CITIZENS' MEDIA MEETS BIG DATA: THE EMERGENCE OF DATA ACTIVISM

Stefania Milan
Miren Gutiérrez

"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 aerial sensors, ubiquitous mobile devices, and 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".

Big data embrace, for instance, the various databases generated by governmental agencies in their functions, and sometimes released as "open data" (i.e., data that citizens can use, reuse and redistribute). Big data include also the extraordinary amount of video and audio files, texts, links and tags that result from online distribution and archiving, and the information generated by human interactions in social networking platforms. Further, big data refers also to the indexing processes of web activities, and the communication metadata resulting from mass-interception and government snooping. But they can result from the process of "datafication", that is to say the "ability to render into data many aspects of the world that have never been quantified before" (Cukier and Mayer-Schoenberger, 2013, pp. 28–40), such as friendships in the form of "likes".

The big data technology and services market is expected to grow from the USD 6 billion of 2011 to USD 23.8 billion in 2016, for a "annual growth rate of over 30 percent, which equals about seven times the information and communication technology market" (Vesset, 2012). Not surprisingly, analyst Abhishek Mehta (2011) sees data is a key "raw material for a variety of socio-economic business systems". Although analysts in fields as diverse as meteorology, finance and genomics have always faced the challenge of making sense of large data sets, the unparalleled scale of the information generated today poses brand-new challenges to a variety of professions dealing with information. Journalism, as one of the key professions specialized in making information visible and accessible to large audiences, is at the forefront of this "data revolution".

Not only do big data pose new challenges; they also open up opportunities. Says again Rogers: "this is a big deal because it means we have access to data as we have would never have 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 given rise to a new combination of big data and computational practices which allow for massive, latent data collection and sophisticated computational modelling, increasing the capacity of those with resources and access to use these tools to carry out highly effective, opaque and unaccountable campaigns of persuasion and social engineering in political, civic and commercial spheres" (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-

bling dynamic real-time experimentation, and the growth of new power brokers who own the data or social media environments" (Braman, 2009).

But the informational state and its computational politics require "corporate collaboration" (MacKinnon, 2012) to implement laws and regulations: as the Snowden leaks made evident, governments increasingly rely on "private sector entities as regulatory agents, turning private centres of power to state purposes" (Braman, 2009). Beyond contributing to monitoring users, corporate actors are behind an increasingly "appliancised" internet, where "tethered" devices (which can be modified only by the manufacturer) end up jeopardizing the generative qualities of the network, that is to say its capacity of finding ways out of government control at the end-user level (Zittrain, 2008), de facto reducing user freedom. Furthermore, disempowered people end up paying the highest price, as often online surveillance serves to discriminate against "members of chronically underserved communities" (Gangadharan, 2012)digital inclusion policies designed to introduce poor people, communities of color, indigenous, and migrants (collectively, "chronically underserved communities" or "the underserved".

Access to data is power, thus data and information has become an increasingly important currency in contemporary politics, as shown by the WikiLeaks and Snowden cases. In recognition of the crucial role of access to information in contemporary society, hackers and hacktivists (the term is a portmanteau of "hackers" and "activists") increasingly take action to counter-act the power of governments in shaping the internet and limiting freedom of expression. Hackers can be seen as the "immune system" of the Internet: by exposing vulnerabilities, they push the internet to become stronger and healthier, wielding their power to create a better world (Elazari, 2014). The very same technology that allowed governments and corporations to amass and exploit digital data about private citizens now offers citizens the opportunity to target governments" and companies" data and computer systems. According to Deiber (2010), WikiLeaks was just "a symptom of a much larger trend (...) the means to engage in cyber espionage have expanded dramatically because of the shift to networked infrastructures and social networking habits. (...) Cyberspace has brought us the world of do-it-yourself signals intelligence."
In such a scenario, there emerge grassroots practices that bring progressive citizens to the core of the "data revolution".

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 one, whereby also non-skilled users take part in the game. It involves a series of practices at the intersection of the social and the technological dimension of human action, with two aims: either resisting massive data collection, or actively pursuing the exploitation of available data for social change.

For analytical purposes, we have identified two sub-fields of data activism: re-active and pro-active. Re-active data activism comprises the practices of resistance to the threats to civil and human rights that derive from corporate and government privacy intrusion. Pro-active data activism embraces those individuals and civil society organizations that take advantage of the possibility for social change and civic

engagement offered by big data. Re-active and pro-active data activism constitute two facets of the same phenomenon, which has data and information at its core. They are enabled by software to manipulate data or to shield one"s online interactions from intrusion and automatized collection. In this article, however, we focus on the pro-active side of data activism, and explores its connections with existing journalism and media practices.

The notion of data activism represents a conceptual innovation at the crossroads of a sociological process (organising collectively in order to take action), a cognitive activity (making sense of complex information), and a sociotechnical practice (the technology, namely the software to manipulate the data, is central to data activism). Studying such an empirical phenomenon, hence, calls for an interdisciplinary approach able to take into account its multidimensional nature. In the next section, we survey contributions from a variety of disciplinary fields, namely: media studies, social movement scholarship, journalism studies, and international relations, in view of grounding the notion of data activism on existing literature.

A Kaleidoscopic Perspective on Data Activism Seen From Different Disciplines

Data activism is an emerging empirical phenomenon for which we need to develop a new vocabulary of interdisciplinary concepts and mechanisms. How do we study a social phenomenon that is rooted in communications technology (and information), and is positioned between the social and communicative dimensions of human action? Here we offer an overview of useful concepts as they emerge from various disciplines at the interplay of social sciences and humanities. The concepts we survey serve the purpose of illuminating different aspects of data activism, and will allow us to refine our notion in light of the literature.[1]

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 (Downing, 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

1 Science and technology studies offer interesting points of entry to study contemporary data activism. By providing a way to think about socio-technical objects and processes, it allows us to rethink the interplay between technology (and data) and society, and to explore simultaneously the technological and social dimension of data activism. Actor-Network Theory (Latour, 1987), and the notions of socio-technical artifacts, object conflicts, and boundary work are particularly useful. However, although it is part of our work, discussing the science and technology perspective on data activism goes beyond the scope of this article.

rules, although rarely all of them in every aspect" (Downing, 2001: v-ix). However, it is the concept of citizens" media that offers the most powerful insights for the study of data activism.

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, empowerment is the process through which individuals and groups, by taking active part in the actions that reshape their communicative processes, exercise control over their communication resources and messages. It is in this active exercise of control over technologies that people, including non-experts, are empowered (Milan, 2013a). Similarly, Rodriguez (2001) sees citizens" media as a space for people to enact their democratic agency beyond traditional means like voting. Furthermore, she argued that "these practices and strategies of resistance constitute the politics of the quotidian." They expand and multiply spaces for political action, which is not confined to institutional spaces but embedded in social life and daily peer interactions (p. 21).

Interestingly, Rodriguez draws on radical democracy and feminist theories to define citizens" media. In her words, citizens" media imply that "a collectivity is enacting its citizenship by actively intervening and transforming the established mediascape; second, that these media are contesting social codes, legitimized identities, and institutionalized social relations; and third, that these communication practices are empowering the community involved, to the point where these transformation and changes are possible" (p. 20). All of the above apply to data activism as well, albeit to different degrees. If the contestation ethos and the challenge to established identities and social relations is common to the two approaches, data activism differes from citizens" media in that is anticipates also a variety of individual practices. There is within the field of data activism a novel tension between the individual and the collective dimension of organized collective action, which risks to sideline the community terms of reference so central in the definition of alternative, community and citizens" media.

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

In spring 2010 WikiLeaks released a video entitled Collateral Murder, featuring a US Army helicopter shooting Iraqi civilians. The public debate spurred by the release forced US authorities to react. Over the last few years, WikiLeaks" revelations have offered an unprecedented amount of raw material for investigative journalism, which resulted in the partnership 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 sides, declared: "Maybe it"s time to say a fond farewell to an old canon of journalism: objectivity. But it will never be time to kiss off the values and principles that undergird the idea".

Looking at WikiLeaks and at the recent trajectory 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. Investigative journalism is believed to be journalism at its best. It entails objective reporting on subjects that are of interest to the citizenry, when "others are attempting to hide these matters from the public" (Northmore, 2001). On the contrary, advocacy journalism deliberately

adopts a partisan perspective in reporting. It is fact-based, but it openly takes sides (Careless, 2000), as these journalists feel that the public interest is better served by media diversity and transparency.

Since the 1990s, internet has revolutionized journalism practices, and has "busted open the system of gates and gatekeepers" typical of mainstream media (Rosen, 2006). As a result, citizen journalism emerged, opening up the newsroom to non-professionals. According to Rosen (2008), we see citizen journalism in action "[w]hen the people formerly known as the audience employ the press tools they have in their possession to inform one another." Curiously, citizen journalism has acquired popularity also within mainstream media, as a tool for involving the audience in news-making as "produsers" (Bruns, 2008). It has limited resemblance with what others have called alternative journalism (Atton and Hamilton, 2009), although it differs in that it seeks to safeguard some ambitions of objectivity typical of traditional journalism. The familiarity with the notion of citizens" media described above, however, is very limited, as the latter deliberately positions itself in the realm of subaltern counterpublics (Fraser, 1992), rather than seeking integration with the mainstream.

The advent of big data has signalled an evolution 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 journalism, that is to say journalism that uses advanced research methods to bring a level of rigor to journalistic work beyond anecdotal evidence (Meyer, 2002). In El cuarto poder en red, Víctor Sampedro (2014) argues that the new journalistic practices rooted in collaborative technology and data are the start of an innovative and better journalism. He quotes WikiLeaks and the whistle-blower Edward J. Snowden as examples of how to put into practice 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. Similar claims belong to data activists as well.

Data journalism offered a breath of fresh air to a profession in deep crisis, facing a generalised and continued drop in circulation, readership, audiences, finances and even credibility. If "journalism is under siege" (Gray and Chambers, 2012), the advent of data journalism becomes particularly significant, as it signals also an evolution in the (social) function of journalists: "Using data the job of journalists shifts its main focus from being the first ones to report to being the ones telling us what a certain development might actually mean" (Gray and Chambers, 2012). And, as journalists are using data to move further, they progressively move from observation into analysis and denunciation, raising awareness and recommendations. This is the case of "advocacy 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 grassroots network Hacks/Hackers connects journalists ("hacks") and technologists ("hackers"), facilitating skill exchange between the two groups (Hacks/Hackers n.d.). A variety of plat-

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 OpenSpending 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" worldwide 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" (Bonechi, 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 their own radio transmitter). She pointed 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 activisms 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, emphasising the links between individual users rather then between organisations and movements. We have observed this dynamic at play also within more radical forms of internet activism (c.f. Milan and Hintz, 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 are particularly useful to map emerging transnational networks of data activists precisely because they have information at their core, in an era in which big data represent the newest form that information has taken.

Data activists mobilizing at the transnational level represent the quintessential TAN. This is easier to understand by looking at a concrete example of transnational data activism in action. "Price of Oil" (http://priceofoil.org) is a website of the organizational Oil Change International. It seeks to expose "the true costs of fossil fuels", including the environmental and social consequences of fuel extraction and consumption. Its transnational campaigns take the shape of TANs, and are grounded on data as currency of activism. Looking at TANs of this kind will allow us to explore how the transnational civil society interact with big data, while introducing in the research a new form of critical engagement with data, that of advocacy, which is grounded on data analysis but at the same time goes beyond it by engaging directly with institutions such as companies and governments. In addition, looking at data activist TANs will enable us to observe closely how the advent of big data alters the power dynamics within the transnational civil society. Finally, literature on transnational networks of radical internet activism offer hints for the analysis of data activist TANs: although they usually embody highly principled preferences associated with a call for institutional or policy change, they tend to be excluded from traditional policy arenas (Hintz and Milan, 2009).

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.

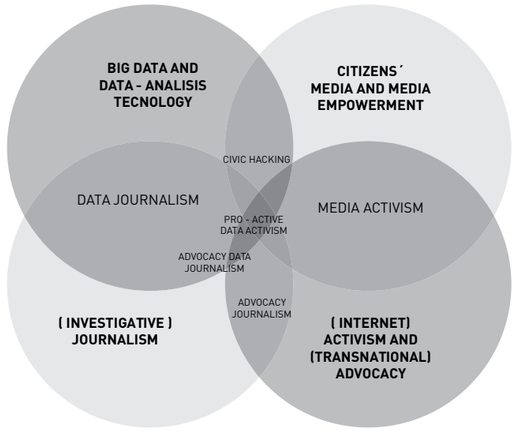

Figure 1. Data activism and neighbouring fields of action.

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.

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 "[n]ew 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

Atton, C., 2001. Alternative media. Sage, Thousands Oaks, CA.

Atton, C., Hamilton, J.F., 2009. Alternative Journalism. Sage, Thousands Oaks, CA.

Benkler, Y., 2006. The wealth of networks: How social production transforms markets and freedom. Yale University Press, New Haven and London.

Bennett, L.W., Segerberg, A., 2012. The logic of connective action: Digital media and the personalization of contentious politics. Information, Communication & Society 15, 739–768.

Bonechi, M., 2012. Interview with Simon Rogers, editor Guardian Data Blog, London. London.

Braman, S., 2009. Change of state: Information, policy, and power. MIT Press.

Brevini, B., Hintz, A., McCurdy, P., 2013. Beyond WikiLeaks: implications for the future of communications, journalism and society. Palgrave Macmillan, Basingstoke, UK.

Bruns, A., 2008. Blogs, Wikipedia, Second Life, and Beyond: From Production to Produsage. Peter Lang, New York.

Careless, S., 2000. Advocacy Journalism. The Interim.

Coleman, G., 2013. Coding Freedom: The Ethics and Aesthetics of Hacking. Princeton University Press, Princeton and Oxford.

Cukier, K., Mayer-Schoenberger, V., 2013. The Rise of Big Data: How it"s Changing the Way We Think about the World. Foreign Aff. 92, 28–40.

Deibert, R., 2010. After WikiLeaks, a New Era. New York Times.

Della Porta, D., Tarrow, S. (Eds.), 2005. Transnational protest and global activism. Rowman & Littlefield, Lanham, MA.

Dickinson, E., 2011. The first Wikileaks revolution? Foreign Policy.

Downing, J.D.H., 2001. Radical Media: Rebellious communication and social movements. Sage, Thousands Oaks, CA.

Dunbar-Hester, C., 2012. Soldering Toward Media Democracy Technical Practice as Symbolic Value in Radio Activism. Journal of Communication Inquiry 36, 149–169.

Dunbar-Hester, C., 2014. Producing "Participation"? The Pleasures and Perils of Technical Engagement in Radio Activism. Public Culture 26, 25–50.

Elazari, K., 2014. Hackers: the Internet"s immune system, TEDtalks.

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. MIT Press, Cambridge, MA, pp. 109–42.

Gangadharan, S., 2012. Digital inclusion and data profiling. First Monday 17.

Garcia, D., Lovink, G., 1997. The ABC of Tactical Media.

Gillmor, D., 2005. The End of Objectivity (Version 0.91).

Hackett, R.A., Carroll, W.K., 2006. Remaking media: the struggle to democratize public communication. Routledge, New York; London.

Hellerstein, J., 2008. The Commoditization of Massive Data Analysis - Data [WWW Document]. URL http://strata.oreilly.com/2008/11/the-commoditization-of-massive.html (accessed 2.26.14).

Hess, D.J., 2005. Technology-and product-oriented movements: Approximating social movement studies and science and technology studies. Science, Technology & Human Values 30, 515–535.

Hintz, A., Milan, S., 2009. At the margins of Internet governance: grassroots tech groups and communication policy. International Journal of Media & Cultural Politics 5, 23–38.

Jasper, J., 1997. The Art of Moral Protest: culture, Biography, and Creativity in Social Movements. Chicago University Press, Chicago.

Keck, M., Sikkink, K., 2008. Activists beyond borders. Advocacy Networks in International Policy. Cornell University Press, London.

Langlois, A., Dubois, F. (Eds.), 2005. Autonomous Media: Activating Resistance & Dissent. Cumulus Press, Montréal.

Latour, B., 1987. Science in Action: How to Follow Scientists and Engineers through Society. Harvard University Press, Cambridge, MA.

Levy, S., 1984. Hackers: Heroes of the Computer Revolution. Dell/Doubleday, New York.

MacKinnon, R., 2012. Consent of the Networked. The Worldwide Struggle for Internet Freedom. Basic Books, New York.

Maxigas, 2014. Internet governance is a trap.

Mehta, A., 2011. Big Data: Powering the Next Industrial Revolution.

Melucci, A., 1996. Challenging Codes. Collective Action in the Information Age. Cambridge University Press, Cambridge, UK.

Meyer, P., 2002. Precision Journalism. A Reporter"s Introduction to Social Science Methods. Rowman & Littlefield, Lanham, MA.

Milan, S., 2013a. Social Movements and Their Technologies: Wiring Social Change. Palgrave Macmillan.

Milan, S., 2013b. WikiLeaks, Anonymous, and the exercise of individuality: Protesting in the cloud, in: Brevini, B., Hintz, A., McCurdy, P. (Eds.), Beyond WikiLeaks: Implications for the Future of Communications, Journalism and Society. Palgrave Macmillan, Basingstoke, UK, pp. 191–208.

Milan, S., Hintz, A., 2013. Networked Collective Action and the Institutionalized Policy Debate: Bringing Cyberactivism to the Policy Arena? Policy & Internet 5, 7–26.

Milan, S., Padovani, C., 2014. Communication Rights and Media Justice Between Political and Discursive Opportunities: A Historical Perspective, in: Padovani, C., Calabrese, A. (Eds.), Communication Rights and Social Justice: Historical Accounts of Transnational Mobilizations. Palgrave Macmillan, Basingstoke, UK, p. forthcoming.

Northmore, D., 2001. Lifting the Lid: A Guide to Investigative Research. Continuum International Publishing, New York.

Postigo, H., 2012. The digital rights movement the role of technology in subverting digital copyright. The MIT Press, Cambridge, Mass.

Rodriguez, C., 2001. Fissures in the mediascape. An international study of citizens" media. Hampton Press, Cresskill, NJ.

Rosen, J., 2006. TEN INTERNET YEARS Web Users Open the Gates. Washington Post.

Sampedro, V., 2014. El cuarto poder en red. Por un periodismo (de código) libre. Icaria, Barcelona.

Touraine, A., 1985. An Introduction to the Study of Social Movements. Social Research 52, 749–787.

Tufekci, Z., 2014. Engeneering the Public: Big data, surveillance, and computational politics. First Monday 19.

UN DESA, 2014. Securing reliable data for development.

Vesset, D.O., 2012. Worldwide Big Data Technology and Services: 2012 – 2016 Forecast (No. 1(238746)), ig Data: Global Overview: Market Analysis. IDC.

Zittrain, J.L., 2008. The Future of the Internet - And How to Stop It. Yale University Press.