



HAL
open science

Data politics

Didier Bigo, Engin Isin, Evelyn Ruppert

► **To cite this version:**

Didier Bigo, Engin Isin, Evelyn Ruppert. Data politics. Big data & society, 2017, 4 (2), 10.1177/2053951717717749 . hal-03574146

HAL Id: hal-03574146

<https://hal-sciencespo.archives-ouvertes.fr/hal-03574146>

Submitted on 15 Feb 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Data politics

Evelyn Ruppert¹, Engin Isin² and Didier Bigo³

Big Data & Society
 July–December 2017: 1–7
 © The Author(s) 2017
 Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
 DOI: 10.1177/2053951717717749
journals.sagepub.com/home/bds



Abstract

The commentary raises political questions about the ways in which data has been constituted as an object vested with certain powers, influence, and rationalities. We place the emergence and transformation of professional practices such as ‘data science’, ‘data journalism’, ‘data brokerage’, ‘data mining’, ‘data storage’, and ‘data analysis’ as part of the reconfiguration of a series of fields of power and knowledge in the public and private accumulation of data. Data politics asks questions about the ways in which data has become such an object of power and explores how to critically intervene in its deployment as an object of knowledge. It is concerned with the conditions of possibility of data that involve things (infrastructures of servers, devices, and cables), language (code, programming, and algorithms), and people (scientists, entrepreneurs, engineers, information technologists, designers) that together create new worlds. We define ‘data politics’ as both the articulation of political questions about these worlds and the ways in which they provoke subjects to govern themselves and others by making rights claims. We contend that without understanding these conditions of possibility – of worlds, subjects and rights – it would be difficult to intervene in or shape data politics if by that it is meant the transformation of data subjects into data citizens.

Keywords

Politics, fields, power, rights, citizens, professions

Introduction

Data enacts that which it represents. This deceptively simple thought signifies two things. First, to collect, store, retrieve, analyse, and present data through various methods means to bring those objects and subjects that data speaks of into being. Data sciences such as statistics, probability, and analytics have emerged not because they have merely quenched our curiosities but because these sciences have been useful for the objects and subjects they have brought into being for the purposes of governing and/or profit. Second, to speak constantly about data as though it either represents or records subjects and objects and their movements independent from the social and political struggles that govern them is to mask such struggles.

If not for the rapid development of the Internet and its connected devices ‘data’ would have probably remained a relatively obscure concept or term confined to these sciences. Yet, data has become a social and political issue not only because it concerns anyone who is connected to the Internet but also because it reconfigures relationships between states and citizens.

Being connected to the Internet may conjure up images of hipsters with shiny aluminium devices but just about every device is now connected and generating vast quantities of digital traces about interactions, transactions, and movements whether users are aware or not. What started as an ostensibly liberated space the Internet rapidly became the space over and through which governments and corporations began collecting, storing, retrieving, analysing, and presenting data that records what people do and say on the Internet. This ranges from who communicates with whom, who goes where, and who says what – and much more besides. This is now being augmented with data that people collect about themselves, especially their relations, body movements, and measurements; the amount and

¹Goldsmiths, University of London, UK

²Queen Mary University of London, UK

³Sciences Po, FR

Corresponding author:

Evelyn Ruppert, Goldsmiths, University of London, Warmington Tower, Lewisham Way, London SE14 6NW, UK.

Email: e.ruppert@gold.ac.uk



range of data that has become available is, as everyone now knows, staggering. There has never been a state, monarchy, kingdom, empire, government, or corporation in history that has had command over such granular, immediate, varied, and detailed data about subjects and objects that concern them. What exactly governments, corporations, and a whole series of agencies and authorities collect, analyse, and deploy is complex but it is now generally understood that data has become a major object of economic, political, and social investment for governing subjects. This development has been captured by the term Big Data to mark a departure from conventional forms of data and statistical knowledge. While first coined by industry, Big Data has come to have different meanings and uses but significantly, and along with the increasing ubiquity of data in everyday life, the term has become less prominent. Notably, attention has started shift to a focus on computation and analytics such as algorithms, machine learning, and artificial intelligence. Yet, data remains a key matter of concern as both the product and condition of computation and analytics.

Reactions to these developments have largely focused on issues concerning surveillance, privacy, anonymity, and types of conduct that the Internet cultivates about always connected, always measured selves. Perhaps equal to the measure of the influence of the Internet there has been a massive literature on data ranging from warnings about its consequences (surveillance, privacy, isolation) to types of conduct (racism, misogyny, bullyism). Along with this, numerous studies, reports, guidelines, regulations, and legislation concerning data protection and the rights of data subjects have proliferated. To summarize such a huge literature and responses is beyond our remit but we are convinced that dominant responses have three shortcomings that brought us to propose the phrase ‘data politics’ as a field of power and knowledge.

The first concerns the focus on the politics *of or in* data rather than data as a force that is generative of politics. In this view, rather than settled in a database or archive, data has potential force that can be realized in myriad ways through its uptake and deployments. Rather than thinking about data as inert representation, we consider it like language in the way elaborated by Bourdieu (1993) and Butler (1997) to argue that data has a performative power that is resignifying political life. That is, data politics is concerned with not only political struggles around data collection and its deployments, but how data is generative of new forms of power relations and politics at different and interconnected scales.

That data is generative of new power relations and politics is evident in the recent controversies about how Big Data was used in the US election and UK

referendum to create personalized political advertising to influence how people voted. Referring to these electoral uses, George Monbiot writing in *The Guardian* noted that we must act now to own these new political technologies before they own us. He was of course referring to the work of a company called Cambridge Analytica, which is partly owned by US billionaire Robert Mercer, who also happens to be a friend of former UKIP leader Nigel Farage. It was widely reported that the company helped to swing both the US election and the UK referendum by mining data from Facebook and using it to create profiles predicting people’s personalities and then tailoring advertising to their psychological profiles. While some of the claims that this happened were brought into question including denials from Cambridge Analytica, UK’s privacy watchdog – the Information Commissioners Office – has deemed there is sufficient cause to launch an inquiry and it is doing so.

So, in the wake of already uneven power and influence over electoral processes – such as campaign financing and media alliances – we now have misinformation, disinformation, and techniques such as bot-swarms whereby fake online accounts are created to give the impression that large numbers of people support a political position. These are some of the threats that Tim Berners Lee pronounced on the 28th anniversary of the world wide web, which he is credited for founding. He names these threats as the sharing of personal data, misinformation, and political advertising. One of his solutions is a call for new ways for people to own their own data. He argues that the web was built by all of us, and so we all can, and should, play a role in defining its future. The Web We Want campaign is one initiative that he has supported in this regard.

What these examples illustrate is that data and politics are inseparable. Data is not only shaping our social relations, preferences, and life chances but our very democracies. And that is how we want to speak of data politics. However, a problem with these views on data politics is the subjects who are constituted as the addressee and are presumably the affected Internet subjects. This is the second shortcoming that has led us to articulate what we call data politics. It concerns atomism: often such pronouncements address atomized individuals who need to protect themselves from the dangers of the Internet and its manipulations. It is based on the ontological premise of ‘hyper-individualism’ whereby persons, events, and phenomenon are treated as independent and ‘atomistic’ entities (Lake, 2017). Data politics that emerges from this reaction is one of urging people to protect themselves as individuals. It is almost as if the narrative says ‘yes, there is collective work that needs done but ultimately it is up to you to change your behaviour to protect yourself

from the dark forces of the Internet'. The addressee in other words is the atomized subject whose data is individualized rather than understood as a product of collective relations with other subjects and technologies (Socialising Big Data Project, 2015).

A third shortcoming concerns the immediacy that pervades these reactions or responses. They are predominantly exercised by the immediacy of a threat, danger, menace, risk, or peril or insecurity or unease that the Internet ostensibly engenders. Even those who have fought battles with governments and corporations to expose their data practices fall prey to a Messianic creep in articulating political problems by decrying their immediacy.

The obverse response to these reactions has been to extol the virtues of the Internet and illustrate that if it is not liberating it is at least making our lives better organized, measured, improved, whatever. Yes, there may be dangers and insecurities but this is a small price to pay for the benefits it brings. This response is still riddled with immediacy and atomism. Its calculative logic is from the point of view of the atomized subject weighing the pros and cons of the Internet against the threats of immediacy.

All this has led us to the conclusion that data politics is yet to find its subjects. In this commentary, we step back from the inertness, atomism, and immediacy of the dominant points of view of the Internet and the data it generates and ask questions about data politics within a broad historical perspective. What do we then mean by data politics?

We start with the assumption that the will to knowledge and the will to power are two aspects of how we conduct ourselves and the conduct of others and thus we approach data not as a representation (i.e., information collected, stored, and presented without interest) but as an object whose production interests those who exercise power. This was at least one of the lessons we have learned from Michel Foucault's relentlessly focused studies of the ways in which modern societies come to depend on governing subjects with data collected over not only their physical and social attributes (life, language, labour) but also about the conduct of their behaviour (Foucault, 2007). Our second assumption is that the production of data is a social and often political practice that mobilizes agents who are not only objects of data (about whom data is produced) but that they are also subjects of data (those whose engagement drives how data is produced). Our question thus shifts to social practices and agents. Data does not happen through unstructured social practices but through structured and structuring fields in and through which various agents and their interests generate forms of expertise, interpretation, concepts, and methods that collectively function as fields of power *and* knowledge.

This was at least one of the lessons we learned from Pierre Bourdieu's studies on the ways in which fields of knowledge constitute fields of power (Bourdieu, 1988) that involve struggle and change, fragile moments, and the emergence of new kinds of practices (Bigo, 2011).

Foucault and Bourdieu influenced a generation of scholars who have taken up the relations between power, knowledge, and fields and investigated the ways in which states, agencies, organizations, corporations, and institutions – often assembled in different combinations as governments – constituted their authority, legitimacy, and legality by producing knowledge about objects and subjects through establishing method and data regimes such as censuses, indexes, indicators, registers, rolls, catalogues, logs, and archives. We now understand much better the relationships between state formation and statistics, probability, and data regimes (Desrosières, 1998; Hacking, 1990; Porter, 1986). Statistics, from their very beginning, combined 'the norms of the scientific world with those of the modern, rational state' (Desrosières, 1998: 8). These data regimes have now been extensively studied as historical developments. The birth of objects of knowledge such as the economy, population, society, and their sciences – originally called political arithmetic and now statistics – have also been studied extensively. Although it would be impossible to summarize what we now know about these data regimes and the state, the overall insight we have gained can be stated as follows. While Max Weber's argument that the sovereignty of the state consists in its monopoly of the means of violence is often cited, following the studies of Foucault and Bourdieu and the literature inspired by them, we have come to recognize that this sovereignty depends on numerous practices beyond the organization of violence. Historically, the state performs sovereignty with control over and dependence on especially education, fiscal, and cultural data regimes. This does not mean that citizens in each state did not influence, interfere, or intervene in the ways which data regimes constituted them as data subjects. On the contrary, scholars have also investigated and documented how citizens have developed democratic practices to challenge social categories of data regimes and their effects (Anderson and Fienberg, 2000; Kertzer and Arel, 2002; Nobles, 2000). There are many cases that illustrate how, for example, census categories such as race, ethnicity, gender, and other indexes have been called into question, subverted, and transformed.

Nonetheless, the state, or rather organizations, institutions, agencies, agents, and authorities that make up the complex field of government, maintained an effective monopoly on data regimes concerning whole populations. This is not to say that corporations did not also generate data about their customers especially over the

last century or so but this was largely limited to specific population groups and in relation to narrow concerns. Beginning in the early 20th century, opinion polling and marketing research were considerable developments in corporate forms of population data generation (Osborne and Rose, 1999). And although there have been various international organizations that have entered into fields of data generation and accumulation such as the United Nations, the European Union, OECD, OSCE, and ILO, the primary site and scene of population data collection and its various regimes remained the monopoly of the state for nearly four centuries.

This monopoly of the state over data production, collection, and even interception is increasingly challenged. Or, at least, state sovereignty over data regimes is now shared by the birth of entirely new assemblages of the production of data (Kitchin, 2014). Not least has been the increasing accumulation and mobilization of data by corporations (Thrift, 2005). It is tempting to immediately single out the Internet and its connected devices as the source of this challenge. But it is much more complicated than that as our argument above anticipates. It would be folly to assume that Internet technologies develop independently from the interests that constitute the fields through which various data regimes have been invented. However, beyond technological developments, the sovereignty of the state in accumulating and producing data about its population, territory, health, wealth, and security is being challenged by corporations, agencies, authorities, and organizations that are producing myriad data about subjects whose interactions, transactions, and movements traverse borders of states in new and complicated patterns. Not least these traversals challenge the methodological nationalism that has dominated statistical thought and practice and their corresponding boundaries of population data, knowledge, and power for centuries (Scheel et al., 2016). While Bourdieu's studies focused on the nation and in particular France, others have taken up his conception to understand fields as international and transnational (Dezalay and Garth, 1996; Madsen, 2011, 2014). For Bigo, the transnational exists in the form of transnational networks and practices of professionals who 'play simultaneously in domestic and transnational fields' (Bigo, 2011). In this view, a transnational field is constituted by networks and practices between and amongst professionals who act at various non-hierarchically ordered scales of the transnational, national, and local.

Through this broad framing, data politics is a contribution to the tradition that intersects historical sociology and political sociology. We not only draw inspiration from major works such as Karl Polanyi's *The Great Transformation* (1944) and Norbert Elias's

The Civilising Process (1939) but also more recent work at the intersections of historical sociology and political sociology such as Christophe Charle's *The Birth of Intellectuals* (1990) and Paul Veyne's *Bread and Circuses* (1975). Despite their different subject matters what these works share is the importance of placing events in the present into historical series where we can discriminate between contingent and transformative events. We similarly aim to identify what is precisely transformative of the regimes of data accumulation that mark our present and which are generative of data politics.

What are the conditions of possibility of data? And relatedly, how are these conditions generative of data politics? In what follows we articulate this broad framing through three interrelated conditions of the making of data politics: worlds, subjects, rights. Our discussion is schematic and principally poses key questions that are not exhaustive of possible inquiries. Along with the foregoing they are meant to set out political questions about the ways in which data has been constituted as an object vested with certain powers, influence, and rationalities.

Worlds

The Internet is an elaborate infrastructure composed of objects, equipment, cables, routers, servers, switches, and devices that constitute a unique technological materiality. Unlike other massive material transformations of industrial and post-industrial cities and their transportation and communication infrastructures, the materiality of the Internet is mostly out of sight and located elsewhere. The data servers and data farms are often in faraway and remote locations or nestled within cities that are inaccessible and unknown to most people. Its connectors are often buried under the earth or sea. Its wireless communications are invisible but routers, switches, and masts create strange yet recognizable objects within and outside cities. Without this massive infrastructure and its maintenance and production, the internet of things, communications, and exchanges would be impossible. The material infrastructure of the Internet not only generates logics of borders and capacities of control that remain often invisible but also protocols and platforms that make people think the Internet is made up of a seamless and invisible flow of information. How are these worlds created and governed? What are the material conditions of possibility, configurations, and stratifications of these worlds? How do these worlds straddle or cross between actual and virtual spaces? To think of worlds is to trace how material conditions of the Internet are critical infrastructures that are generative of politics and struggles.

The Internet may depend on a massive infrastructure of servers, devices, and cables but what brings them together or more precisely what holds them together and enables them to communicate with each other is a special kind of language called code. But to understand code is anything but straightforward because code itself embodies various programming and communication languages such as binary machine code to algorithms (Galloway, 2006). The Internet has a language but it is hardly visible or even comprehensible to those who do not write such code. How does the language of the Internet traverse both actual and virtual worlds of data?

Through the Internet a new space is being made – a cyberspace perhaps – but understanding this space is fraught with difficulties. The Internet has not only blurred the boundaries between online and offline worlds but it has also rendered the distinction between the two spurious and perhaps untenable. With always-connected devices it is impossible to say when people or things are offline or online or indeed to separate embodied subjects from their operation. What kind of space does the Internet generate? What is the role of data in such a space and how does data make it possible? In turn, how does the Internet and the space it generates make data politics possible and with what effects?

The emergence of Big Data with its focus on production, accumulation, mining, circulation, aggregation, analysis, and interpretation has engendered the formation of various professions from data scientists to data journalists. Each of these professions is engaged in competitive struggles between each other and with other professions and yet at the same time also reinforce the broader practice of investing data with powers. These emergent professions and their practices have not only begun reorganizing existing fields of data production such as the official statistics of states (state and statistics share common etymologies) but also have given birth to new forms of data accumulation and valuation whose source of authority and legitimacy traverse the boundaries of state sovereignties and produce international effects.

In this light, data is not an already given artefact that exists (which then needs to be mined, analysed, brokered) but an object of investment (in the broadest sense) that is produced by the competitive struggles of professionals who claim stakes in its meaning and functioning. They engage in struggles over the valuation of different forms of capital conceived by Bourdieu including cultural, economic, social, and symbolic capital (Bigo, 2013). It is through the accumulation of these various forms of capital that their relative positions are established within the field (Bourdieu and Wacquant, 2006). The emergence of data as a field and data professionals as its custodians and gatekeepers shapes competitive struggles not only in

defining an object but also the principles of how to understand and intervene in data politics. At the same time, algorithms increasingly call into question the very expertise that data accumulation has spawned through the automating practices of judgement. Who decides whether to invest, what to listen to, where to eat, where to stay, and where to go? How do algorithms embed expert judgements and normative assumptions without appearing to do so?

The accumulation of data procures not only cultural capital but also economic capital. An economy of data is founded on the ‘voluntary’ input of personal data in exchange for Internet services. This creates the conditions for the making of a stock market of data involving data brokers and profit shares generated by deep data mining and data discoveries. How do individuals contribute to this production and what is the political economy of desire that generates a material economy of services? What are the consequences of subjects giving up data in return for so-called free services? What are the legal conditions that enable and disable the circulation of data within and across states? From questions of data commons to data ownership, how are legal regimes being challenged and remade by struggles over data as property?

Subjects

The emergence of data as an object of government engenders the emergence of subjects who take positions in and through the various resignifications and challenges that it spawns. Rather than occupying already existing positions, subjects are produced through various digital interactions and at the same time their digital traces shape and organize their subjectivities and how they are known and governed. How are subjects part of the work and making of data through which they then come to be known? Through procedures of channelling, filtering, and sorting data, various devices and platforms configure not only transactions and interactions but the data they generate recursively shapes and forms subjects in never fixed but modulating ways. With the increasing circulation, mining, and combining of data, how are subjects and their affiliations, connections, and relations multiplied and governed via ever more dispersed micro data politics?

Data not only captures but also colonizes minds, souls, bodies, and spaces. It subjectifies through practices of production, accumulation, aggregation, circulation, valuation, and interpretation. These practices call upon subjects who are not separate from but submit to and are active in the various ways that data is made and colonizes lifeworlds to constitute ‘data’s empire’ (Isin and Ruppert, forthcoming). People govern their health by making themselves data subjects of health.

Measuring their own performances with Internet-enabled devices and benchmarking their performance against other performers, data subjects of health increasingly calibrate a model body not through images circulated by the advertising industry but by literally working themselves out through their data performances and for others. How is data part of the making and shaping of bodies and the body a site of data politics?

Being a data subject entails the radically shifting meaning of being a consumer from a subject making choices to a choice-making and sorted subject. Being constantly a reviewer, modern data subjects are caught in a spiral of evaluations: they are evaluated and evaluator all at once and all the time. Recommender platforms and evaluation data generated by transacting ever more sort subjects into categories of cultural preferences that narrow and channel choices. How is consuming through the Internet generative of data politics?

Are we now witnessing the changing relations of production in the generation of data and if so, then who are its labourers? Are we moving from the logic of having a job to a logic of contributing something to the fulfilment of a task? The data-generated market of global tasks has now created a vast meeting place for those who need and will pay for accomplishing specific and often micro tasks and those who can and need to fulfil these tasks to make a living.

To consider the data subject also calls upon consideration of the uncanny convergence between robots and humans not in the way in which the cyborg manifesto (Haraway, 1991) envisaged it but perhaps more in the manner in which Star Trek anticipated. How does the automated generation and analysis of data based on artificial intelligence and machine learning appear autonomous and yet inseparable from struggles and relations between programmers, subjects, and technologies?

Rights

If the accumulation of data traverses subjects it also constitutes them with claims to certain rights that concern its accumulation: who owns, distributes, sells, accesses, uses, appropriates, modifies, and resignifies data become objects of struggles for claiming rights to such modalities. The rights-claiming subject is the figure of the citizen that we have inherited as a political subject who is now making rights claims about being a subject of data. How do subjects exercise and claim such rights through what they say and do through the Internet? How do they perform rights and claims about being subjects of data through how they communicate, share, express, and engage with digital devices and platforms? How do they invent data practices that challenge and subvert state and corporate forms of data and struggle for rights through legal and regulatory mechanisms?

This third condition of data politics considers rights-claiming subjects such as citizen data scientists as part of material–political arrangements and struggles over who generates, legitimizes, and has authority over data and how data is mobilized to make claims for environmental and other rights. It concerns how citizens make data an object of transnational politics and engage in struggles around free expression; privacy and ethics; and the forums, practices, and networks through which these struggles are being fought.

The relationship between the right to privacy and that of data protection is illustrative of the transversal relations and legal and political tensions that make up data politics. On the one hand, international human rights laws and obligations seek to secure and universalize the former and various national regimes have emerged to address the latter. How are rights not only claimed through national and international institutions and their regulations, laws, and protocols but by citizen subjects who make claims and in turn perform what is data politics through their everyday digital acts? As such, how does this call for a figure of a citizen that is different from the subject we have inherited and instead one who can make rights claims that traverse national borders? (Isin and Ruppert, 2015).

Conclusions

We have provided an outline of data politics as a field of power and knowledge that traverses virtual and actual worlds, produces subjects that come into being through such worlds, and whose rights become objects of struggles through such a field. We have specified worlds, subjects, and rights as the conditions of possibility of such a field of data politics. We want to argue that without understanding these conditions of possibility – of worlds, subjects, and rights – and their relations, it would be difficult to intervene in or shape data politics if by that it is meant the transformation of ourselves from data subjects into data citizens.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Funding from a European Research Council (ERC) Consolidator Grant (615588) supported the writing of this commentary. Principal Investigator, Evelyn Ruppert, Goldsmiths, University of London.

References

- Anderson M and Fienberg SE (2000) Race and ethnicity and the controversy over the US census. *Current Sociology* 48(3): 87–110.
- Bigo D (2011) Pierre Bourdieu and international relations: Power of practices, practices of power. *International Political Sociology* 5(3): 225–58.
- Bigo D (2013) The transnational field of computerised exchange of information in police matters and its European guilds. In: Kauppi N and Madsen MR (eds) *Transnational Power Elites: The Social and Global Structuration of the EU*. London: Routledge, pp. 155–182.
- Bourdieu P (1988) Social space and symbolic power. *Sociological Theory* 7(1): 14–25.
- Bourdieu P (1993) *Language and Symbolic Power*. Cambridge, MA: Harvard University Press.
- Bourdieu P and Wacquant L (2006) *Reflexive Anthropologie*. Frankfurt am Main: Suhrkamp.
- Butler J (1997) *Excitable Speech: A Politics of the Performative*. London: Routledge.
- Desrosières A (1998) *The Politics of Large Numbers: A History of Statistical Reasoning*, trans. Camille Naish. Cambridge, MA/London: Harvard University Press.
- Dezalay Y and Garth BG (1996) *Dealing in Virtue: International Commercial Arbitration and the Construction of a Transnational Legal Order*. Chicago, IL: University of Chicago Press.
- Foucault M (2007) *Security, Territory, Population*, ed. Arnold Davidson, trans. Graham Burchell. Lectures at the Collège De France, 1977–78. Basingstoke: Palgrave Macmillan.
- Galloway AR (2006) *Protocol: How Control Exists After Decentralization*. Cambridge: MIT Press.
- Hacking I (1990) *The Taming of Chance*. New York: Cambridge University Press.
- Haraway DJ (1991) *Simians, Cyborgs and Women: The Reinvention of Nature*. London: Free Association.
- Isin E and Ruppert E (2015) *Being Digital Citizens*. London: Rowman & Littlefield International.
- Isin E and Ruppert E (forthcoming) Data's Empire. In: Bigo D, Isin E and Ruppert E (eds) *Data Politics: Worlds, Subjects, Rights*. New York and London: Routledge.
- Kertzer DI and Arel D (2002) Censuses, identity formation, and the struggle for political power. In: Kertzer DI and Arel D (eds) *Census and Identity: The Politics of Race, Ethnicity, and Language in National Censuses*. Cambridge/New York: Cambridge University Press, pp. 1–42.
- Kitchin R (2014) *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences*. London: SAGE.
- Lake R (2017) Big Data, urban governance, and the ontological politics of hyper-individualism. *Big Data & Society* 4(1): 1–10.
- Madsen MR (2011) Reflexivity and the construction of the international object: The case of human rights. *International Political Sociology* 5(3): 259–275.
- Madsen MR (2014) The international judiciary as transnational power elite. *International Political Sociology* 8(3): 332–334.
- Nobles M (2000) *Shades of Citizenship: Race and the Census in Modern Politics*. Stanford, CA: Stanford University Press.
- Osborne T and Rose N (1999) Do the social sciences create phenomena?: The example of public opinion research. *British Journal of Sociology* 50(3): 367–396.
- Porter TM (1986) *The Rise of Statistical Thinking, 1820–1900*. Princeton, NJ: Princeton University Press.
- Scheel S, Cakici B, Grommé F, et al. (2016) Transcending methodological nationalism through transversal methods? On the stakes and challenges of collaboration. ARITHMUS Working Paper. Available at: http://arithmus.eu/?page_id=13 (accessed 14 January 2017).
- Socialising Big Data Project (2015) A social framework for Big Data (CRESC). Available at: <http://bit.ly/28Ye3kI> (accessed 2 December 2016).
- Thrift N (2005) *Knowing Capitalism*. London: Sage.

This commentary is a part of special theme on Data Politics. To see a full list of all articles in this special theme, please click here: <http://journals.sagepub.com/page/bds/collections/data-politics>.