reality and to represent it, mobilizing "numbers, measurements and indicators as means of denunciation and criticism" (Bruno/Didier/Vitale 2014: 199). Similar to data activists, statactivists operate simultaneously inside and outside institutions, and numbers are a means of both disclosure and affirmation (*ibid.*). However, data activism as a theoretical category embraces a broader range of encounters with datafication: on the one hand, it comprises but also exceeds numbers as source of truth claims, and on the other, it understands datafication as conveying risks as well. Consequently, data activism bears the promise to incorporate a wider set of tactical responses.

Finally, Jordan's expansive notion of information politics (2015) engages with the contentious reactions to mass digital cultures along with network and control protocols, pointing to antagonistic formations in which information politics becomes part of conflicts over exploitation and liberation. While our notion of data activism presupposes the possibility for contention and rebellion, we understand data activism as a series of nuanced phenomena that position themselves in a continuum between contestation and recognition. In this vein, we offer the notion of data activism as yet another possible manifestation of activism in the information society – one that, however, explicitly engages with the new forms information and knowledge and their production take today, challenging dominant understandings of datafication. As such, thinking in terms of data activism has something to offer to the reflection on the ubiquity of digital communication and mediation dynamics in the platform society (e.g., Castells 2009; van Dijck 2013), as well as on the nature of contemporary activism and its evolving protest/media configurations (e.g., Bennett 2012; Bennett/Segerberg 2013; Milan 2015b).

The data activism rubric embraces the composite series of sociotechnical practices that, emerging at the fringes of the contemporary activism ecology, interrogate datafication and its socio-political consequences. We deliberately unite under the same label two variations of grassroots data politics – affirmative engagement with data and resistance to massive data collection – which are often considered contrasting with each other. Yet, they both address and interrogate the fundamental paradigm shift brought about by datafication.

Our notion of data activism foregrounds democratic agency vis-à-vis datafication. Two elements rest at its core: its sociotechnical nature and the mobilisation factor. First, following the path traced by STS, we emphasise the fundamental interaction between people, information, and technology that constitutes emerging formations of data activism. Data activism is deeply rooted in, and thus enabled and constrained by data and software, both its availability and its pursuit – and this

special relation shapes tactics, identities, and modes of organizing. Second, from a SMS perspective data activism can be seen as an form of socio-political mobilisation, as it brings people (and information and technology) together for some kind of action variably contentious in nature, and explicitly addressing, confronting, or engaging with datafication. Mobilisation here embraces both discrete events – individual and collective acts of appropriation of data, but also of dissent, subversion, and resistance to data collection – and the overall process that subtends to the emergence of data activism – namely, the raising popular concern signaling a fundamental change in perspective and attitude towards datafication that is slowly emerging within civil society.

The social forces supporting these incipient activism practices are not new. Users have long appropriated and repurposed media technology for self-expression and social change, experimenting with up-and-coming media technology – be it phone, print, radio, fax, or the internet (see Marvin 1990; Downing 2011). The most tech-aware amongst radical openness activists have experimented with technical ways of circumventing state control since the 1980s – advocating, for example, for a crypto anarchy in communications which would be made possible by cryptography (May 1992). Emerging from the smoking ashes of their predecessors and variably intersecting other contemporary trends like citizen science (Irwin 1995) and sensing (Gabrys 2016), biohacking (Delfanti 2013), citizen data journalism (Gray/Chambers/Bounegru 2012), the quantified self (Lupton 2016) and the transparency movements (Sifry 2011), current data activism practices are largely modelled after hacker cultures and rituals (Levy 1984) and the do-it-yourself approach of makerspaces, hacklabs, and hackerspaces (Maxigas 2012). Similar to grassroots radio producers (Dunbar-Hester 2012), the radical techies of the 1990s (Milan 2013), and the open-source movement (Coleman 2013), data activists often preach technical engagement as a way of confronting elite expertise and taking control over the technology of daily use. From their forerunners, data activists have borrowed the hands-on attitude, and the notions of access to information, code tinkering, collaboration, and world improvement through technical fixes. And because datafication is such a prominent feature in public life, data activism, as a mode of dealing with it, might progressively appeal to more diverse communities of concerned citizens, beyond the expert niche of previous incarnations of tech activist engagement.

Data activism shapes novel epistemic cultures

Datafication alters “the conditions under which we can make sense of our world and our own actions”, affecting “our capacity to act with agency” (Baack 2015: 1). Big data has far-reaching epistemological consequences, affecting “how knowledge is produced, business conducted, and governance enacted” (Kitchin 2014: 2). It “reframes key questions about the constitution of knowledge” (boyd/Crawford

2012: 665). Algorithms, too, have the “ability to shape perceived realities [...] not only by enabling certain representations of the world around us, but also by enticing us to internalise these realities and make them our own” (Renzi/Langlois 2015: 202). Further, big data and their representations tend to reify the future, crystallizing *a reality* – usually quantified, and presented as neutral and infallible – as the only possible one and narrowing down the options for alternatives (cf. Chun 2011). In other words, what is known as big data constitutes a novel, powerful system of knowledge with its own epistemology that is to say a specific way of framing, packaging, presenting and activating information and knowledge.

Framing, packaging and presenting data are productive exercises, which have the potential to alter not only our vision of the world, but also our own theory of knowledge, so to speak. The current emphasis on computational analysis and machine learning as core (and qualitatively superior) ways of understanding the social world, moulds the way people relate to information and knowledge. Furthermore, “big” data are often seductively staged, accompanied as they typically are by attractive visualisations and graphs that simplify reality and communicate it in immediate and efficient ways (see Cairo 2012). What is then the role of data activism in these dynamics? If data is not a given nor is it ever raw (Gitelman 2013), data activism can be seen as an exercise in creating alternative ways of seeing the world, while opening up questions about the positivism ethos of the so-called “data revolution”.

Fraser (2005) criticised the state-centric “politics of representation”, identifying in the control by the elites over the framing of political representation a form of injustice. “Making discourses”, on the contrary, entails “producing new languages or modifying old ones so as to find words for novel phenomena” (Jasanoff 2004: 39–41). We argue that data activism practices signal that new epistemic cultures are emerging within the civil society realm: a way (or, better, ways) of making counter-discourses and data countercultures that challenge the mainstream readings of reality.

Epistemic cultures shape the way we relate to knowledge and its validation, how we understand and filter the world around us as well as our experiences. They represent “cultures of creating and warranting knowledge” (Knorr Cetina/Reichmann 2015: 873). When the notion was introduced in laboratory studies in the late 1990s, it disrupted the idea of “epistemic unity” of the sciences, by stressing diversity in modes of knowledge making (*ibid.*). This concept invites us to look at the “specific strategies that generate, validate, and communicate scientific accomplishments”, and to take into account the complex “relationships between experts, organisational formats, and epistemic objects” (873–4). The emerging epistemic cultures propelled by data activism point to innovative and potentially transformative ways of relating to big data and its consequences.

Postulating a critical/active engagement with data, its forms, dynamics, and infrastructure, data activists function as producers of counter-expertise and alternative epistemologies, making sense of data as a way of knowing the world

and turning it into a point of intervention. They challenge and change the mainstream politics of knowledge, and operate as mediators between the dominant “algorithmic culture” (Striphias 2015) and the citizenry at large. They operate as a “critical community” composed of “critical thinkers who have developed a sensibility to some problem, an analysis of the sources of the problem, and a prescription for what should be done”. These critical thinkers “seek acceptance of a new conceptualisation of a problem [...] [and] attempt to influence the conceptual framework used to think about a cluster of issues” (Rochon 1998: 22–23).

For its rich history of uncovering the politics and partiality of technology and scientific knowledge (Winner 1980; Haraway 1988), STS can help to understand fully grasp the potential of the alternative epistemic cultures of data activism. STS shows that science and technology are neither purely technical nor social, but are co-produced through very specific settings that bring along particular (material) affordances, situated practices, and tacit knowledge. This means that values and modes of working stemming from these settings become *inscribed* in the way science and technology develop. STS scholars have always been interested in what happens when the range of actors participating in this process of co-production is stretched up. They have paid attention to these arguably democratic moments in which various actors have influenced technological development, for example through the study of activist appropriations of science and technology by patient movements (Epstein 1995) or technology activists (Hess 2005). We propose two possible focus areas for an STS analysis of the alternative epistemic cultures of data activism.

STS has explained how particular objects and methods that traditionally belong to the epistemic community and culture of scientists and experts have been opened up to a different population, sometimes in democratically designed settings, and how along this trajectory these objects and methods have been malleable to change. For example, Epstein’s study (1995) on clinical trials on HIV patients showed how the mainstream scientific epistemology can be altered by the emergence of a grassroots, alternative one. The protocol for clinical trials wanted these to be conducted on a clean population with no prior medicine use. But clinical trials were among the few ways HIV patients could get access to treatment. Patients contested this practice on the ground that it did not mirror the “real world”, given only few individuals had a clean history of medicine use. A group of patients resorted to scientific literature to educate themselves, constructing novel discourses about clinical trials. An alternative epistemic culture emerged in which up-and-coming knowledge and arguments collided with the official values and discourses concerning access to medicine. Eventually, the patients managed to have their alternative epistemology taken serious. Epstein’s analysis provides a valuable lesson for observers of citizens’ responses to datafication: in a context in which “big data” becomes an important reference point for “new” ways of knowing, one can track alternative data narratives and emerging forms and dynamics of counter-expertise.

STS are also known for mobilizing a particular vocabulary and method by carefully tracing how science and technology develop through sociotechnical assemblages. Especially scholars of Actor Network Theory (ANT) – which treats objects like machines as part of heterogeneous networks with humans – have argued that the material agency of artifacts should not be neglected (Callon 1986; Akrich 1992; Latour 2005). In later works, ANT scholars have also tackled the issue of the democratic agency of things and have probed the exploratory question of what could be the potential of their democratic agency (Latour 2005; Latour and Weibel 2005; Marres 2011). In his *Politics of Nature* (2004), Latour designed a “parliament of things” in which he outlined how proper representation of “things” would look like, repurposing rather classic notions of representation and voice. Focusing instead on “issue politics” (Marres 2005), Marres (2011) looked at how material devices participate in the “articulation of issues”. In her view, everyday material devices facilitate the articulation of concerns because they provide particular logics for action and for dealing with problems. Material devices therefore contribute to what we understand as matters for the public, and shape our social behavior, too. An era in which we see “data inscriptions” everywhere around us, be it through narratives based on numbers or data visualisations and infrastructure, there is a real need for a critical look at the agency gained (or lost) by and through these inscriptions.

STS thus helps us, on the one hand, to zoom in on the participation of new critical actors in knowledge generation and, on the other, to appreciate the political and potential democratic impact of the material agency enshrined in data, algorithms and infrastructure. Both appear very relevant for our exercise of understanding how new epistemic data cultures shape up. If we consider data not as something given but as a *techne*, i. e. a form of “making” as argued above, we ought to take these critical forms of data making seriously.

Data activism as a heuristic tool

We argue that data activism constitutes a valuable heuristic tool for the study of political participation and civil engagement in the age of datafication, a sort of “lens” through which one can investigate how activism evolves in relation to big data. Data activism is a composite, polyfunctional, holistic and polysemic notion. First, it is a composite concept because, as its interdisciplinary origin suggests, it is made of parts: at the minimum people, (variably sustained forms of) contention, information, and technology. Second, it is polyfunctional, as it can be read through diverse disciplinary lenses, and can be domesticated to investigate different dynamics and relations, between and within people, information, technology, and the state/industry complex. Third, data activism encourages us to adopt a holistic perspective, as it allows us to bring under the same umbrella the two facets of the “data revolution”: the productive and the harmful qualities of the “big data”

phenomenon, the opportunities and the threats brought forward by datafication, as well as the varied response of the grassroots, be it recognition or resistance. By taking data activism as a whole, we are able to bypass these dichotomies, acknowledging that they ultimately adhere to the same phenomenon of data countercultures/practices, whose components cannot be fully understood if taken in isolation. Finally, as a consequence of its holistic nature, the notion of data activism is polysemic, since “activism” embraces practices of resistance and instances of appropriation as discrete but complementary means to achieving political goals, allowing very distinct attitudes towards institutions and social norms to coexist.

As a heuristic device of sociological nature, data activism is simultaneously flexible and specific enough as to guide and support the analysis of a multifaceted empirical phenomenon where the “social” dimension of taking action (i.e., mobilising and organising) is mediated by the “technical” of information and technology. However, the notion has yet to enter in conversation with the empirical field, and thus can be considered in its guise of theoretical construct, or “the consequence[s] of theories or conventions” that “exist[s] primarily to serve the interest of investigators” (Ragin 1992: 8). This does not deny the existence of empirical units out there. On the contrary, it affirms that these (and especially the borders defining them) are the consequence of the researchers’ choice on the basis of theoretical considerations and empirical intuitions. In this respect, we expect the theoretical category of data activism to “coalesce in the course of the research. Neither empirical or given, [it is] gradually imposed on empirical evidence” (Ragin 1992: 10). In sum, the notion can be seen both as a product of the attempt to define the object of investigation by connecting two approaches, respectively resistance and engagement, that might not be necessarily or explicitly connected by activists themselves; and a working hypothesis susceptible to contestation, revision and refinement in the course of the research.

The collective action dimension of data activism at the heart of this notion can be best analysed using the conceptual toolkit of social movement studies. In what follows, we provide suggestions for the operationalisation of data activism as a heuristic tool, engaging in an exercise of preliminary mapping based on desk research as well as Milan’s earlier studies on radical tech activism (Milan 2013; 2015a), which bears some resemblance to data activism.

Zooming in on the many ways in which individuals and groups engage with data politics, we identify two main approaches: datafication is interpreted as a challenge to individual rights *or* as a novel set of opportunities for advocacy and social change. This translates into a varied action repertoire, i.e. the range of tactics activists may adopt to pursue their goals (Taylor/Van Dyke 2004), that is positioned along a continuum between two kinds of responses that are not necessarily in contradiction with each other: contentious attitudes such as obfuscating and resisting *vs.* embracing and making the most of datafication. Under the data activism umbrella we therefore identify and assemble at least two forms of action repertoires that are often considered as antithetical. Those activists who perceive

massive data collection as a threat to their values, freedoms and activities, often use technical means like encryption or anonymity networks to resist monitoring by state and corporations. We understand this approach as “*re-active* data activism”, underscoring the fact that activists react to exogenous threats trying to defend their values, beliefs and practices and/or undermine those dynamics and mechanisms they reject. On the opposite end of the spectrum we position those activists who consider the increasing availability of data as an unprecedented, powerful opportunity to provoke social change. They create, mobilise, solicit, appropriate, or crunch data in view of supporting alternative narratives of the social reality, questioning the truthfulness of other representations, denouncing injustice and advocating for change. We label this form of data activism “*pro-active*”, in order to highlight how activists take charge and engage in hands-on practices of appropriation and re-use. “Reactive” and “proactive” represents two facets of the same phenomenon: both take information as a constitutive force in society capable to shape social reality (Braman 2009). Analytically, these labels constitute ideal-types: while they are guilty of approximation and generalisation, they serve the purpose of drawing attention to what various approaches to data activism might have in common.

When we look at action repertoires, we can detect some similarities with other technology-oriented movements (Hess 2005) and subcultures. Like the open-source movement, data activism might concern the promotion of alternative technologies and shaping the surrounding policies; sometimes this might entail some form of collaboration with the industry (*ibid.*). This is particularly true amongst proactive data activists, when they engage in software development or collaborate with state institutions. Reactive data activists, on the contrary, tend to uphold an adversarial attitude towards the state/industry complex; their action repertoire, which includes self-defence, civil disobedience and disruption, is inspired to the radical tech activists of the early days (Milan 2013). Both share the hands-on approach that postulates first-person engagement with information and technology seen as objects of intervention.

The selection of tactics, however, is known to reflect the activists’ cultural and ideological preferences (Milan 2013). Some data activists share the “engineering philosophy to ‘make things work’” of open source developers, coupled with a certain “insistence on adopting a technocratic approach to solving societal problems and to bypassing (‘hacking’) legislative approaches” (Berry 2008: 102). Ideologically, though, data activists appear to cover a broad ground, from the self-organisation and -determination of anarchism to anti-state anarchocapitalism tendencies, to an emphasis on transparency and openness to one on human rights, social justice and the fight against inequality. The notion of collective identity can help us mapping ideological and cultural inclinations, the way they are built, reproduced and maintained over time. Collective identity refers to the process through which “a collective becomes a collective” (Melucci 1996: 84): a sort of “esprit de corps” (Blumer 1939) that holds people together. Some data activists are characterised by a “technical identity”, which is suggestive of a closer

relationship with technology than that of lay users (Dunbar-Hester 2012). Unitary and adversarial ideas of democracy often coexist, bearing distinct ideological preferences that have consequences, for instance preventing collaboration with state institutions. Certainly, data activism as a whole still lacks a collective identity of its own, visible for example in the fact that only seldom do activists define themselves primarily as “data activists”.

Finally, data activism concerns both individuals and groups, as taking action does not necessarily call for a collectivity to mobilise as well. Many contentious actions in data activism are performed at the individual level: think, for example, of engaging in programming or inserting data into a spreadsheet. While it is mainly the group that takes action, and it takes peer-to-peer recognition and interaction to give meaning to action, individual acts of data activism like encrypting private communication do matter, because “there is protest even when it is not part of an organised movement” (Jasper 1997: 5).

Towards a research agenda on data activism

Cycling back to where we started, we can revisit the introductory examples in light of our reflections on data activism. The first two cases represent instances of proactive data activism, while the third represents both. *La Cura* engages in a sort of big-data update of the patient activism described by Epstein. Faced with the enclosure of medical knowledge, Iaconesi managed to lay the ground for an innovative epistemic culture in which the values of openness and collaboration were inscribed into the discourses and practices of health data. His hack enriched the action repertoire by bringing technology and alternative data cultures to bear on “official” knowledge discourses. By repurposing OSINT, otherwise usually associated with intelligence services, and offering trainings and tools to perform investigative data journalism, Bellingcat activists engage in a public learning experiment inspired to open source principles (cf. Glassman and Kang 2012). Apps like CameraV, built to facilitate the uptake of image material in evidential arenas, perform data activism by carefully curating metadata. Here the (encrypted) storage and formatting of data becomes a socio-technical configuration accommodating particular causes (cf. van der Velden 2015). These cases constitute data activism because they *question data agency*, inviting other people to produce their own data inscriptions and to actively “shape issues” in the datafied society, meanwhile contributing to the creation of alternative epistemologies of what data means and represents. But they raise questions, too: how do activists make “data count”? In other words, how do they determine what constitutes relevant and “true” data? To what extent are these “critical” approaches to datafication? How do we draw the line between critique and positivism?

In this article, we suggested that data activism supports the emergence of novel epistemic cultures within civil society, contributing alternative narratives of

our datafied social reality. We offered the notion of data activism as a heuristic tool to study citizens' engagement with datafication, emphasizing data activism as an evolving theoretical construct – a working hypothesis susceptible to contestation and revision throughout the empirical research. But how might a research agenda on data activism look like?

Potential questions address both the emerging ontologies and epistemologies of data activism as an empirical phenomenon. Thinking about ontologies of data activism implies investigating the nature of collective action in grassroots data politics, by means of naming and defining of types, properties and interrelations. To this end, we proposed a first typology of data activism that situates two seemingly contradictory approaches – proactive engagements with data (e.g., data-based advocacy) and reactive attitudes (e.g., using and developing encryption tools) – in a continuum that encapsulates the nuanced grassroots responses to the fundamental paradigm shift imposed by datafication. We might then ask who is involved in particular practices of data making (or, in STS vocabulary, who are the main actors and what are the implications of this composition?) How do technical practices, social values and ideologies play out in these emergent alternative epistemic cultures? What kinds of practices are performed to make these alternative data and narratives count? (In other words, how are questions of relevance, truth or political goals determined, and do these decisions connect to technical affordances of, e.g., analytics tools?) Do the activists differ from the state or the industry in the way they treat and mobilise data? How do data activists influence the agenda of other movements? How do they incorporate reflections on gender, privilege, digital labour? What does “critical engagement” mean in the context of data activism?

We should also ask higher-order questions concerning the interplay between data activism and datafication, and between data activism and the evolution of democracy more in general. How is data constructed or enacted, and what values and modes of understandings are inscribed in this process? (How is data “co-produced”?). What are the political and democratic possibilities of data? How do data and data activism partake in shaping larger societal issues, and how do they affect our behaviour in relation to those? What lessons might data activism hold for contemporary social movements and the citizenry at large? How does data activism contribute to (re-)define the ways we think citizenship, engagement, and democratic deliberation and participation?

But the assumption that activist data interventions express articulations of unfolding alternative epistemic cultures interrogates our modes of knowing about data activism in itself, calling for a reflection on research epistemology and triggering a set of meta-questions that query and situate the key role of researchers in labelling and interpreting social reality. We consider reflecting on our own epistemologies as a moral imperative, if we are to respect grassroots data politics and the activists' effort to question dominant narratives and ways of knowing the world. For example, what constitutes “critical” data practices, and who and

how sets the parameters for assessing such critique and its validity? This contemplation of research epistemologies must go hand in hand with methodological questions concerning our engagement with the field. What are the key moments for investigating data activism? (In an STS fashion, what is the “controversial” moment when data epistemic cultures get interrupted and reorganised?) What methods are the most conducive to the investigation of data activism practices? What new, participatory methods can we envision, which can best incorporate the criticism to dominant epistemologies of knowing advanced by data activism? Are there ways to do *data-activist research* and what is specific to them? We define data-activist research as a type of co-generative inquiry and a way of conducting “engaged research” (Milan 2014) that turns (research) data into points of intervention, supporting grassroots efforts. To envisage how data-activist research might look like, how it might be practiced, and whether it is desired at all, is a challenging task for both the research and the activist community.

References

- Akrich, Madeleine (1992): “The De-Scription of Technical Objects.” In: Wiebe E. Bijker/John Law (eds.), *Shaping Technology/Building Society. Studies in Sociotechnical Change*, Cambridge, MA: MIT Press, pp. 205–24.
- Aouragh, Miriyam/Gürses, Seda/Rocha, Jara/Snelting, Femke (2015): “Let’s First Get Things Done! On Division of Labour and Techno-Political Practices of Delegation in Times of Crisis.” In: *Fiberculture* 26, pp. 208–35.
- Baack, Stefan (2015): “Datafication and Empowerment: How the Open Data Movement Re-Articulates Notions of Democracy, Participation, and Journalism.” In: *Big Data & Society* 2/2, pp. 1–11.
- Bennett, Lance W. (2012): “The Personalisation of Politics: Identity, Social Media, and Changing Patterns of Participation.” In: *The Annals of the American Academy* 644, pp. 20–39.
- Bennett, Lance W./Segerberg, Alexandra (2013): *The Logic of Connective Action: Digital Media and the Personalisation of Contentious Politics*, Cambridge, UK: Cambridge University Press.
- Berry, David M. (2008): *Copy, Rip, Burn: The Politics of Copyleft and Open Source*, London: Pluto Press.
- Birchall, Clare (2014): “Radical Transparency?” In: *Cultural Studies ↔ Critical Methodologies* 17/4, pp. 77–88.
- Blumer, Herbert G. (1939): “Collective Behavior.” In Robert E. Park (ed.), *An Outline of the Principles of Sociology*, New York: Barnes & Noble, pp. 221–80.
- boyd, dana/Crawford, Kate (2012): “Critical Questions for Big Data. Provocations for a Cultural, Technological, and Scholarly Phenomenon.” In: *Information, Communication & Society* 15/5, pp. 662–79.

- Braman, Sandra (2004): "The Meta-Technologies of Information." In: Sandra Braman (ed.), *Biotechnology and Communication: The Meta-Technologies of Information*, Mahwah, NJ: Lawrence Erlbaum Associates, pp. 3–36.
- Braman, Sandra (2009): *Change of State: Information, Policy, and Power*, Cambridge, MA: MIT Press.
- Bruno, Isabelle/Didier, Emmanuel/Prévieweux, Julien (2014). *Statactivisme. Comment lutter avec des nombres*, Paris: La Découverte.
- Bruno, Isabelle/Didier, Emmanuel/Vitale, Tommaso (2014): "Statactivism: Forms of Action between Disclosure and Affirmation." In: *Partecipazione & Conflitto* 7/2, pp. 198–220.
- Brunton, Finn/Nissenbaum, Helen (2015): *Obfuscation: A User's Guide for Privacy and Protest*, Cambridge, MA: MIT Press.
- Cairo, Alberto (2012): "Facing the Dataclysm." In: *The Functional Art: An Introduction to Information Graphics and Visualisation*, San Francisco: New Riders (<http://www.thefunctionalart.com/2014/09/facing-dataclysm.html>).
- Callon, Michel (1986): "Some Elements of a Sociology of Translation: Domestication of the Scallops and the Fishermen of St Brieuc Bay." In: John Law (ed.), *Power, Action and Belief: A New Sociology of Knowledge?*, London: Routledge, pp. 196–223.
- Castells, Manuel (2009): *Communication Power*, Oxford: Oxford University Press.
- Chun, Wendy (2011): *Programmed Visions Software and Memory*, Cambridge, MA: MIT Press.
- Coleman, Gabriella (2013): *Coding Freedom: The Ethics and Aesthetics of Hacking*, Princeton: Princeton University Press.
- Couldry, Nick/Powell, Alison (2014): "Big Data from the Bottom up." In: *Big Data & Society* 1/2, pp. 1–5.
- Cukier, Kenneth/Mayer-Schoenberger, Viktor (2013): "The Rise of Big Data: How It's Changing the Way We Think about the World." In: *Foreign Affairs* 92/3, pp. 28–40.
- Dalton, Craig M./Taylor, Linnet/Thatcher, Jim (2016): "Critical Data Studies: A Dialog on Data and Space." *Big Data & Society* January-June, pp. 1–9.
- Delfanti, Alessandro (2013): *Biohackers: The Politics of Open Science*, London: Pluto Press.
- della Porta, Donatella (1995): *Social Movements, Political Violence, and the State*, Cambridge, UK: Cambridge University Press.
- Downing, John D. H. (ed.) (2011): *Encyclopedia of Social Movement Media*, Thousand Oaks, CA: Sage.
- Dunbar-Hester, Christina (2012): "Soldering Toward Media Democracy: Technical Practice as Symbolic Value in Radio Activism." In: *Journal of Communication Inquiry* 36/2, pp. 149–169.
- Epstein, Steven (1995): "The Construction of Lay Expertise: AIDS Activism and the Forging of Credibility in the Reform of Clinical Trials." In: *Science, Technology & Human Values* 20/4, pp. 408–437.

- Fraser, Nancy (2005): "Reframing Injustice in a Globalizing World." In: *New Left Review* 36, pp. 69–88.
- Gabrys, Jennifer (2016): *Program Earth: Environmental Sensing Technology and the Making of a Computational Planet*, Minneapolis: University of Minnesota Press.
- Ganesh, Maya Indira/Hankey, Stephanie (2015): "From Information Activism to the Politics of Data." In: *Fiberculture* 26, pp. 275–286.
- Gitelman, Lisa (2013): "Raw Data" is an Oxymoron, Cambridge, MA: MIT Press.
- Glassman, Michael/Kang, Min Ju (2012): "Intelligence in the Internet Age: The Emergence and Evolution of Open Source Intelligence (OSINT)." In: *Computers in Human Behavior* 28/2, pp. 673–82.
- Gray, Jonathan/Chambers, Lucy/Bounegru, Liliana (eds.) (2012): *The Data Journalism Handbook*. Sebastopol: O'Reilly.
- Gürses, Seda/Kundnani, Arun/Hoboken, Joris van (2016): "Crypto and Empire: The Contradictions of Counter-Surveillance Advocacy." In: *Media, Culture & Society* 38/4, pp. 576–590.
- Haraway, Donna (1988): "Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective." In: *Feminist Studies* 14/3, pp. 575–599.
- Heemsbergen, Luke J. (2014): "Designing Hues of Transparency and Democracy after WikiLeaks: Vigilance to Vigilantes and Back Again." In: *New Media & Society* 17/8, pp. 1340–1357.
- Hellerstein, Joseph (2008): "The Commoditisation of Massive Data Analysis." In: *Radar*, November 19 (<http://strata.oreilly.com/2008/11/the-commoditisation-of-massive.html>).
- Hess, David J. (2005): "Technology-and-Product-Oriented Movements: Approximating Social Movement Studies and Science and Technology Studies." In: *Science, Technology & Human Values* 30/4, pp. 515–535.
- Iaconesi, Salvatore (2016): "[Air-L] La Cura: Possibilities?" Association of Internet Researchers Mailinglist, May 31 (<http://listserv.aoir.org/pipermail/air-l-aoir.org/2016-May/033196.html>).
- Irwin, Alan (1995): *Citizen Science: A Study of People, Expertise and Sustainable Development*, London-New York: Routledge.
- Jasanoff, Sheila (ed.) (2004): *States of Knowledge: The Co-Production of Science and the Social Order*, London-New York: Routledge.
- Jasper, James M. (1997): *The Art of Moral Protest: Culture, Biography, and Creativity in Social Movements*, Chicago: Chicago University Press.
- Jordan, Tim (2015): *Information Politics. Liberation and Exploitation in the Digital Society*, London: Pluto Press.
- Jordan, Tim/Taylor, Paul (2004): *Hacktivism and Cyberwars: Rebels with a Cause?*, London-New York: Routledge.
- Kitchin, Rob (2014): "Big Data, New Epistemologies and Paradigm Shifts." In: *Big Data & Society* April–June, pp. 1–12.

- Knorr Cetina, Karin/Reichmann, Werner (2015): "Epistemic Cultures." In: International Encyclopedia of the Social & Behavioral Sciences, 2nd ed., New York: Elsevier, pp. 873–80.
- Latour, Bruno (2004): Politics of Nature: How to Bring the Sciences into Democracy, Cambridge, MA: Harvard University Press.
- Latour, Bruno (2005): Reassembling the Social: An Introduction to Actor-Network Theory, Oxford: Oxford University Press.
- Latour, Bruno/Weibel, Peter (2005): Making Things Public: Atmospheres of Democracy, Cambridge, MA: The MIT Press.
- Levy, Steven (1984): Hackers: Heroes of the Computer Revolution, New York: Dell/Doubleday.
- Lupton, Deborah (2016): The Quantified Self, New York: Wiley.
- Lyon, David (2014): "Surveillance, Snowden, and Big Data: Capacities, Consequences, Critique." In: Big Data & Society 1/2, pp. 1–13.
- Lyon, David (2015): Surveillance After Snowden, Cambridge-Malden, MA: Polity Press.
- Marres, Noortje (2005): "No Issue, No Public: Democratic Deficits after the Displacement of Politics." PhD Diss., University of Amsterdam (<http://dare.uva.nl/record/165542>).
- Marres, Noortje (2011): "The Costs of Public Involvement: Everyday Devices of Carbon Accounting and the Materialisation of Participation." In: Economy and Society 40/4, pp. 510–33.
- Marvin, Carolyn (1990): When Old Technologies Were New: Thinking About Electric Communication in the Late Nineteenth Century, Oxford: Oxford University Press.
- Maxigas (2012): "Hacklabs and Hackerspaces – Tracing Two Genealogies." In: Journal of Peer Production 2.
- May, Timothy C. (1992): "The Crypto Anarchist Manifesto" (<http://www.activism.net/cypherpunk/crypto-anarchy.html>).
- McCaughey, Martha/Ayers, Michael D. (eds.) (2003): Cyberactivism: Online Activism in Theory and Practice, London-New York: Routledge.
- Melucci, Alberto (1996): Challenging Codes. Collective Action in the Information Age, Cambridge, UK: Cambridge University Press.
- Milan, Stefania (2013): Social Movements and Their Technologies: Wiring Social Change, Basingstoke: Palgrave Macmillan.
- Milan, Stefania (2014): "The Ethics of Social Movement Research." In: Donatella della Porta (ed.), Methodological Practices in Social Movement Research, Oxford: Oxford University Press, pp. 446–464.
- Milan, Stefania (2015a): "Hacktivism as a Radical Media Practice." In: Chris Atton (ed.), Routledge Companion to Alternative and Community Media, New York: Routledge, pp. 550–60.

- Milan, Stefania (2015b): "When Algorithms Shape Collective Action: Social Media and the Dynamics of Cloud Protesting." In: *Social Media and Society* 1/2, pp. 1–10.
- O'Brien, Paraic (2013): "Brown Moses: The British Blogger Tracking Syrian Arms." In: *Channel 4 News*, March 30 (<http://www.channel4.com/news/brown-moses-blog-syria-arms-weapons-croatia>).
- Ragin, Charles C. (1992): "Introduction: Cases of 'What Is a Case?'" In: Charles C. Ragin/Howard Saul Becker (eds.), *What Is a Case? Exploring the Foundations of Social Inquiry*, Cambridge, UK: Cambridge University Press, pp. 1–15.
- Renzi, Alessandra/Langlois, Ganaele (2015): "Data Activism." In: Greg Elmer/Ganaele Langlois/Joanna Redden (eds.), *Compromised Data: From Social Media to Big Data*, London: Bloomsbury, pp. 202–25.
- Rochon, Thomas R. (1998): *Culture Moves. Ideas, Activism, and Changing Values*, Princeton: Princeton University Press.
- Russon Gilman, Hollie (2016): *Participatory Budgeting and Civic Tech: The Revival of Citizen Engagement*, Washington, DC: Georgetown University Press.
- Schrock, Andrew R. (2016): "Civic Hacking as Data Activism and Advocacy: A History from Publicity to Open Government Data." In: *New Media & Society* 18/4, pp. 581–99.
- Sifry, Micah L. (2011): *Wikileaks and the Age of Transparency*, New York: OR Books.
- Sismondo, Sergio (2003): *An Introduction to Science and Technology Studies*, Malden, MA: Wiley-Blackwell.
- Striphas, Ted (2015): "Algorithmic Culture." In: *European Journal of Cultural Studies* 18/4–5, pp. 395–412.
- Tactical Tech Collective (2013): *Visualising Information for Advocacy*, Bangalore: Tactical Tech Collective.
- Tactical Tech Collective (2016): "Decoding Data." In: *Exposing The Invisible*, October 18 (<https://exposingtheinvisible.org/guides/decoding-data/>).
- Taylor, Verta/van Dyke, Nella (2004): "Get Up, Stand Up: Tactical Repertoires of Social Movements." In: David A. Snow/Sarah A. Soule/Hanspeter Kriesi (eds.), *The Blackwell Companion to Social Movements*, Malden, MA and Oxford: Blackwell, pp. 262–93.
- The Guardian Project (n.d.): "CameraV: Secure Verifiable Photo & Video Camera" (<https://guardianproject.info/apps/camerav>).
- Van der Velden, Lonneke (2015): "Forensic Devices for Activism: Metadata Tracking and Public Proof." In: *Big Data & Society* 2/2.
- Van Dijck, José (2013): *The Culture of Connectivity: A Critical History of Social Media*, Oxford: Oxford University Press.
- Winner, Langdon (1980): "Do Artifacts Have Politics?" In: *Daedalus* 109/1, pp. 121–36.
- Zuboff, Shoshana (2015): "Big Other: Surveillance Capitalism and the Prospects of an Information Civilisation." In: *Journal of Information Technology* 30/1, pp. 75–89.