“We are living in a time in which we are surrounded by data. Governments around the world are opening up their data vaults, allowing anybody access to it,” explained Simon Rogers, editor of the Guardian Data blog. We indeed live in a time of data abundance. Vast data sets are continuously generated and automatically stored by a variety of technologies such as radio-frequency identification readers. As we move in cyberspace, our activities leave behind digital traces of our doings, in a myriad of software logs and communications metadata collected by service providers. The ability of generating and making sense of ever-larger quantities of data has prompted observers to speak of a new breakthrough phase in human history, which Hellerstein (2008) termed the “industrial revolution of data”.

Not only do big data pose new challenges; they also offer new opportunities. Says again Rogers: “this is a big deal because it means we have access to data as we have never had in the past. This gives us stories, and new ways of looking at the world” (in Bonechi, 2012). Citizens are becoming increasingly aware of the potential of data for social change. This awareness gives rise to new social practices rooted in technology and data, that we have termed “data activism”.

Data activism indicates social practices that take a critical approach to big data. Examples include the collective mapping and geo-referencing of the messages of victims of natural disasters in order to facilitate disaster relief operations, or the elaboration of open government data for advocacy and campaigning. But data activism also embraces tactics of resistance to massive data collection by private companies and governments, such as the encryption of private communication, or obfuscation tactics that put sand into the data collection machine.

Data activism is a theoretical construct grounded on empirical observations. We propose it as a heuristic tool useful to explore how people engage politically with big data and massive data collection. In this theoretical paper, we explore the potential use of data, data-based narratives, and data crunching software as tools for activism and social change. In what follows, we present the socio-political and economic context in which data activism emerges, and offer a preliminary definition of data activism. We then ground the concept on a multidisciplinary literature review that places it in context and explores some of its facets. Finally, we offer a revised definition of data activism, and provide a tentative agenda for the study of the phenomenon.

Setting the Scene: Informational States and Computational Politics

Since 2012, the non-profit media organization WikiLeaks has released millions of classified documents, including US diplomatic cables and field reports from the Afghan and Iraqi wars (c.f. Brevini et al., 2013). Similarly, the top-secret information that whistle-blower Edwards J. Snowden, a former US intelligence contractor, offered to media outlets, has exposed blanket data collection strategies by Western national security agencies. More often than not, the disclosure of critical information spurred public outrage and forced authorities to react. The Snowden revelations, for example, mobilized heads of state and were discussed at the United Nations’ General Assembly. The WikiLeaks cable leaks, too, had social consequences, to the point that they are considered the catalyst of, for instance, the 2010-2011 Tunisian uprising. Some claim that “We might also count Tunisia as the first time that WikiLeaks pushed people over the brink” (Dickinson, 2011).

As the WikiLeaks case shows, the growth in magnitude and complexity of information has changed the nature of the nation state as well as its relation with citizens. Governments “deliberately, explicitly, and consistently control information creation, processing, flows, and use to exercise power” (Braman, 2009). The state has become “informational”: in other words, it is characterised by “shifts in the nature of power and its exercise via information policy” (Braman, 2009). But there is more. According to Tufekci, digital technologies “have pushed people over the brink” (Tufekci, 2014). The expansion of such “computational politics” is intimately linked to “the rise of big data, the shift away from demographics to individualised targeting, the opacity and power of computational modelling, the use of persuasive behavioural science, digital media ena-
Defining Data Activism

Data activism emerges in the fringes of society, in the realm typically associated with grassroots activism and civic engagement. However, it is rapidly evolving from a peripheral, elitist form of activism to a diffused and democratized form of protest, enabled by new technologies. (…) Cyberspace has brought us the means to engage in cyber espionage have expanded dramatically because of the shift to networked infrastructures and social networking habits. (…) The internet allows us to do-it-yourself signals intelligence." In such a scenario, there emerge grassroots practices that bring progressive citizens to the core of the “data revolution”.

Data activism signals a change in perspective and attitude towards massive data collection emerging at the core of the civil society realm. There are evident links with the long tradition of alternative, independent and community media that bring citizens at the forefront of media production.

From (Alternative) Media Studies: Citizen Media and Empowerment

Data activism signals a change in perspective and attitude towards massive data collection emerging at the core of the civil society realm. There are evident links with the long tradition of alternative, independent and community media that bring citizens at the forefront of media production.

Over the years, media scholars have proposed different labels to describe non-commercial grassroots media: radical media (Dowling, 2001), citizens’ media (Rodriguez, 2001), alternative media (Atton, 2001), tactical media (Garcia and Lovink, 1997), autonomous media (Langlois and Dubois, 2005) and (counting). Hackett and Carroll (2006) referred to “oppositional communication practices”, emanating from the lifeworld and centred on lifeworld change, and seeking to cultivate alternative public spheres. Alternative media scholar John Downing defined radical media as “media, generally small-scale and in many different forms, that express an alternative vision to hegemonic politics, priorities, and perspectives". By nature, these media “break somebody’s
rules, although rarely all of them in every as-
pect" (Downing, 2001: vi-x). However, it is the concept of citizens’ media that offers the most powerful insights for the study of data activ-

What makes the notion of citizen’s media so appropriate for the subject matter is its focus on empowerment and the politics of daily life. Within the communication sphere, empower-

From Journalism Studies: Journalism Meets Data (And Social Change)

In spring 2010 WikiLeaks released a video en-
titled Collateral Murder, featuring a US Army helicopter shooting Iraqi civilians. The public debate spurred by the release forced US authori-
ties to react. Over the last few years, Wiki-
Leaks’ revelations have offered an unprece-
dented amount of raw material for investiga-
tive journalism, which resulted in the partner-
ship with five prominent media organisations, including El Pais and The New Yorks Times. Dan Gillmor (2005), one of the most prominent representatives of a journalism that takes si-
des, declared: “Maybe it’s time to say a fond farewell to an old canon of journalism: objec-
tivity. But it will never be time to kiss off the values and principles that undergird the idea”.

Looking at WikiLeaks and at the recent tra-
jectory of the journalism field offers some insights into the emergence of data activism. Three genres, in particular, directly speak to data activism practices, namely investigative, advocacy, and citizen journalism. Investigati-

The advent of big data has signalled an evo-
lution in journalism too, exemplified by “data journalism”. Data journalism makes use of advanced social science and computational methods to find stories in complex data sets. It represents an evolution from precision jour-

Table: Data Journalism

Data journalism calls for the development of specific skills in programming and advanced statistical methods, that are not typically part of the skillset of old-school journalists. Thus, the journalists’ engagement with data prompts unprecedented alliances, which are of paramount relevance for observers of data activism. For example, the international grass-

innovative and better journalism. He quotes WiKiLeaks and the whistle-blower Edward J. Snowden as examples of how to put into prac-
tice a prototype of critical and free press for the common good. In his view, hackers and whistle-blowers retrieve the ideals of the old muck-racking investigative journalism: “The trip that hacktivists have initiated is a return trip: the journalism that is coming is returning to its roots”. In order to make it possible, the press must “recover their objectives” that give “democratic legitimacy” to this profession. Si-

124 125
forms that ease the task of navigating and making sense of data facilitates the connections and collaboration between the professionals of information and common citizens with interest for their social reality. For instance, the Open Knowledge Foundation has developed the open-source data portal CKAN, used by laymen to explore datasets made available as “open data” by their governments; the OpenSensing online platform allows non-experts to explore over 13 million government financial transactions from 66 countries.

From Social Movement Studies: Technical Identities and Informal Communities

September 2011. Anti-austerity protesters try to occupy the New York Stock Exchange on Wall Street, but are turned away by the police. They withdraw to the nearby Zuccotti Park, and start an encampment that will set in motion the so-called “Occupy Wall Street” world-wide protests. The Occupy slogan (“We are the 99%”) calls attention to the disempowered 99 percent of the world population as it is disenfranchised by a privileged minority managing resources like access to credit. “It was really interesting to see how interested in data that movement has been. The 99% versus the 1% is a data query,” argued Simon Rogers, editor of the Guardian Data blog. Occupy protesters “have managed to use data to show what is happening around the world” (Bonacci, 2012).

As the Occupy protests show, the organized civil society increasingly acknowledges the evocative power of numbers and statistics for civic engagement, protest, and advocacy. But how do social movements use data to foster social change? Data activists position themselves as interpreters of data, acting as facilitators in the contemporary data-rich public sphere.

They take advantage of the decentralised peer production and distributed human capital of what Benkler (2006) called the “networked public sphere”: individuals and groups who share their expertise in software development and data analysis in order create and make available to others data-crunching tools and platforms.

While data activism is still relatively unknown to scholarship, the social forces behind it are not new. Activists have appropriated and repurposed communications technology at least since the 1950s, both for self-expression and social change: it is what has been termed “media activism”, that is to say the creative and tactical use of media for advancing social change (Milan and Padovani, 2014). Furthermore, several social movement currents and sub-cultures influence today’s data activism: for example, the hacker culture of the 1970s and beyond (Levy, 1984); the US media reform activists advocating technical engagement as a way of confronting elite expertise (Dunbar-Hester, 2014); the radical technology activism of the 1990s (Milan, 2013a); the digital rights activists as they engage in technological resistance (Postigo, 2012); the do-it-yourself attitude of hackerspaces, hacklabs and makerspaces (Maxigas, 2014); the technology- and product-oriented movements such as the open-source culture (Hess, 2005). Data activists have borrowed from these sub-cultures the emphasis on hands-on activism, and the notions of knowledge sharing, peer collaboration, access to information, code tinkering, and world improvement through technical fixes. Surveying closely this scholarship will provide generous hints to interpret present-day data activists and their impact on society as a whole.

In Social Movements and their Technologies, Milan (2013b) explored the interaction between social movements and their “liberated technologies”, focusing on the upsurge of radical internet projects (“emancipatory communication practices”) as a political subject. She explained how radical tech groups, which create autonomous, under-the-radar alternatives to ordinary communication systems and networks, contribute to change how people communicate in daily life. Similarly, data activists engage in practices that will, in the long run, change the ways activists campaign and media (including citizens’ media) disseminate information. Both contribute peculiar progressive values and mechanisms to the society they live in, as their value systems and organisational models emphasize collective, open and decentralised creation and decision-making.

Dunbar-Hester (2012) has studied low-power radio broadcasting in the US, focusing in particular on those groups promoting critical engagement with the technology (for example, building radio stations that transmit the “people’s music”). She points to the emergence of a “technical identity” rooted in the hands-on attitude activists develop through their technology-mediated activism. Practitioners use this technical identity to actively “mark the boundaries between their group and others in the terrain of media democracy work” (Dunbar-Hester, 2012, p. 149).

There is something to be learnt for the study of data activism from the literature on scientific and intellectual movements that has emerged from science ad technology studies, and has intersected social movement studies. Activists of technology-and product-oriented movements such as open-source software activism are inherently political, as their hands-on practices seek ultimately to alter power distribution (Latour, 1987). They engage in the “development of new or alternative forms of material culture” (Hess, 2005). Occasionally, they might establish connections with private sector firms, which they see as potential allies in the promotion of alternative technology. Similarly, data activists might partner up with small, alternative software companies to market their data-crunching products.

Finally, there is much to learn from recent studies on technology-mediated collective action, whereby activists use electronic communication technologies to communicate, lobby, build communities, organise, and coordinate action, but also from radical forms of internet activism such as cyberactivism and hacktivism. There is a burgeoning of literature on the subject matter, so here we highlight only those contributions that appear to be more useful to understand data activism. Bennett and Segerberg (2012), for example, elaborated on the notion of “connective action”, where the mediation of digital technologies promote personalised interactions, emphasizing the links between individual users rather than between organisations and movements. We have observed this dynamic at play also within more radical forms of internet activism (c.f. Milan and Hints, 2013), as well as data activism.

Further similarities exist between present-day data activists and radical internet activism, that is to say those groups that exploit the technical infrastructure of the web for social change (Milan, 2013b). For both, technology is simultaneously the means to provoke change in society and a site of struggle in its own right, in recognition of the politics and power relations embedded in software (Milan, 2013a). Likewise, this type of activism requires a high degree of specialisation and technical expertise, which is then typically rewarded with individual reputation and a central role within one’s group (Hintz and Milan, 2009; c.f. Coleman, 2013).
From International Relations: Transnational Advocacy Networks as Information Agents

In 2013, the United Nations’ Department of Economics and Social Affairs Under-Secretary-General Wu Hongbo called for more data in development. “Statistics is shaping our understanding of the world,” he said addressing the UN Statistical Commission (UN DESA, 2014). This renewed interest by multilateral organizations and governments alike for data and statistical evidence is likely to alter the advocacy strategies of a myriad of civil society organizations, who will move away from moral arguments into statistical evidence in advocating for their causes.

Whereas much is known about the evolution of transnational activism in the internet era (see, for example, della Porta and Tarrow, 2005), to date scholars have not addressed the evolution of transnational civil society structures and dynamics in relation to the dramatic change in the global information flows. International relations scholars Keck and Sikkink offered a conceptual tool, namely the notion of transnational advocacy networks (TANs), which is potentially useful in interpreting data-prodded dynamics within the global civil society. These are value-rich networks of social actors characterised by the production, exchange, and strategic use of information. They seek to influence policymaking, but also to “transform the terms and nature of the debate” (Keck and Sikkink, 2008, p. 2). TANs include a vast assemblage of groups operating across borders and beyond the reach of governments. They tend to be excluded from traditional political mechanisms as they intersect community. On the contrary, when citizen’s media encounters activism and advocacy, we witness the emergence of media activism. In turn, when activism and advocacy join journalistic practices and values we observe the emergence of advocacy journalism. When big data intersects journalism, and investigative journalism in particular, we have data journalism. It is, however, at the intersection of more quadrants that the most noteworthy trends come into being. When advocacy journalism and investigative journalism cross paths with available data and data-analysis software, we have the sub-field of “advocacy data journalism”, which combines the traditional values of journalism with crunching data and a social change ethos. At the juncture of the four circles, finally, we find “pro-active data activism”.

What notions and mechanisms are apt to study such a complex and inherently interdisciplinary field of action? From journalism studies, we have learnt that there are values of objectivity, fairness and accuracy that pro-active data activists, too, embody. From (alternative) media studies, we derive the focus on empowerment mechanisms as they intersect communities, and the notion of activism as an every-day critical engagement with the surrounding social reality. International relations offer hints on how to integrate transnational power dynamics into the analysis of a phenomenon that is local and transnational at the same time. The emerging data activism networks take the form of trans-border de-localised communities, while operating as transnational advocacy networks. Similar to other instances of radical internet activism, these “innew forms of networked action and informal collaboration

Revisiting Pro-Active Data Activism: A Research Agenda

Pro-active data activism identifies an emerging field of action that combines communicative practices and the social elements of collective organizing with information at its outermost complexity: “big data”. It is characterised by a multiplicity of arrangements and actors, both collective and individual, as well as a wealth of different tactics of engagement with data. Pro-active data activism tactics range from technology development projects and platforms for the manipulation of data and the visualization of data patterns for campaigning and advocacy. In light of the literature review, we now revise and expand our notion of data activism.

Data activism emerges at the intersection of several fields of human action, including both communicative process (activism and advocacy, and the subaltern counterpublics interested in empowerment through media and technology) and information-related professions (data analysis and journalistic investigation). Figure 1 shows the dynamic interactions between these different fields.

![Diagram of data activism](https://example.com/diagram.png)

Figure 1. Data activism and neighbouring fields of action.

The figure can be explored clock-wise, starting from the big data quadrant, which sets in motion the empirical phenomenon of data activism. When big data, and data analysis practices and tools, meet citizens’ media and other grassroots experiences that put empowerment though active engagement with media at the centre, there emerge instances of “civic hacking”. Civic hackers are technologists and open-data activists who engage with datasets to address challenges relevant to their community. On the contrary, when citizen’s media encounters activism and advocacy, we witness the emergence of media activism. In turn, when activism and advocacy join journalistic practices and values we observe the emergence of advocacy journalism. When big data intersects journalism, and investigative journalism in particular, we have data journalism.

What notions and mechanisms are apt to study such a complex and inherently interdisciplinary field of action? From journalism studies, we have learnt that there are values of objectivity, fairness and accuracy that pro-active data activists, too, embody. From (alternative) media studies, we derive the focus on empowerment mechanisms as they intersect communities, and the notion of activism as an every-day critical engagement with the surrounding social reality. International relations offer hints on how to integrate transnational power dynamics into the analysis of a phenomenon that is local and transnational at the same time. The emerging data activism networks take the form of trans-border de-localised communities, while operating as transnational advocacy networks. Similar to other instances of radical internet activism, these “innew forms of networked action and informal collaboration
are challenging traditional notions of civil society” (Hintz and Milan, 2009), and are radically different from traditionally organized collective action. Nonetheless, we suggest approaching data activism using the tools of social movement studies.

Social movement studies help us to understand data activism as a new form of grassroots engagement with technology, one that entails direct action, cultural forms of resistance, and coding. Similar to other activism tactics, these strategies provide activists with a “moral voice”, which gives them the opportunity to articulate and affirm their principles and preferences (Jasper, 1997). A few key concepts derived from the literature on social movements are particularly useful to study data activism: for example, one can study data activists’ collective identity, that is to say the “interactive and shared definition” by individuals who recognise that they share certain orientations and, on that basis, decide to act together (Melucci, 1996), asking how is a collective identity created on the basis of a technical identity. We can also investigate the organizational forms typical of data activism: what groupings are more conducive to data activism activities and practices? What internal mechanisms regulate them? Finally, we can study the relationship between pro-active data activists and institutions and social norms, including national governments and multilateral organizations.

Ultimately, pro-active data activism can be considered a new, advanced form of citizens’ media, one that has a critical approach to big data at its core. Similar to citizens’ media, pro-active data activism involves a politics of the quotidian, as it alters the everyday relationship between citizens and automatized data collection. As such, it brings back into the data collection machine the fundamental elements of agency and politics. Similar to other social movements, pro-active data activists work towards long-term norm change: they struggle for “the social control of the main cultural patterns (...) through which our relationships with the environment are normatively organised” (Touraine, 1985). Thus, in the long run, data activism is likely to change the way citizens approach computational politics and the informational state, as well as the way we see and practice social change.

References


Maxigas, 2014. Internet governance is a trap.


