The grounded theory method to study data-enabled activism against corruption: Between global communicative infrastructures and local activists' experiences of big data

Alice Mattoni
University of Bologna, Italy

Abstract

Anti-corruption efforts from the grassroots that make use of digital media to hinder corrupt behaviors are flourishing worldwide. In many cases, these efforts see activists interact with big data along with other types of data. They do this in the framework of broader communicative infrastructure in which activists create, employ, and spread big data to support their struggles. As well, they do so differently, according to a diverse range of activists' local situations across the world. The article uses examples of anti-corruption efforts in Brazil, India, and Spain to illustrate how the grounded theory method might help researchers to produce knowledge that escapes a universalistic and global vision of datafication detached from activists' lived and local experiences. The article first briefly outlines what grounded theory is, the main steps in a grounded theory study, and its applications in media and communication studies. It then moves

Corresponding author:
Alice Mattoni, Department of Political and Social Sciences, University of Bologna, Strada Maggiore 45, Bologna 40125, Italy.
Email: alice.mattoni@unibo.it

to a broader discussion of two relevant elements of grounded theory – sensitizing concepts and theoretical sampling – in taking into consideration data-enabled activism as an emergent phenomenon that might take many shapes. Then, it considers the emphasis on the situation in which data-enabled activism spreads out through a brief discussion of one relevant development of grounded theory, which is situational analysis, to solve the tension between the global and the local in data-enabled activism.

Keywords
Social movements, grounded theory, data-activism, big data, anti-corruption movements

Introduction

In 2013, the Spanish activist organization *XNet*, based in Barcelona, developed *BuzonX*, a digital platform that allowed for the secure, anonymous leaking of relevant information on corruption and other related crimes. Through this platform, activists were able to gather crucial information that led to a criminal investigation on the misbehaviour of Bankia's managers in the years that preceded the economic and financial crisis. In 2010, the Indian civil society organization *Janaagraha* developed *I Paid A Bribe*, a digital platform that permits citizens to denounce cases of petty corruption among public officials in the local administrators. With more than 100,000 bribes reported all over the country, *I Paid A Bribe* rendered visible corrupt behaviour to the broader public in a dynamic way. In 2017, a group of citizens launched a software named *Rosie*, in Brazil, that can check the public expenditures of elected members of the chamber of deputies. The software runs algorithms to establish if these expenditures are suspicious and then posts this information automatically on Twitter through a bot, also named *Rosie*.

These three examples make use of big data, data analytics and algorithmic automation in combination with other types of data to denounce corruption and related crimes from the grassroots, through people's intervention, and beyond the actions of institutional actors, like governments and their anti-corruption agencies. In so doing, *BuzonX*, *I Paid a Bribe* and *Rosie* are telling examples of the intertwining between big data and contentious politics. With this regard, Davide Beraldo and Stefania Milan speak about the emergence of data-enabled activism, a form of activism that employs big data as leverage to sustain activists' struggles (Beraldo and Milan, 2019). Data-enabled activism includes big data within activists' repertoire of contention, hence exerting a certain level of agency concerning datafication instead of being subjected to it (Beraldo and Milan, 2019).

When considering data-enabled activism against corruption, big data and the algorithms that work with them might be relevant, but not as the only leverage.

Activists employ digital media that run on algorithms, like social media platforms, to sustain their mediated communication beyond their direct engagement with big data. For instance, they use digital media to organize, participate and protest (Rotberg, 2017), and also as spaces for discussion and awareness, like in the case of the Facebook group '1,000,000 Facebookers Support Chandra Hamzah and Bibit Samat Riyanto' in Indonesia between 2009 and 2011 (Sulistyo and Azmawati, 2016). Activists, then, employ data about corruption that they did not necessarily create through the combined use of digital media platforms and algorithms that work with them. The anti-corruption sector is heavily relying on the production and circulation of massive amounts of data related to corrupt behaviours. This large amount of information comes from multinational surveys, like the well-known Perception of Corruption Index created by the transnational civil society organization Transparency International.

Data-enabled activism, thus, might be a significant label to point out the relevance that big data have for certain types of activism. However, at the same time, activists engaged in data-enabled activism against corruption develop, appropriate and employ big data within a broader repertoire of communication. It is from that repertoire that activists select their communication strategies to meet various goals, from engaging more and more supporters to sustaining their internal organization activities. In some cases, big data could be central for their communication strategies; in others, they could stay in the background; almost all the time, they combine with other media technologies and communication channels. In short, activists who struggle against corruption, embed big data, data analytics and algorithmic automation in the broader range of activists' experiences with communication and information technologies. Perhaps even more importantly, they do this from different countries and cities across the world, in which various types of corrupt behaviour do exist, and there is an unequal distribution of the skills necessary to deal with big data in grassroots politics. To consider these differences, some scholars recently called for an understanding of datafication rejecting its universalism (Milan and Treré, 2019), and arguing that we need to focus not so much on datafication in itself, but rather on experiences of datafication (Kennedy, 2019).

When thinking about data-enabled activism, this means that we should not consider its interactions with datafication in abstract terms, but appreciate how activists and their activists' organizations actually exert some type of agency towards big data, also by combining them with other types of data or, as it is often the case, matching them with diverse forms of communication, both mediated and non-mediated. Which is, though, the methodological path that would lead us to understanding the whole experience of dealing with big data and datafication from an activists' viewpoint and in the framework of mobilizations, protest campaigns and social movements? While I acknowledge that the roads to be taken might be varied and all equally valid in pursuing this endeavour, in what follows, I show how one of the paths to take could be grounded theory as a family of methods rooted in the tradition of qualitative methodology.

In the remainder of this contribution, I seek to illustrate how the grounded theory method might help researchers to produce knowledge that escapes a universalistic reified vision of datafication detached from the lived experiences of the many social actors that deal with it, including social movement actors. At the same time, I also attempt to illustrate how this might consequentially lead to the production of a grounded theory on data-enabled activism that is able to situate them within the broader communicative ecology in which activists create, employ and spread big data to support their struggles. In the next section, I briefly outline what grounded theory is, the main steps in a grounded theory study, and its applications in media and communication studies. I then move to a broader discussion of two relevant elements of grounded theory – sensitizing concepts and theoretical sampling – in taking into consideration data-enabled activism as an emergent phenomenon that might take many shapes. Then, I consider the emphasis on the situation in which data-enabled activism spreads out through a brief discussion of one relevant development of grounded theory, which is situational analysis, to solve the tension between the global and the local in data-enabled activism. Finally, in the last section, I will propose some further comments on grounded theory and situational analysis for the study of data-enabled activism.

To provide concrete examples that might sustain my arguments, I employ the three cases of data-enabled activism against corruption that I outlined earlier. As for *BuzonX*, I gathered information during previous qualitative research that I conducted on the 15 MPaRato campaign in Spain (Mattoni, 2017). As for *Rosie* in Brazil and *I Paid a Bribe* in India, I rely on information that I found through secondary sources and desk research (Cordova and Gonçalves, 2019; Venkatesh, 2016).

A very brief introduction to grounded theory

The grounded theory method has its roots in the seminal work that Barney Glaser and Anselm Strauss began in the 1960s when investigating the awareness of dying of terminal patients in the field of nursing studies (Glaser and Strauss, 1967). It then became central in the qualitative tradition, developed to a great extent over the years, and ended up including many versions. The most notable variants are the classic grounded theory of Glaser and the early Strauss (Glaser and Strauss, 1967), the revised version of grounded theory by Strauss and his co-author Corbin (Strauss and Corbin, 1998) and the constructivist grounded theory of Charmaz and Bryant (Bryant, 2017; Charmaz, 2014). Grounded theory today can be understood as a family of methods whose main objective is to produce middle-range theory starting from a varied array of data – mostly qualitative, but not only – through a process of incremental abstraction (Bryant, 2017) that puts coding and the comparison of codes and cases at the centre of the analytical process (Glaser and Strauss, 1967).

The grounded theory method does not generally seek to test hypotheses or to put at work preconceived concepts, constructs and models. Due to this aspect,

Table 1. Main steps in grounded theory.

Initial case selection	One case study that let the researcher address the initial general research interest/question
Data gathering	In-depth interviews, participant observation, documents and other types of relevant data
Initial coding	Codes data word-by-work, line-by-line or incidents-by-incidents
	Produces transitional codes
	Writing of memos related to the emerging codes
Theoretical sampling	Sample new case studies and/or new types of data based on the categories that emerge from coding;
	Further data gathering, then go back to initial coding
Selective coding	Consider the most significant or the most significant codes to code large amounts of data;
	Writing of memos related to the emerging codes;
	Produces categories and sub-categories
Theoretical sampling	Sample new case studies and/or new types of data based on the categories that emerge from coding;
	Further data gathering, then go back to selective coding
Theoretical coding	Finds out connections between categories;
	Connects categories to the extant literature;
	Connects categories to theoretical codes;
	Produces concepts

scholars often employ it to explain emergent phenomena on which there is limited knowledge or to provide different explanations of phenomena on which existing literature cannot be applied (for an example of this, see Coe, 2009). Over the years, some scholars in the field of communication and media studies have also applied the grounded theory method. Vivian Martin (2008), for instance, employed the classic grounded theory method in audience studies producing a grounded theory of news-attending as a daily regimen that goes unnoticed in the daily lives of people. Astrid Gynnild (2007, 2016), instead, investigated how Norwegian professional journalists performed their job in a multimedia environment and suggested that the method could be suitable to develop theory related to the field of media production studies.

Although there might be some relevant variations according to the type of grounded theory method that researchers use, the starting point is usually a general research problem to which a more specific research question is linked. Then, researchers develop their investigation following the steps that I single out in Table 1 that adapts Urquhart (2017).

At first sight, the grounded theory method might seem not so different from other types of qualitative methods. The reason for this is that many other qualitative traditions also use the data that grounded theorists employ: qualitative interviews, participant observation and documental research are three standard ways of constructing data across the broad field of qualitative research.

Also, the grounded theory method frequently requires immersive fieldwork similar to other types of qualitative traditions, like ethnography. However, its uniqueness stands neither in the type of qualitative data on which it rests nor on its strict relationship with the data that researchers construct when doing fieldwork. Instead, its peculiarity lies in the recursive logic that connects the sampling of case studies, the data gathering and the coding stages, which is a logic of theory building, rather than exploration and verification.

In the next sections, I address two specific elements that are important in such a logic: sensitizing concepts and theoretical sampling. I argue that these two elements render the grounded theory method suitable to grasp data-enabled activism, avoiding the trap of producing knowledge that is too general and too detached from the actual experiences of the people who live datafication while engaging with politics from the grassroots.

Big data as a sensitizing concept for the study of data-enabled activism against corruption

The grounded theory method employs sensitizing concepts as relevant tools that might guide the initial selection of case studies, but also the data gathering and data analysis. They are not prescriptive and, as Blumer (1954) famously put it, a sensitizing concept does not tell researchers what to see, but rather 'it gives the user a general sense of reference and guidance in approaching empirical instances' and 'merely suggests direction along which to look' (p. 7). Along these lines, we can also deal with big data as a concept that is neither entirely defined nor completely definable until we put it at work in actual experiences of data-enabled activism. We might thus interpret big data not as a stable concept that we already defined before commencing our research. Instead, we might consider it as an ever-evolving concept, whose boundaries are fuzzy and which suggest where to look instead of providing answers on what we should find out there in societies and, more specifically, in the realm of grassroots politics when it comes to big data.

BuzonX, *I Paid a Bribe* and *Rosie* involve a varied ensemble of social actors, both human and non-human. Such a wide array of actors contributes to the construction of and, at the same time, is positioned within a complex communicative infrastructure, where a variety of media technologies, information flows and communication processes intersect. For this reason, these actors enter in different relationships with big data, leading to the deployment and recombination of different types of meanings related to what big data are according to the viewpoint from which actors look at them. In communicative infrastructures where a varied ensemble of social actors come together, the meanings of big data are necessarily polysemic and never univocal. For the study of data-enabled activism and its relationship with datafication, then, the grounded theory method would suggest starting from big data as a sensitizing concept, that remains open to many different interpretations at the inception of the empirical research and guides

researchers to look for case studies that might not seem to put interactions with big data at their centre.

If we look for data that are big in their volume, variety and velocity (Kitchin and McArdle, 2016), for instance, we might not find them in the three cases of data-enabled activism against corruption that I briefly presented earlier. In the case of *BuzonX*, data coming from whistleblowers might not come in large volumes. However, the leak of thousands of emails required a great deal of coordination for activists as well as the management of an appropriate amount of human resources, including lawyers and journalists, to make the leaked contents available to and readable for the broader public. Furthermore, in the case of the software *Rosie* data do not come in great varieties. All related to the same public dataset registering the expenses of elected parliamentarians in Brazil, such data are rather homogeneous. Finally, data come into existence not according to a real-time velocity, as in the case of *I Paid a Bribe*, in which people do not necessarily create and post online denunciations of bribes as soon as the corrupt behaviour happens. In short, these examples show that we are not strictly speaking about big data; and yet, activists support their collective actions against corruption through the creation, transformation and distribution of what, for them, might be abundant amounts of data, that are difficult to digest and that come at a swift pace.

Thanks to the role that it gives to sensitizing concepts, the grounded theory method makes it possible to select the initial case studies with a certain degree of freedom. They do not have to be cases of a definite concept that sets clear boundaries between what is a case of that concept and what is not. When thinking about data-enabled activism against corruption, it is clear that the three cases above are not immediately and explicitly linkable to the concept of big data, if we use it as a definite concept. However, when considering big data as a sensitizing concept, it would be possible to consider these three cases from different perspectives, also assigning different meanings to them according to the position from which activists look at big data within the communicative infrastructure that sustain the grassroots opposition to corruption. What would emerge, then, are concepts and theories that might not be universally valid, but certainly make sense for the situated experiences on data-enabled activism. It is from these experiences, rather than from purely deductive theories of datafication, that we might fully appreciate the challenges and opportunities that big data pose to grassroots politics in connection to the other media technologies, information flows and communication processes that activists engage with.

Theoretical sampling as a guiding principle for comparison within the grounded theory method

Sensitizing concepts are relevant to select the initial case study, gather some data and begin coding in the grounded theory method. However, the grounded theory method does not usually stop after one case study and rest, instead, on robust

comparative procedures. With this regard, when Barney Glaser and Anselm Strauss first presented their method to the academic community, they dubbed it the *constant comparative method* and not grounded theory (Glaser and Strauss, 1967), the latter instead being the product of the inquiry through the former. Comparison, hence, is the essence of the grounded theory method and lays at its heart in at least two ways.

First, during the coding stages, the codes that emerge from the data are continuously contrasted across the whole dataset, to refine the coding scheme and, more importantly, to pass from one stage of coding to the other. In so doing, researchers produce more abstract categories grounded in such codes and, eventually, that focus on the relationships between categories (Kelle, 2019). It is indeed through a continuous comparative attitude towards the coded data that grounded theorists can produce constructs, models and theories that are deeply tied to their empirical materials while being able to provide more general understandings of the phenomena under investigation.

Second, we find a strong comparative stance in the grounded theory method also when we move from initial sampling to theoretical sampling (Glaser and Strauss, 1967). Initial sampling allows researchers to start gathering data according to different types of criteria and techniques, like the representativeness of the sample or the snowball sampling strategy. However, after some rounds of open and/or selective coding, some categories and concepts begin to emerge that need further refinement. It is at that point that another crucial step in sampling occurs, named theoretical sampling, that expands both the types of data and the groups to which these data refer, based on analytic reasons.

For instance, after initial coding and focused coding, some categories might emerge on data-enabled activism against corruption that calls researchers to gather other data than in-depth interviews and/or to consider other groups of people beyond activists. Researchers, however, do not make these decisions in advance: it is through the first rounds of coding and the preliminary abstractions that they sample new data and/or new groups in order to discover more properties in their categories as well as new relations among these properties. In this case, sampling aim is to further expand theory development on the subject matter. For instance, to understand better one of the concepts emerging from one case study on *I Paid a Bribe,* researchers might decide to look at how the technological infrastructure sustaining the platform works daily. To reach this goal, they need to engage with participant observation of the daily work of software developers that take care of the technological side of the endeavour. Alternatively, they can go and look for additional case studies, in which third parties outside the realm of grassroots anti-corruption efforts designed the technological platforms used to denounce corruption. Indeed, theoretical sampling 'may prompt grounded theorists to sample in entirely new empirical areas from those in which they began their study' (Charmaz and Bryant, 2012: 375).

Theoretical sampling, therefore, works as a compass: each time pointing our attention to the most relevant and telling data on data-enabled activism; each time

suggesting to us where to look for new case studies to be taken into consideration to develop an emerging category further or to refine a concept that begins to take shape. Theoretical sampling works in different directions: researchers can look for similar data within similar case studies, different data within the similar case studies, different data within different case studies, or similar data within different case studies (Glaser and Strauss, 1967; Urquhart, 2019). For instance, to refine a concept that emerges concerning several examples of data-enabled activism, theoretical sampling might suggest considering case studies in which activists wanted to engage with big data but then decided not to; these would be case studies that do not fall into the definition of data-enabled activism. Nevertheless, they can be precious to refine a concept that is relevant to explain the processes of data-enabled activism and its relationship with datafication. Similarly, theoretical sampling might push research on data-enabled activism towards the gathering of in-depth interviews and activist documents that speak about other ways through which activists set up their communicative strategies, using media technologies that do not involve the use of big data, algorithmic automation or artificial intelligence in the first place. Theoretical sampling allows constructing knowledge based on a range of case studies that speak about the varied lived experiences of data-enabled activism while producing concepts that might resonate beyond the specificities of each of them.

Considering the situation when investigating data-enabled activism

In the previous sections, I showed how sensitizing concepts and theoretical sampling in the grounded theory method are relevant to go beyond a universalistic understanding of datafication when it comes to its relationship with data-enabled activism against corruption. Beyond avoiding overgeneralizing interpretations, another relevant aspect is grasping the experiences related to datafication in the framework of data-enabled activism. Doing this is particularly tricky because data-enabled activism against corruption is both a global and local phenomenon at the same time.

Grassroots anti-corruption efforts tackle a global problem and do this through the use of digital media platforms that also have a global scale. On some occasions, activists might appropriate already existing commercial platforms, like Facebook and Twitter, to denounce corruption and crowdsource data on corrupt behaviours. In other cases, they might adapt locally digital platforms that other activists created elsewhere, like in the case of *I Paid a Bribe*, whose original project in India has been then transferred in other countries across the world.

How can we grasp these types of experiences, which are deeply entrenched with global processes that also have an active local component in the way they unfold? The sound comparative attitude of the grounded theory method already helps developing knowledge that looks at what happens in specific cases, expanding its

scope through a reasoned selection of case studies and data, sometimes even at the cross-country level. However, it is a further extension of grounded theory, that is situational analysis, that might become relevant to grasp the trans-local experiences of data-enabled activism and its relationship with datafication as a process. Adele Clarke developed situational analysis to focus on the ecologies of relations that exist between the various elements found in a given situation (Clarke et al., 2018). She starts from the idea that the situation is the basic unit of analysis that involves 'a somewhat enduring arrangement of relations among many different kinds and categories of the element that has its own ecology' (Clarke et al., 2018: 17). A situation, for instance, might be the anti-corruption activist work to counter corruption through data in Spain.

Situational analysis constructs different maps that allow first to describe the situation under inquiry. First, a situational map that lists and positions all the elements that are part of the situation, ranging from human to non-human, from individual to collective, from concrete to discursive elements, and so on. In the case of *Rosie*, for instance, an initial situational map might depict the situation of activists' work against corruption and include, activists, software developers, elected MPs, elected MPs expenses register, Twitter bots, Twitter as a platform, Twitter users, ReTweets, Rosie itself, technical skills just to name a few. This map, then, renders immediately visible which other types of media technologies, information flows and communication processes, beyond those related to big data, are included in data-enabled activism.

The step that follows the creation of a situational map is usually the design of relations among elements in the construction of a relational map. *Rosie*, for instance, is at the same time a software that employs algorithms and a Twitter bot that creates automated Tweets. However, it is also a loose network of software developers with ties that go well beyond Brazil, in the social coding platform GitHub, and a dispersed crowd of people that contribute to assessing what the Twitter bot suggests when denouncing publicly suspicious expenses. The relational map captures the various relations that characterize the elements – human and non-human alike – of the situational map. In so doing, the relational map is able to grasp the relations that are at work between elements that pertain more to the realm of big data and those that are related, instead, to other types of data, media and forms of communication.

Another map is the social worlds/arena maps that consider 'all of the collective actors and the arenas of commitment within which they are engaged in ongoing discourse and negotiations' (Clarke et al., 2018: 18). Again, when looking at *Rosie*, what we see are not just the different elements that co-constitute the situation of data-enabled activism against corruption in Brazil, but also different types of collective actors, that goes well beyond the activist circle that gathers around the experience of *Rosie*. The social worlds and the arenas might be positioned at the local level, but also at the national and even at the transnational one. When thinking about the case of *Rosie*, some of the social worlds might be those of social media, political parties', the tech industry, activists against corruption and transnational NGOs, just to name a few of them. Again, also in this case, it is clear that the social worlds/arenas maps can

detect those worlds and arenas that are not directly, or even not at all related, with big data in the framework of the data-enabled activism project *Rosie*.

These three maps allow seeing data-enabled activism conflating the local and the global within the same situation, considering the elements that pertain to the local level and those that are linked to the global level at the same time.[1] In so doing, we might see the global at work in the local, without conceiving datafication as part of an external context of data-enabled activism, but rather as a whole part of what activists do to counter corruption through the use of big data and other types of data. Ultimately, then, the combined use of the grounded theory method and situational analysis might let us see, comparatively, how the global entangles with the local in various situations within the same country, but also across the world.

Conclusion

In this contribution, I started with some examples of data-enabled activism against corruption to provide some methodological reflections on how we can construct sounding knowledge on datafication, able to escape the temptation of universalistic explanations that reify the process of datafication with which activists' efforts intertwine. To conclude, three tensions emerged in the previous pages that cross data-enabled activism with regard to its connection with the multifaceted communicative infrastructure in which it is embedded: universal meanings versus situated meanings attached to the concept of big data; modes of communication primarily revolving around big data versus modes of communication pushing big data in the background; global flows of big data versus local usages of big data. The grounded theory method and its extension, situational analysis, allow to take into consideration these tensions looking at data-enabled activism and its relationship with datafication not in abstract terms, but rather as a social processes and relational ecologies that are deeply tied to the flesh and bones of activism that make use of big data, data analytics and algorithmic automation as leverage to counter corruption from the grassroots. Indeed, grounded theory develops 'concepts intimately connected with, and responsive to, actually lived social life' (Wasserman et al., 2009: 359) that through the use of constant comparison across codes, between codes and categories, and across different cases can generate theories able to speak about datafication from a global perspective while holding its validity also in specific situations of data-enabled activism. This, ultimately, means developing theoretical explanations about processes and relations of data-enabled activism that situate more general categories and concepts into the encounters that activists have with big data, data analytics, and algorithmic automation when they engage in grassroots politics. It is from such a perspective that we can understand which is the type of political agency that activists must nurture to deal with datafication not as passive individuals, but as active collective forces that might use data-enabled activism to disempower the powerful and empower the powerless in their struggle against corruption.

Funding

Note

1. Although equally relevant, I do not discuss here the fourth map that researchers construct in situational analysis: the position map, whose task is to single out the main issues, the positions on the central issues and the lack of positions that we can find in a situation (Clarke et al., 2018).

References

Beraldo D and Milan S (2019) From data politics to the contentious politics of data. *Big Data & Society* 6: 1–11.

Blumer H (1954) What is wrong with social theory? *American Sociological Review* 19(1): 3–10.

Bryant A (2017) *Grounded Theory and Grounded Theorizing: Pragmatism in Research Practice*. Oxford: Oxford University Press.

Charmaz K (2014) *Constructing Grounded Theory*. London: SAGE.

Charmaz K and Bryant A (2012) Grounded theory. In: Given LM (ed.) *The SAGE Encyclopedia of Qualitative Research Methods*. Thousand Oaks, CA: SAGE, pp. 375–377.

Clarke AE, Friese C and Washburn RS (2018) *Situational Analysis: Grounded Theory after the Interpretive Turn*. Thousand Oaks, CA: SAGE.

Coe A-B (2009) 'Being in the spaces where decisions are made': Reproductive rights advocacy and policy influence in two regions of Peru. *Social Movement Studies* 8(4): 427–447.

Cordova Y and Gonçalves EV (2019) Rosie the robot: Social accountability one tweet at a time. Available at: https://blogs.worldbank.org/governance/rosie-robot-social-account ability-one-tweet-time.

Glaser B and Strauss AL (1967) *The Discovery of Grounded Theory. Strategies for Qualitative Research*. Chicago, IL: Aldine.

Gynnild A (2007) Creative cycling of news professionals. *The Grounded Theory Review: An International Journal* 6: 67–95.

Gynnild A (2016) Applying grounded theory in media production studies. In: Paterson C, Lee D, Saha A, et al. (eds) *Advancing Media Production Research: Shifting Sites, Methods, and Politics*. London: Palgrave Macmillan, pp. 115–130.

Kelle U (2019) The status of theories and models in grounded theory. In: Bryant A and Charmaz K (eds) *The SAGE Handbook of Current Developments in Grounded Theory*. London: SAGE, pp. 68–88.

Kennedy H (2019) Introduction. In: Treré E and Stephansen HC (eds) *Citizen Media and Practice: Currents, Connections, Challenges*. London: Routledge, pp. 207–211.

Kitchin R and McArdle G (2016) What makes big data, big data? Exploring the ontological characteristics of 26 datasets. *Big Data & Society* 3(1): 2053951716631130.

Martin VB (2008) Attending the news: A grounded theory about a daily regimen. *Journalism* 9(1): 76–94.

Mattoni A (2017) From data extraction to data leaking. Data-activism in Italian and Spanish anti-corruption campaigns. *Partecipazione e Conflitto* 10(3): 723–746.

Milan S and Treré E (2019) Big data from the south(s): Beyond data universalism. *Television & New Media* 20: 319–335.

Rotberg RI (2017) *The Corruption Cure: How Citizens and Leaders Can Combat Graft*. Princeton, NJ: Princeton University Press.

Strauss AL and Corbin JM (1998) *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*, 2nd edn. Thousand Oaks, CA: SAGE.

Sulistyo PB and Azmawati AA (2016) The roles of Facebook in creating new anti corruption social movement through online community in Indonesia. *International Journal of Scientific and Research Publications* 6(7): 33–37.

Urquhart C (2017) *Grounded Theory for Qualitative Research: A Practical Guide*. Thousand Oaks, CA: SAGE.

Urquhart C (2019) Grounded theory's best kept secret: The ability to build theory. In: Bryant A and Charmaz K (eds) *The SAGE Handbook of Current Developments in Grounded Theory*. London: SAGE, pp. 89–106.

Venkatesh K (2016) *I paid a bribe: Using technology to fight corruption in India*, 18 May. IGC. Available at: http://www.theigc.org/blog/i-paid-a-bribe-using-technology-to-fight-corruption-in-india/.

Wasserman JA, Clair JM and Wilson KL (2009) Problematics of grounded theory: Innovations for developing an increasingly rigorous qualitative method. *Qualitative Research* 9(3): 355–381.