Big Data

Milan, S.

Published in:
Routledge Encyclopedia of Citizen Media

License
CC BY

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Big data
Stefania Milan, University of Amsterdam, The Netherlands


Datafication – or the transformation of various aspects of human life into information that can be processed in an automatized fashion and thus monetized (Mayer-Schönberger and Cukier 2013) – alters the exercise of civic participation. Our being citizens is increasingly mediated by digital platforms and the data they generate (Hintz et al. 2018). Citizen media, too, evolve under the pressure of datafication. Data, irrespective of their magnitude, harbour stories and new ways of looking at the world, and come to constitute a novel terrain of engagement for citizen media (Milan and Gutiérrez 2015). New varieties of media activism, exploiting the latest affordances of media innovation for activism purposes, are emerging on the fringes of the datafied society (Gordon and Mihailidis 2016; Pickard and Yang 2017). Data activism, for example, takes advantage of the possibilities for civic engagement offered by big data, incorporating data-based narratives in advocacy campaigns. It increasingly dialogues with more traditional forms of citizen self-expression and contributes to redefining citizen media today (Milan 2017).

This entry explores how and where citizen media meet big data by addressing two important questions: how big data contribute to empowering citizens, and what this empowerment means for the present and future of citizen media. Looking at emergent forms of engagement with data, the entry thus connects the current fascination with quantified information, apps and databases with the possibilities of participation and practices of empowerment enabled by citizen media. After illustrating four core features of citizen media in the age of datafication, it moves on to explore the rise of two novel configurations of citizen media: proactive data activism and civic hacking. The entry concludes by reflecting on the challenges that big data pose to citizen media and, in particular, the emergence of new divides.

From typewriters to big data: Five key features of citizen media

From low-power radio transmitters to photocopier machines, from the static websites of the 1990s to today’s interactive blogs, from neighbourhood television stations to social media platforms, the evolution of citizen media has closely followed the path of technological innovation. Irrespective of the supporting technology, these varieties of citizen media share four features: they empower people through their engagement with technology; present low entry barriers to participation; are collective and communitarian by nature; and allow individuals and groups to exercise democratic agency through media practice. Each of these features is elaborated below.

Empowerment, which lies at the core of citizen media, can be understood as the process through which individuals and groups come to exercise control over both
messages and technology by playing an active role in reshaping the communicative processes they are involved in (Milan 2013). Empowerment thus derives from first-hand engagement with media production – in terms of both content creation and the making of technology and infrastructure (Dunbar-Hester 2014). Learning to programme and broadcast, articulating and disseminating alternative narratives for social change, or assembling a radio transmitter from a kit bought online are some of the processes that enable empowerment.

Characterized by low entry barriers to participation, citizen media projects are accessible to virtually everyone and do not typically require advanced technical skills (Atton 2002). Instead, they capitalize on the orality of radio broadcasting, the spur-of-the-moment reporting of citizen journalists, or the immediacy of blogging and live broadcasting through a smartphone. When more specialized skills are required – for example to operate switchboards or develop software programs – experienced practitioners tend to share their knowledge through dedicated sessions that are open to newcomers.

In their capacity as media “of the people, by the people, and for the people” (Gillmor 2006:195), citizen media are by nature collective projects with a civic orientation (Downing 2001). They seek to interpret and give voice to the needs and imaginaries of their communities of reference – be they geographically bounded or interest-based (Hollander et al. 2002). In addition, citizen media emphasize the collective dimension of freedom of expression, as opposed to an individualistic approach. In the words of a practitioner, the radio “gave me the power of letting the others speak, speaking myself just a little” (Milan 2013:54).

Finally, citizen media expand and multiply spaces for political action. By making media, “a collectivity is enacting its citizenship by actively intervening and transforming the established mediascape … contesting social codes, legitimized identities, and institutionalized social relations … to the point where these transformations and changes are possible” (Rodríguez 2001:20). In other words, citizen media practice enables people to exercise their democratic agency beyond traditional means such as voting and institutional spaces like local administrations. By acquiring and exercising voice (Hemer and Tufte 2016), people change their perception of self, alongside their reality.

The engagement with big data and data infrastructure – the databases, data portals, apps and platforms – increasingly funnels citizen participation and alters some of the core dynamics of citizen media, with both positive and negative consequences. But we first need to understand how engaging with data can contribute to citizen empowerment.

The availability of large quantities of data and the numerical and visual forms information takes today offer novel opportunities for the exercise of democratic participation (Couldry and Powell 2014; Gutiérrez 2018). People can engage with data in a variety of ways: they can find stories in existing datasets, engage with
innovative storytelling techniques, request data when they are not available, and create their own datasets.

Firstly, citizens can put existing data to new uses, promoting transformative experiments that repurpose information. For example, the Illegal Fishing project used publicly available satellite-tracking datasets to map and denounce prohibited fishing activities in Western Africa (Overseas Development Institute, n.d.). Similarly, the Left-to-die Boat project leveraged data generated by existing surveillance technologies to reconstruct the events leading to the death of sixty-three migrants in the Mediterranean Sea. The report supported a series of legal petitions against NATO member states (Forensic Architecture, n.d.).

Moreover, as data visualization software allows users to represent information with minimal effort, low-skill individuals and organizations can also use data for storytelling and to increase the effectiveness of their messages (Kennedy et al. 2016). The non-profit Tactical Tech Collective, for example, argued that “as activists, we can’t sit and wait for people to wade through sixty-page reports. To influence people we must make strong arguments and communicate them using strong evidence” (Visualizing Information for Advocacy, n.d.). By way of illustration, Out of Sight, Out of Mind is an interactive digital visualization of US drone killings in Pakistan, which allows the public to explore the details and scope of the strikes by themselves (Pitch Interactive, n.d.).

Citizens can also request information from public administrations when this is not already made available. By 2018, over 100 countries across the world had implemented some form of freedom of information legislation, designed to facilitate the general public access to data held by national and local institutions. Data portals providing this sort of information are becoming increasingly common. OpenSpending, for example, gathers publicly available data to enable citizens to check and compare fiscal information from 79 countries (Openspending, n.d.). Websites and entities supporting data requests are also proliferating. ControlaTuGobierno, for instance, is a Mexican non-governmental organization promoting the right to access information in marginalized communities (Controla tu Gobierno, n.d.).

Finally, citizens can create their own datasets to contribute to alternative narratives of a given problem. For instance, the Syrian Archive project (Syrian Archive, n.d.) adopts open source intelligence techniques to curate a body of visual evidence of human rights abuse in the Syrian conflict that can be used for reporting, advocacy and accountability purposes (Deutch and Habal 2018). In the same way, the Bellingcat project, founded by a citizen journalist, engages in “open source and social media investigation” to scrutinize conflicts and the criminal underground, and offers guides “so others may learn to do the same” (Bellingcat, n.d.).

What these data activism projects have in common is the belief that data can be effectively used for social change. As is also the case with citizen media, even non-experts can engage with information and technology for activist purposes. However,
big data have not only introduced innovation in the tactics, but also fostered novel cooperation. In the absence of the necessary skills, unprecedented alliances have emerged: for instance, Hacks/Hackers connects journalists (hacks, who can tell stories) and technologists (hackers, who understand data).

The rise of proactive data activism and civic hacking

The importance of proactive data activism is borne out by a number of recent initiatives. InfoAmazonia, an example of proactive data activism launched in 2012, has produced investigative reports on the status of the biggest tropical forest in the world. By promoting data transparency, InfoAmazonia contributes to blocking the illegal deforestation and wildfires endangering the Amazon area (Infoamazonia, n.d.). At the core of this citizen media initiative are self-collected data about carbon monoxide, forest fires, water quality and level, and deforestation. Data are compiled through a self-organized data infrastructure that consists of sensors and mobile applications generating crowdsourced information. Combining satellite imagery with data gathered by activists, InfoAmazonia publishes interactive maps and alternative cartographies, and trains journalists, campaigners and communities in data collection and storytelling (Milan and Gutiérrez 2017). It illustrates the synergies that may obtain when the values and practices of citizen media capitalize on the novel opportunities offered by big data.

Combining complex information with innovative storytelling and collective organizing, proactive data activism embraces tactics of engagement with data seeking to take advantage of the wealth of information of big data for social change purposes. As illustrated in Figure 1, this form of activism emerges at the intersection between citizen media, journalism, big data, and activism/advocacy. From journalism it takes the investigative ethos and the meticulous search for the truth; from the activism and advocacy realm the desire to make of the world a better place; from the big data domain a taste for verified information and strong evidence, such as statistics. From the citizen media domain proactive data activism derives its collective orientation, first-hand engagement with technology, storytelling with a purpose, and the possibility of exercising democratic agency through media practice. Finally, proactive data activism is also inspired by advocacy journalism and data journalism, a specialty combining traditional journalistic storytelling with advanced data analysis to produce investigations (Anderson 2018).

Civic hacking is another important phenomenon emerging at the intersection of citizen media and the big data domain. Civic hackers take advantage of publicly available information, usually government data, to improve institutional output and democratic governance. Organizing hackatons (i.e., marathons to hack data) that oddly combine the do-it-yourself culture of grassroots movements and Silicon Valley technocratic approaches, civil hackers exercise the right to information in a bid to identify solutions to social, economic and environmental challenges (Schrock 2016). Civic hacking shares with citizen media the premise that democratic agency can be exercised daily though engagement with information. They also share an interest in creating bottom-up infrastructure to match the values of citizen media. In this vein,
the Open Knowledge Foundation launched the Comprehensive Knowledge Archive Network (CKAN, n.d.), a web-based management system for the storage and distribution of open data, used by both non-profit and governmental organizations.

The emergence of novel divides

Proactive data activism and civic hacking signal a change in attitude towards big data that has emerged from within citizen media and storytelling projects. However, while the availability of data and the accessibility of data infrastructure clearly provide unparalleled opportunities for citizen media practitioners, they also present new challenges. In particular, novel divides emerge on a number of fronts: skills and data access capabilities vary among citizen media practitioners; the culture of individualism typical of hacking and other forms of tech activism is at odds with the collective nature of citizen media as understood in this entry; the prevalence of positivism associated with big data might hamper value-based advocacy activities; and the high concentration in the data brokers market might in the long run debilitate the efforts of citizen media to control their own infrastructure.

The skills required to engage with data and data analysis are not readily available within the citizen media realm. While software increasingly simplifies tasks, for example, by allowing even non-experts to produce appealing infographics, the entry barriers remain relatively high, especially at the cultural level. For most citizen media practitioners, handling big data is not yet part of their core skills set. In addition, software is usually available only in English. Language, literacy and income remain important obstacles for practitioners in developing countries in particular. On the whole, big data have introduced a novel divide, between the ‘haves’ – individuals and organizations who have access to data, state-of-the-art hardware and software, and the knowledge to meaningfully make sense of data – and the ‘have nots’, who are excluded from the “data revolution”.

On the other hand, the power of attraction of data and data infrastructure goes hand in hand with a culture of individualism and expertise typical of the hacker scene and, more widely, of Silicon Valley. Although data projects require team work, they are markedly rooted in personal capabilities and individual practices. Big data have therefore introduced a novel tension between the individual and the collective dimensions of citizen media, which risks sideling the cooperative nature of existing projects. Unfortunately, data-related skills are not easily transferable – at least not at the speed and the scale at which citizen media practitioners have traditionally involved newcomers in their activities.

The positivism that surrounds big data (Mosco 2014) represents another important divide, as it results in a tendency towards technocratic solutionism that is particularly visible in civic hacking. In this field of action, quantified data tend to be regarded as irrefutable information and receive higher consideration than norm- and value-based reasoning. Although data activists seek to promote alternative
epistemologies countering the positivist ethos of big data (Milan and van der Velden 2016), there remains a fundamental tension between different ways of understanding and communicating the world around us – and the ways to change it. For example, the use of claims based on quantified data to justify collective demands – e.g. “we are the 99 percent”, the slogan of the Occupy Wall Street mobilization – is increasingly popular, as are strategies aimed at gathering evidence as a fundamental step towards the resolution of problems (Simon 2012). This tension between value- and data-based reasoning deeply affects advocacy organizations and their strategies; it is indirectly reproduced by funding organizations, who prefer to support data-based projects over value-driven ones, thus reinforcing the divide mentioned above.

Finally, with the increase in magnitude and importance of big data, new data brokers such as Google and Facebook have become central to the functioning of citizen media. Although InfoAmazonia, for example, was able to produce its own data-reporting app, most citizen media are not able to do so and have to rely on commercially available products. As a result, big data contributes to weakening the self-organized infrastructure of citizen media, accelerating the move towards commercial services and promoting an even higher level of centralization of power in the hands of those who produce software, hardware and content. Unfortunately, most of these data brokers have values and agendas that are radically different from those of citizen media, and expose citizen media to monitoring and surveillance (Dencik and Leistert 2015).

**Future directions**

Big data have come to constitute the new frontline of citizen media practice. Two novel types of citizen media initiatives emerge, namely proactive data activism and civic hacking. The former seeks to incorporate data-based reasoning in citizen storytelling, while the latter takes advantage of available data to ameliorate democratic governance. Both share the hopeful belief that data and technology can contribute to social change. Datafication, however, presents citizen media with both opportunities and challenges. On the positive side, citizens can use data for advocacy purposes by supporting stories with data, engaging in innovative persuasive storytelling, requesting data when not readily available, and creating their own datasets to monitor phenomena of concern. Conversely, big data contribute to generating new divides within the domain of citizen media, reflecting varying levels of skills, income and access. New tensions arise between the individual capabilities of data experts and the collective nature of citizen media, and between the positivism of big data and the more traditional value- and norm-based advocacy. Finally, the self-organized infrastructure of citizen media is finding it increasingly difficult to compete with novel data brokers offering complex services that would otherwise be out of reach for grassroots practitioners.

**Recommended reading**

Positions a new development in citizen media, namely open data, in relation to civic hacking and journalism, but also democratic participation.


Explores proactive data activism in a number of successful projects at the intersection of advocacy journalism, data journalism, and civic technologies.


Investigates how tactics, identities and worldviews of citizen media practitioners evolve as a consequence of their exposure to data and data infrastructure.


Reviews key aspects of data activism at the intersection with citizen media, including alternative epistemologies and the everyday experience of practitioners in engaging with data.
Figure 1. The emergence of two new configurations of citizen media (adapted from Milan and Gutiérrez 2015).