



In the name of Development: Power, profit and the datafication of the global South



Linnet Taylor^{a,*}, Dennis Broeders^b

^a University of Amsterdam, The Netherlands

^b Erasmus University Rotterdam, The Netherlands

ARTICLE INFO

Article history:

Received 15 October 2014

Received in revised form 2 July 2015

Accepted 3 July 2015

Available online 9 July 2015

Keywords:

Big data

Power

Development

Public private partnerships

Surveillance

Datafication

ABSTRACT

We examine the current ‘datafication’ process underway in low- and middle-income countries (LMICs), and the power shifts it is creating in the field of international development. The use of new communications and database technologies in LMICs is generating ‘big data’ (for example from the use of mobile phones, mobile-based financial services and the internet) which is collected and processed by corporations. When shared, these data are also becoming a potentially valuable resource for development research and policy. With these new sources of data, new power structures are emerging within the field of development. We identify two trends in particular, illustrating them with examples: first, the empowerment of public–private partnerships around datafication in LMICs and the consequently growing agency of corporations as development actors. Second, the way commercially generated big data is becoming the foundation for country-level ‘data doubles’, i.e. digital representations of social phenomena and/or territories that are created in parallel with, and sometimes in lieu of, national data and statistics. We explore the resulting shift from legibility (Scott, 1998) to visibility, and the implications of seeing development interventions as a byproduct of larger-scale processes of informational capitalism.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

There is a process of ‘datafication’ (Mayer-Schönberger and Cukier, 2013) underway in low- and middle-income countries (LMICs),¹ where the use of new communications and database technologies in LMICs is generating digital data that is machine-readable and computationally manipulable, particularly for ‘big data’ analytics. These born-digital datasets are of unprecedented size and detail, especially compared to the statistical records previously available on lower-income countries (Jerven, 2013). In contrast to traditional state survey data, however, these data are generated, collected and processed under the auspices of private-sector corporations and are shared, often on a pro-bono basis, at the level of international academic research institutions or development actors such as the UN. Where previously development donors (governments or international NGOs) worked with LMICs’ own statistical apparatuses

to generate population data, it is becoming increasingly possible and cost-efficient for donors to turn to corporations for consumer-generated data that can proxy for traditional household surveys and other statistical products (Taylor and Schroeder, 2014). The discourse on big data as a resource for development (World Economic Forum, 2012; Global Pulse, 2012; Taylor and Schroeder, 2014) indicates that a shift is underway from the predominance of state-collected data as a way of defining identities and sorting and categorising individuals, groups and whole societies to a big-data model where data is primarily collected and processed by corporations and only secondarily accessed by governmental authorities.

The central question addressed in this paper is how datafication is influencing the way that LMIC populations are made legible in the context of development, and what this means for power dynamics amongst development actors. We examine a power shift from the traditional collector and user of statistics – the state – to a messier, more distributed landscape of governance where power accrues to those who hold the most data. This power shift has its roots in the larger neoliberal trend in governance worldwide – what Cohen (2013: 1928) has termed the ‘new governance,’ dominated by public–private partnerships engaged in ‘informational capitalism’ (2013: 1912, following Castells (1996)), a system where ‘information flows in circuits that serve the interests of powerful

* Corresponding author.

E-mail address: l.e.m.taylor@uva.nl (L. Taylor).

¹ We use the World Bank’s definitions grouping countries, see: <http://data.worldbank.org/about/country-classifications>, where LMICs have incomes of US\$1,036 – \$12,616 per capita and high income countries (HICS) above that threshold. Our particular focus is the low- and lower-middle-income countries, with an upper threshold of \$4,085 per capita, which includes India and most of Africa.

entities, both private and public' (2013: 1916) and where those who trade in information – primarily corporations – are able to modulate people's behaviour, activities, and relationship to the state. She argues that the creation of surveillance infrastructures are an inevitable part of governance through informationalism, and the possibility of surveillance is therefore pervasive. We argue that 'data-driven development' is characterised by the same features as informational capitalism, and results in increased visibility for the populations of lower-income countries – though not necessarily in greater governability or representation.

This shift towards a combination of datafication and privatisation is still in its early stages and the evidence is not yet available to draw conclusions about its medium or longer-term impacts. It is potentially wide-ranging, however, and has powerful implications for how LMICs will be understood by the development field as economic and political actors in the coming years. So far, little research has been conducted on the political implications of the production of big data about lower-income countries (exceptions include [Thomas \(2014\)](#), [Burns \(2014\)](#) and [Taylor and Schroeder \(2014\)](#)). We therefore assess the longer-term risks of the secondary use of such data under the rubric of economic or human development, based on illustrative examples of current data collection by corporations and public–private partnerships in LMICs and aim to contribute towards a research agenda on this issue by drawing out and illustrating key trends and their potential repercussions.

We define two key trends using illustrative examples: first, the increasing empowerment of private sector actors in the field of international development due to their ownership of data, and second, the emergence of development/surveillance assemblages with the potential for ongoing monitoring of population dynamics and people's activities. The first trend involves leverage and opportunities for corporate actors who are proprietors of vast quantities of data on citizens of developing countries, whilst states continue to prioritise traditional survey data such as censuses which convey a different type of detail and allow for different types of sorting and monitoring. We look at how the deputisation of technology corporations by states through public–private partnerships (PPPs) may lead to the merging of these different analytical perspectives. We then frame the second trend, the emergence of country data doubles and various forms of shadow mapping – in relation to these partnerships.

This paper is based on 60 interviews on the use of big data in development policy and planning, along with two years' research (2013–2015) on the interface of data science and international development policy. The research also involved attending international conferences, workshops and public discussions hosted by large institutional actors, and by gathering knowledge from mailing lists and other online discussions. This project was partly funded by the Alfred P. Sloan Foundation at the Oxford Internet Institute. The interviews on big data in the field of international development were conducted with academic researchers and private-sector data scientists working with big data on questions relevant to LMICs. Interviewees were selected using a purposive sampling process focused both on the most relevant projects, and on those with an overview of the state of the art in big data research.

2. Power through LMIC data

2.1. Legibility, visibility and emergent power structures

[Scott's \(1998\)](#) famous book *Seeing Like a State* could instead have been titled 'reading like a state' since the registrations upon which the modern bureaucratic state was built, made society – in Scott's own words – 'legible'. 'Reading like a state' in the era of

big data, however, increasingly involves remotely performing data analytics to make populations visible. This constitutes a step beyond – or away from – Scott's formulation because such visibility creates power over data subjects via volume of data rather than accuracy and detail. Unlike legibility, which depends on data that people are aware of providing, the data that provides visibility is often observed, not volunteered ([Hildebrandt, 2013](#)), derived as a byproduct of technology use rather than collected by authorities through survey methods. Unlike the data that provides legibility, these data are of unknown reliability because they reflect, in [Shearmur's words \(2015\)](#), not populations but 'users and markets' and are therefore biased towards those with connectivity. This unknown bias is compounded because the visibility such data offers is created by data scientists, not social scientists, and frequently lacks contextual information to explain what is being seen ([Taylor and Schroeder, 2014](#)). The gap in terms of distance – but also cultural dissimilarity – between the processors and end users of LMIC big data and those living in these countries may also increase power asymmetries. This suggests that the power to make visible is different from the power to make legible. Legibility increases governability (in Scott's formulation), but visibility offers the power to influence and intervene to a wider, more distributed set of actors: the corporations who gather and analyse the data, plus whoever they choose to share it with (or can capture it through other means), who may be state actors, international development institutions, or other corporate partners.

Traditional visualisations such as mapping are being transformed by spatial big data analytics. GPS-derived big data from mobile phones in particular lends itself to real-time, highly detailed mapping of mobilities ([Taylor, 2015a](#)), and sensor data offer new possibilities for understanding urban space in particular ([Kitchin, 2014](#)). However, this big data is mainly created and accessed within government–corporate networks ([Crampton et al., 2014](#)), so that big data is a tool of the powerful in the corporate and governmental spheres who can afford to collect, merge and analyse it. In this way, little has changed since [Harley's assertion \(1989\)](#) that maps are essentially political despite the claim of scientific objectivity put forward by the profession of cartographers. It also reinforces and broadens [Taylor and Johnston's \(1995: 58\)](#) warning to geographers that by using digital data generated by state apparatuses, they are 'creating the state's geography' and determining 'what is included and what is not'. Today, a similarly critical approach is needed in relation to corporations that collect and own data, and determine how they may be made accessible. What kind of geography is emerging though the corporate-led Data for Development?

Power is thus exerted through the generation of data, control of access to data, the merging of various data, but also through their analysis and interpretation. At each of these points, however, the private sector's role is increasingly important but also increasingly hard to pin down. [Bowker and Star \(2000\)](#) in their seminal work on classification remarked that information architectures (and the politics that they contain) are hard to analyse because 'Good usable systems disappear almost by definition. The easier they are to use, the harder they are to see. As well, most of the time, the bigger they are, the harder they are to see' ([Bowker and Star, 2000: 33](#)). At this moment, big data analytics constitute the biggest usable system around and the corporate and governmental enthusiasm for its potential does not motivate many to analyse what lies under the hood. Similarly, [Schneier \(2015: 30\)](#) points out that data collection in HICs has become so engrained in modern life that it fades into the background for the individual user. Moreover, most of the attention that has been paid to projects under the rubric of Data for Development focuses on the perceived and expected benefits at the expense of the epistemological and power implications of the data analytics used.

For these reasons, data analytics are becoming the new battleground of representation. Ultimately, all data lead to actionable knowledge (Degli Espositi, 2014: 212), or in the words of Amoores (2011: 27): ‘the processes of data integration, mining and analytics draw into association an amalgam of disaggregated data, inferring across the gaps to derive a lively and alert new form of data derivative – a flag, map or score that will go on to live and act in the world’. Yet data and representation ultimately serve their ‘owners’ and are at least in part built on power asymmetries. Graham (2009) writing about ‘virtual earths’ and web 2.0 applications within webglobs such as Google Earth and Microsoft Virtual Earth contends that these are not simply mirrors of their physical counterparts but are characterised by ‘information black holes’ as well as ‘hubs of rich description and detail’. He warns that: ‘Although myriad representations of place are now easily accessible, it remains important to note that the virtual earth remains highly shaped by dominant power structures, software algorithms and the cultural links between producers of information’ (Graham, 2009: 432). Being seen and not being seen, looking and not looking are in his view part of a power structure that is almost always asymmetrical. Visual equality is a rare thing, especially when states are involved. (In)visibility is therefore a power struggle: ‘Invisibility is a relationship between those who have the power to see or to choose not to see, and, on the other hand, those who lack the power to demand to be seen or to protect themselves from the negative effects of imposed visibility’ (Polzer and Hammond, 2008: 421).

2.2. Distributed governance: PPPs for data and development

If having the data to see what was previously invisible, or ill-defined, is a source of power, then this power seems to be shifting to a new group of actors. Worldwide, big data is largely proprietary, being generated and analysed by technology-providing corporations. Now that data analytics are being advocated as a new solution to development problems, those with the data will also increasingly gain power over how development interventions are conceptualised and practiced. As corporations acquire the power to make data subjects visible, they also acquire the power to monitor (or to ‘aid surveillance’, as Privacy International has put it (Hosein and Nyst, 2013)). Lyon (2007) argues that surveillance is always somewhere of the continuum between ‘care’ and ‘control’, with the latter more closely associated with the (negative) political label. In the world of Development, monitoring through data is constant and widespread, and the power to see is also the power to influence: Lyon (2007) defines surveillance as ‘(...) the focused, systematic and routine attention to personal details for purposes of influence, management, protection or direction’. Organising aid, development interventions or even welfare states require data for monitoring inclusion, progress and impact – but in the development field, where there is an accountability gap with regard to international interventions, it is important to understand how data may also enable the abuse of power.

Jerven (2013) shows that statistical apparatuses in lower-income countries are widely judged to be inadequate and under-resourced, a problem which has recently led researchers and policymakers to identify ways in which new sources of digital data, arising from the use of digital technologies by people in LMICs, may provide a way to supplement traditional survey-based methods of data-gathering such as censuses (e.g. Center for Global Development, 2014; Green, 2014; Global Pulse, 2013). This high-level institutional interest in ‘data-driven development’ (World Economic Forum, 2012) has been accompanied by the notion of ‘data philanthropy’ (Kirpatrick, 2011), where the corporations providing digital services to populations of LMICs give development authorities access to that data under the rubric of Corporate Social Responsibility (CSR). This process

involves high-level intermediaries, notably UN Global Pulse, one of the highest-profile nodes in an informal network of international researchers and development organisations dealing with data-driven development projects, publishing and discussing results, and sharing the knowledge gained from interventions. Examples of past projects within this network include the 2013 Data for Development challenge established by Orange (2012), the donation of mobile phone data to epidemiologists to track the cholera outbreak after the 2010 Haiti earthquake (Bengtsson et al., 2011), and UN Global Pulse’s use of social media to track food price shocks in Indonesia (Global Pulse, 2014). Each of these research projects has been conducted with ‘big data’ donated free of charge by large multinational corporations such as Orange and Twitter for purposes of monitoring and evaluating social or economic dynamics in LMICs.

The discourse on big data as a resource for development (World Economic Forum, 2012; Global Pulse, 2012; Taylor and Schroeder, 2014) highlights that data is primarily collected and processed by corporations and only secondarily accessed by governmental authorities. This shift has two effects: first, it translates the individual from citizen to data subject, ‘a conditional form of existence whose rights are dependent upon its behaviour within digital networks’ (Bauman et al., 2014: 129). Second, it underlines the move to a more distributed governance model with regard to population data, ‘organized neither horizontally, in the manner of an internationalized array of more or less self-determining and territorialized states, nor vertically in the manner of a hierarchy of higher and lower authorities.’ (ibid: 124). Instead, those who hold the data increasingly have the power to intervene, or to inform interventions.

Decentralised or distributed governance may come about due to conditions of limited statehood (Livingstone and Walter-Drop, 2014), but may also arise through the involvement of international development institutions, including private-sector actors. The legacy of the household survey sponsored by large development institutions (e.g. USAID’s Demographic and Health Surveys; the InDepth Network’s health and demographic surveillance system, and UNICEF’s Multiple Indicator Cluster Surveys), along with a myriad country-level examples run by Non-Governmental Organisations (NGOs), have created a landscape where data flows in complex circuits which sometimes intersect with the host state’s own, and sometimes do not. Thus foreign institutions, non-governmental organisations and public–private partnerships may all play a role in collecting data to make populations visible, and in creating and monitoring interventions in social and economic life.

Another repercussion of this shift is that the power to profile is in new hands. According to Lyon (2014: 6–9) the use of big data results in the automation of analysis, which gives those who develop algorithms for sorting and classifying data subjects significant power over those people’s lives – more so if the analysis is being used to direct development interventions. In the era of digital communications and big data this distributed governance of the means of sorting and monitoring populations is translated into unrestrained digital data maximisation and, in turn, to what Haggerty and Ericson (2000) have termed ‘data doubles’ – abstracted representations of people through their data, formed with the aim of monitoring or targeting people for intervention. Where corporations are responsible for profiling, the data doubles that are created are a source of power in themselves. Ruppert (2012) has updated this concept for the era of big data, noting that data doubles are now constructed by a plethora of different data processors, and Lyon (2014) adds that big data analytics give the data double more power (and its subject greater visibility) by making it ubiquitous and constantly self-generating. In line with these contributions we scale up the notion of the data double, showing how it now extends to the collective level of cities and countries.

Where corporations conduct or facilitate development research on LMICs, it is worth examining the risks of repurposing data analytic techniques developed for purposes of profit-making. As Lyon notes, ‘The enthusiasm for big data “solutions” may lead to inappropriate transfer of techniques from one field to another’ (Lyon, 2014: 6).

Given all these factors, we therefore draw attention to a shift in power relating to LMIC development: namely, that the power to count, categorise and visualise those in LMICs is increasingly being acquired by corporations, particularly through the use of new communications technologies by LMIC citizens. This shift is towards practices of informational capitalism, which has as its byproduct a new level of visibility for previously less-surveyed populations, or at least those in contact with digital technologies. We posit that this shift extends unprecedented power to monitor, and therefore also to surveil, to new and unregulated actors; that it gives corporations a new type of influence over the lives of individuals with potential for misuse, and that it may exclude smaller-scale actors with local knowledge and understanding. In what follows, we offer illustrations of this shift and draw out the implications for its longer-term repercussions.

3. Two trends in datafication and development

In this section we explore two trends which illustrate the issues identified above. We first show how corporations’ technical capacity leads to their deputisation by LMIC states and foreign donors through the institutional mechanism of Public–Private Partnerships (PPPs), and how such projects are using the development discourses that underpin these partnerships to justify the search for new markets. Secondly, we examine how private big data are becoming the foundation for national-scale data doubles and shadow maps, i.e. digital representations of social phenomena and/or territories that are created in parallel to, and sometimes in lieu of, national data and statistics. These PPPs are thus leading to new data doubles and mapping systems which cannot be separated from the power structures and political economy that create them, and that they reproduce through their manipulation of data.

3.1. Corporations as development actors, development as market expansion

As corporations expand into emerging markets through services which generate digital data, they now find themselves simultaneously expanding into the development field. Although the private sector has long played a role in development through Corporate Social Responsibility (CSR) activities, corporations are becoming involved in the field in new ways as digital data emerges as a resource for analysing development and informing policy. This is taking place in two different configurations: first, with corporations as partners or deputies of states or donors where the latter are still developing digital capacity, a configuration which reflects the distributed modes of governance facilitated by the new data flows. Second, there are projects which carry the label of development but which are managed and rolled out entirely by corporate interests. This section draws on three different projects to examine how discourses of development are mingling with those of profit as corporations increasingly move into the space previously occupied by states and nonprofit donor institutions.

An example of the first configuration can be seen in India’s biometric identification project, the Unique Identification Authority of India (UIDAI, also known as Aadhaar), the largest-scale public–private partnership currently underway in a developing country in terms of its coverage of the population. This scheme, in operation since 2008, forms the world’s largest biometric database and aims to provide every Indian resident with a 12-digit Unique Identity

number (UIDAI.gov.in, 2014). By mid-2013, 400 million people’s details had been entered (Business Standard News, 2013). The public–private structure of the scheme was inevitable due to the amount of technical expertise necessary to deal with the big data analytics that are essential to real-time crosschecking and updating of a database of hundreds of millions of records (Krishnan, 2012). Besides acting as identification for welfare recipients, a UID is also the gateway to a bank account, a mobile phone contract, and various commercial products which corporations previously were unable to offer to lower-income citizens partly because of a lack of reliable identification. The project, however, is only dubiously ‘public’. Officially UIDAI is a quasi-governmental organisation attached to the national planning authority, but effectively it is an autonomous organisation, operating without a legal framework or parliamentary oversight (Ramanathan, 2013). Besides being non-governmental, UIDAI is also a collaboration between and amongst corporations. It is managed and chaired by Nandan Nilekani, a co-founder of InfoSys, one of India’s largest technology and consulting firms, and the day-to-day work of gathering, processing and storing data is done by private companies (India Today, 2013; Yadav, 2013). The collaboration bears out Mirafab’s claim (2004) that corporate involvement in the Development project is conceptualised as a way of importing technical and managerial capacity to under-resourced countries; but as he points out, since the oversight role ascribed to government authorities in those countries is often only sketchily established, corporations often remain essentially autonomous actors and these arrangements may come to underpin distributed governance structures.

The UIDAI project is extremely broad in its aims: it is presented by Nilekani as facilitating development and poverty alleviation; ‘a huge project of social inclusion ... [that will] improve the quality of government spending on various public programs’, and a way of becoming visible to the ‘welfare state’ (Nilekani, 2013). He also states, however, that the project could be a blueprint for ‘a next-generation healthcare record system for developing countries, or you could use it for massive online education, or whatever.’ (ibid). The other functions constituted by ‘whatever’ range from giving marketers direct access to potential consumers to creating linkable data on individuals’ location, consumption and other characteristics. These are hugely commercially powerful, and possibly eclipse the ID’s role as a connection to the state’s welfare apparatus. As Nilekani describes it, the database is designed to be a complement to state records, and can therefore enable a variety of consumer functions as much as operations of citizenship. Nilekani’s speech to the Center for Global Development (ibid) also includes a vision of the UID scheme as a form of instant direct marketing on an unprecedented scale: ‘The network effect also comes into play. Let’s say that we have 300 million people who have this ID, and let’s say that somebody launches a product for a new kind of an annuity product. With the help of our system, all 300 million people can instantly get access to the product. So, it really creates this huge network effect where you can quickly get people to ramp up and get services.’ (Nilekani, 2013). Nilekani has subsequently suggested that the UID database will be opened up to app developers in order to speed the invention of ways to connect corporations directly to customers (LiveMint, 2013). These are the features which have led Jayaram (2014) to refer to the scheme as a ‘welfare industrial complex’ with ‘implications for how citizens relate to private sector entities, on which the UID rests and which have their own vested interests in the data’.

As Nilekani’s speech points out, the UID was never designed as just a public development project, but has a commercial rationale at its core. This combination of functions both extends the UID project into the realm of governance and raises serious questions about privacy (Jayaram, 2014). This both aligns with Cohen’s vision of informational capitalism (Cohen, 2013), and is also an example

of data gathering moving from objectives of care towards a reality of control (Lyon, 2007). As Dreze (2010) has pointed out, the linking of the UID to the National Population Register is also potentially a way of relating individual biometric ID numbers to tax returns, bank records, SIM (subscriber identity module) registers and the Indian Home Ministry's National Intelligence Grid (NATGRID), which in turn links to 21 national databases. In fact, Dreze's predictions have come true as domestic terrorism has become more of an issue in India over the six years of the UID's existence. The scheme's security implications have become an important selling point – demonstrating how population-level information systems easily become subject to function creep – the biometric database has been planned to be compatible with 'national security and border control applications' (Zelazny, 2012).

Alternatively, corporations may choose autonomously to carry out work they define as development-related, rather than aiming to fill a gap left by state capacity. This can be seen occurring in important emerging markets such as East Africa, as with the recent rollout of IBM's Project Lucy in Kenya. IBM promises to 'solve Africa's grand challenges' including 'healthcare, education, water and sanitation, human mobility and agriculture' (IBM, 2014) using artificial intelligence and big data analytics. The project will take place within a private–public consortium structure led and managed by IBM, including 'government leaders, university professors, business leaders, heads of nonprofits and people involved in the burgeoning tech startup communities' (IBM, 2014: 4). It will involve feeding all the published economic and social data available from Sub-Saharan African countries into IBM's 'Watson' supercomputer, which will then data-mine for answers to questions. (So far the company is not responding to researchers' requests for more information about Project Lucy, and the materials available do not specify whether the project will also have access to data which is not in the public domain.) Despite the promise to address problems which have their roots in poverty, the lack of a strong civil society, inequality and inadequate governance, the leader of IBM's Project Lucy is clear that the standard for success is profitability, saying that 'We're creating a strong business-oriented culture in the lab... I want my people to constantly ask themselves, "Who's going to buy it?"' (IBM, 2014: 7). Similarly, the chief data scientist for the project again blends the discourses of development and corporate profit when he warns that 'You have to be aware of the price point (for Project Lucy's findings). You have to be aware of cultural issues, of education. You have to make technology easily consumable.' (IBM, 2014: 4). Project Lucy epitomises how corporations' power over data translates into power to determine what constitutes development. Any organisation that can pay may ask questions of the Lucy database, but the structure and content of that database has been determined by IBM – a company whose motives in the region are primarily profit and market share.

Perhaps the most high-profile project in the corporate/development marketplace currently is that of Facebook's internet.org project, which promises to offer the internet to those previously excluded by poverty. As of 2013 69 per cent of the world's mobile subscribers had only 2G service (GSMA, 2013b), suggesting that brand recognition and subscribers captured whilst connections are slow will yield an important market advantage as they speed up over time. Facebook's internet.org aims to capture this market by extending internet connectivity, primarily via mobile phones, 'to the two thirds of the world's population that doesn't have it' (internet.org, 2014). The project will use 'drones, satellites and lasers to deliver the internet to everyone'.² In combination with

providing connectivity, Facebook also promises mobile subscribers in LMICs free data service on their phones when they access a text-only version of Facebook and selected 'partner sites',³ although for other sites users will have to pay data charges. Facebook markets this 'Facebook Zero' service as a way to enable low-income people in areas with low levels of technological infrastructure to go online in a simpler internet environment.

However, despite Facebook's rhetoric about the benefits of 'connecting the world' to the internet (see note 3), they are instead connecting a subset of the world. The Facebook Zero service is being marketed to the most populous emerging economies, where technology corporations are looking to build market share: initially Indonesia, the Philippines, and Malaysia, and soon thereafter Nigeria and South Africa (MTN, 2010). In terms of the global South, then, the project focuses mainly on those countries on the verge of connecting themselves, where awareness of the internet is high, and desire for online social networking so powerful that the mobile phone becomes just a vehicle for a particular application. Therefore reducing the asymmetries in connectivity between more profitable markets and remoter, lower-income areas will not be an outcome of the current strategy, although the company's rhetoric strongly implies it. Instead the project is likely to reduce the asymmetry in the use of Facebook. An activist in India interviewed for this project said of Facebook and other social networks:

That's being used to sell phones now. I mean, I remember doing this at one point, just taking lots of photographs of hoardings that were advertising phones, and they're not advertising the phone, the phone is a very small part in the corner of the billboard, it's Facebook or Twitter which is... you know, it's connected to Find Your Friends, that's how the phone is being sold now.

[(Digital rights activist, India, interviewed 12.11.2013)]

Providing internet connectivity based on access to Facebook, then, is likely to convince first-time users that they are accessing the most important services available, i.e. social networking. In doing so, the project reduces the internet to what the Electronic Frontier Foundation (EFF, 2015) has called a 'walled garden' of tethered partner services which restricts the possibility for creative use of the web, and for making it impossible for users to use secure connections. In the words of the Electronic Frontier Foundation, 'Internet.org is not neutral, not secure, and not the internet' (EFF, 2015).

Facebook's strategy is logical in a saturated market: the next billion people online will be in the countries internet.org focuses on. The project, however, does more than capture new markets. It underlines how corporate projects publicised as promoting development and inclusion may still reinforce existing power and knowledge asymmetries, and it highlights how the data corporations gather from developing countries is likely to be skewed by particular business strategies, with consequences for whether and how donor institutions should align with these corporate actors. The project will also have a less obvious but possibly more insidious effect, however. Facebook's connection to millions of previously unreachable subscribers will result in new datasets on their characteristics and behaviour, and consequently new types of visibility for those people. If the power to visualise is also the power to influence and intervene, Facebook's project empowers Facebook first, but next its strategic allies such as app developers and advertisers, creating new axes of visibility and power as a byproduct of the aim to create new digital consumers. Informational capitalism, in this case, will create both visibilities and the power to represent. In the process it will also, by capturing

² Statement via Facebook by Mark Zuckerberg, 27.3.2014: <https://www.facebook.com/zuck/posts/10101322049893211>.

³ Statement via Facebook by Sid Murlidhar, 18.5.2010: <https://www.facebook.com/notes/facebook/fast-and-free-facebook-mobile-access-with-Ofacebookcom/391295167130>.

data doubles in the global South, distribute the power to govern to anyone who can afford it.

The examples above suggest that corporations' data-gathering and analytical advantage means they no longer have to conceal for-profit motives under development discourse, as was the case with CSR projects, but instead are now free to connect the expansion of markets, and the attendant expansion of digital records, with human development. This perspective is being answered by an increasingly technocratic and commercial discourse amongst the new coordinating bodies such as UN Global Pulse and the Stanford Peace Innovation Lab. Like corporations, they use the language of the private sector to frame human and economic development aims. One example comes from the director of the Peace Innovation Lab at Stanford:

'If you can measure something, you can design for it; if you can design for it you can create new value; if you can create new value you can monetize it. Our aim is to create peace businesses.'⁴

Similarly, the methodologies used by Global Pulse's analytics lab are common to marketing researchers and business analysts, and the lab's director has said that Global Pulse is 'trying to track unemployment and disease as if it were a brand'.⁵

The rise of the 'welfare-industrial complex' (Jayaram, 2014) combined with the availability of vast new digital datasets has increased the risk that human development will become framed as an engineering problem.⁶ Cherlet (2014) notes that this engineering approach to development is based on both classic technological determinism and a new form of 'epistemic determinism', the latter denoting an approach which does not distinguish between technical knowledge produced remotely and that based on local understanding. This determinism tends to distance human development priorities from local concerns and understandings of development problems, and from academic research domains which might add a historical or political perspective. For example, the 'grand challenges' the leaders of Project Lucy propose to solve, including disease, illiteracy, contaminated water and food insecurity can also be seen as grand challenges for politics, governance and activism since they are strongly related to embedded social and economic inequality, governance priorities and, often, international power asymmetries around issues such as trade and debt. Framing these as data and engineering problems has two repercussions. First, it shuts out the expertise and knowledge of small-scale organisations, qualitative researchers and local activists, who have traditionally acted as the extension agents of states and development organisations in actually applying pro-poor policies in LMICs; and second, it may divert power, money and resources from these smaller-scale actors. Meanwhile, the resources and institutional power currently being marshalled to support the engineering approach makes for an extensive and in-depth way of monitoring activities and trends as a whole, making this kind of collaboration arguably more apt as a surveillance infrastructure than as a way of solving development problems.

3.2. Shadow maps and state data doubles

States have always gathered data on their own populations, and more recently those with less resources have been funded to do so by international donors such as USAID (organiser of the worldwide

Demographic and Health Surveys [DHS] program) or UNICEF (which carries out the Multiple Indicator Cluster Surveys [MICS]). The sheer cost of building large, longitudinal datasets on developing countries, however, has led the Center for Global Development, an influential development think tank, to suggest that development donors should turn to the private sector and 'experiment with new institutional models, such as public-private partnerships or crowdsourcing, to collect hard-to-obtain data or outsource data collection activities' (Center for Global Development, 2014). This can already be seen occurring through the growing visibility of institutions such as UN Global Pulse and corporate data-gathering efforts such as Project Lucy. States, rather than generating statistics, must now find ways to access them from the private sector.

And as the new sources of digital data start to parallel or even supplant national data collection efforts, they will increasingly offer the opportunity to re-visualise countries and populations – another element of the gradual shift from legibility towards visibility. As observed data at the population level becomes a possibility, big data analytics performed on LMICs are starting to generate state-level versions of the 'data doubles' theorised by Haggerty and Ericson (2000). The examples above, particularly that of Project Lucy, denote a process where instead of human bodies becoming abstracted, whole cities and countries can be datified, analysed and interpreted largely in isolation from national processes of data gathering and analysis. Grumbach and Frénot (2013) have written in *Le Monde* that Google now knows more about France than its national statistical agency, INSEE. If this is the case, then the amount of new information big data can provide about a country such as Angola, which did not conduct a census between 1970 and 2014, is exponentially greater (though also less representative, since a smaller share of Angola's population uses digital technologies).

Orange's 2013 Data for Development project demonstrates two potential problems with the new country data doubles. The first 'D4D' challenge, organised in 2012–13, involved the first major release of mobile phone calling records from an African country for research purposes. It was designed to encourage both basic and applied research using anonymised Call Detail Records (CDR) from the company's subscribers in Côte d'Ivoire. The aim of the challenge was to 'help address the questions regarding development in novel ways' (Orange, 2012). The company received much positive publicity in its core (European) markets, and aimed to use the research findings to plan development interventions in Côte d'Ivoire. The first iteration of the project resulted in 74 research papers, dealing with human mobility, social and economic development (mapping poverty, tracking economic and social activity), data mining and health (Netmob, 2013).

The Côte d'Ivoire project also demonstrates how the visibilities produced by big data analytics (the research findings were expressed predominantly through visualisations and network analysis) may be at the same time rhetorically powerful and yet unreliable due to their unknown bias. They illustrate Shearmur's point (2015) that such datasets reflect users and markets, with one research team (Mao et al., 2013) pointing out that their findings were based on 'a significant sample of all Côte d'Ivoire mobile users that we assume is not strongly biased, lacking data to the contrary' (italics added). Gutierrez et al. (unpublished), similarly warned of the 'unreliability of census reports' from Côte d'Ivoire provided as ground truth for the remote research performed. The analysis by Berlingerio et al. (2013), for example, shows the effects of a lack of supplementary data to give context to digital traces. Creating a transport optimisation model for Abidjan, Côte d'Ivoire's capital city, the researchers were provided with a dataset that offered only a partial image of the transport network that did not show what factors might be impeding traffic flows (Taylor, 2015b). Nevertheless, the data double assembled from the available (online) sources was used to create a model which was then

⁴ Mark Nelson, Stanford Peace Innovation Lab Co-Director. Panel at the Leiden University Big Data for Peace Summer School, August 16–22, The Hague.

⁵ Robert Kirkpatrick, director, UN Global Pulse, panel at Leiden Innovation Lab, 18.8.2014.

⁶ The discourse on development as a unitary issue lending itself to engineering solutions is common to many big data projects, for example the 'Big Data for Social Good' project at Harvard: <http://www.hsph.harvard.edu/ess/>.

put forward as a finished data assemblage, ready to guide transport policy. Although the model was designed to apply to Côte d'Ivoire's whole transport sector, the data in fact reflected only the parts of it that were expressed through digital data. This mismatch shows on the one hand how big data can mislead through bias and lack of context, but on the other how corporations that hold data can be empowered to draw conclusions on a policy level about developing countries, despite the potential inaccuracy of the data doubles they can create.

The new data sources allow not only country-level data doubles, but what might be termed shadow-maps: the use of spatial data to empower new forms of visibility separate from state mapping efforts. For example the new maps being derived at the time of writing through data donated by mobile phone corporations to track and predict the Ebola epidemic in West Africa (Talbot, 2014) may have huge implications for foreign donors who wish to target interventions and funding, but also for those who wish to enforce quarantine and close borders. The new shadow maps are set to become ever-more detailed and ubiquitous. Manyika et al. (2011) have projected that by 2020 LMICs will be the primary producers of geolocated data from the use of digital communications technologies such as mobile phone use, social media such as geolocated Tweets, and the traceability of IP addresses when individuals use the internet. These are being complemented by the use of powerful imaging technologies in LMICs including satellites and drones, with the latter becoming an important resource for commercial, development and humanitarian purposes. Whilst satellites provide the data used in most types of conventional maps, the production of spatial data is an inevitable by-product of drones' primary functions, which are often commercial. Drones are advocated as a way of gaining access to rural and remote customers in LMICs whilst saving those countries the need to build roads (BBC News, 2012); they are already deployed by the UN in peacekeeping operations in the Democratic Republic of Congo (Crowe, 2013) and by entrepreneurs for humanitarian response after natural disasters (Churchill, 2014). The World Bank is seeking to demonstrate the potential of drones for predictive and planning purposes 'in many sectors including: cadastral mapping/registration, infrastructure projects (roads, energy and dams), urban planning, and disaster risk management' (Volkman, 2014).

Each of these LMIC spatial data projects has a different purpose, is made up of different institutional configurations (private sector firms such as Orange, multilaterals such as Global Pulse, or NGO/multilateral collaborations such as that between the World Bank and various partners in Haiti,⁷) and is subject to different forms of governance. However, together they are generating shadow maps that offer an alternative, dynamic and highly detailed account of the spatial dynamics of even very remote places where accurate mapping has previously been limited or nonexistent. They also present new challenges of interpretation which demand new skills and levels of contextual understanding from researchers, since different technologies make people visible in different ways. For example, mobile phone location data displays people's movement and activities, but not their interactions with their physical surroundings (Taylor, 2015a), whilst satellite data shows people's physical context, but does not offer insights into what they may be doing. As Nathaniel Raymond of the Harvard Humanitarian Initiative has said, '[there is a] lack of specific humanitarian pedagogy for interpreting this data',⁸ so that those in possession of detailed maps formed for use in humanitarian or longer-term development work are effectively flying blind in terms of how to deal with it.

The implications of the new shadow maps are far-reaching. Just as military maps have been put to civilian use, so methods for generating spatial data may see function creep from civilian to military uses. Similarly, these maps-as-byproducts of other processes are likely to find uses beyond their original purposes. For example, the Harvard Humanitarian Initiative's Satellite Sentinel Project (SSP) uses satellite data to map and monitor human rights abuses on Sudan's southern border. Nathaniel Raymond, leader of the initiative, relates:

'What we realised with Satellite Sentinel was that we were changing the battle space. We had gone from this idea that we were passive observers – our monitoring hat ... [to] changing the tempo at which the emergency happens and the very nature and outcome of the emergency itself. It is doing that by providing unique situational awareness otherwise not available to armed actors in real time. There is intelligence-grade situational awareness in the case of sensors: what can be used to document a human rights abuse can also be used to target an artillery strike.'

[Nathaniel Raymond, interview 25.2.2015]

Spatial data projects may be dedicated to providing security, development planning information or reaching a new customer base, yet they increasingly provide as byproducts data that can map some of the world's remotest places. A humanitarian, peace-keeping, business or 'development' drone in a remote and virtually unmapped area is inevitably collecting data not only on its targets but on everyone. Spatial data derived from drones, mobile phone records, social media and other emerging sources offer real-time, dynamic information on where people are located and what they are doing socially, economically and politically. Such datasets may have particular power where they present an alternative mapping of low-income and remote communities, in ways which have already been recognised as highly sensitive in HICs (ACLU, 2011). As Raymond explains, in the case of digital humanitarianism:

'You have data that is publicly available that by itself is not weaponised to be harmful. But then digital humanitarians come in and start creating cross-cutting correlating datasets that allow seemingly static data [...] to turn into a bomb.'

[Interview, 25.2.2015]

The new shadow-maps and country data doubles are the results of a structural change in the way digital data is produced and used. Corporations are now embedded to an unprecedented extent in the production of data that can be used to track, monitor, map and analyse activities in LMICs, and their data trickles upward towards more powerful, technologically adept collaborators such as the UN (e.g. Global Pulse), researchers (such as the Harvard Humanitarian Initiative), development donors and, inevitably, national intelligence services. This complex structure of information gatherers and users gives rise to a diverse set of visibilities which underpin more distributed governance structures. Multiple possible futures are created when others access these data, both good and bad – but in every case power accrues to wealthier international actors, who have the access to the data and the capacity to analyse it, at the expense of less well-resourced local ones who do not. Thus the deep embedding of corporations in processes of creating country data doubles creates a default where those doubles are not accessible to states, and where mapping, counting and sorting become distributed rather than centralised functions.

4. Conclusion

We have outlined two trends which demonstrate an institutional and political power shift with regard to digital data in LMICs. First, the way in which practices of informational capitalism are shaping development interventions via public-private

⁷ <http://www.fastcompany.com/1554072/open-source-maps-are-helping-world-bank-save-lives-haiti>.

⁸ Nathaniel Raymond, Harvard Humanitarian Initiative, presentation at Humanitarian Innovation Conference, 19–20 July 2014, Keble College, Oxford.

partnerships. Multinational technology firms are involved in interventions to map, sort and categorise, often also connecting people to new technologies for the first time, and in return capture data on new technology users and, through that data, new consumers and markets. The second trend is for the visibilities created by these corporate interventions to constitute new population-level databases and maps. We characterise these new data infrastructures as shadow maps and state data doubles because they operate in parallel to state data collection infrastructures, and give rise to distributed forms of governance and powers to intervene.

This power shift from state to corporate data collectors and processors suggests that there is a need to move beyond Scott's influential 1998 framing of population-level data collection as providing legibility and governability. In the cases we have explored, the visibility provided by big data offers a 'god's eye view' (Pentland, 2011) in terms of potential scale and detail – but if the new data create gods, they are gods with limited understanding. Unlike survey data, the new visibilities reflect those using particular digital technologies or services, and without analogue sources such as censuses and qualitative research to provide a baseline or contextual information, there is no way to pin down their bias and understand what they represent (González-Bailón et al., 2012). Despite offering an unprecedented level of detail on individual activities, movements and behaviour, big data is therefore also very fuzzy as a basis for intervention. Furthermore, the new visibilities are different from legibility in that the data are collected by unaccountable third parties rather than states, and therefore do not offer the potential for increasing accountability or representation.

What, then, are the risks of such visibilities in a context of development? First, a lack of accountability may give rise to a culture of experimentation amongst technology giants (as has already been seen in the mass experiments conducted on users by various online services such as Facebook and OKCupid). Greenleaf (2013: 11) shows that the overwhelming majority of states without privacy or data protection laws are low- or lower-middle-income countries. For example, in Sub-Saharan Africa in 2013, only 8 states out of 55 had data protection laws (ibid). Although many countries are adopting versions of the EU or US data protection standards, these do not so far restrict corporations from sharing data for research or humanitarian reasons. If big data analytics in the name of Development remain effectively unregulated, human-subject experiments framed as development are more than a possibility. Reid-Henry (2011) has warned of the use of 'the global borderlands as an appropriate site for experimentation in the government of peoples': in this paper we have charted one route that this type of experimentation may take.

Next, if visibility distributes the power to govern and diminishes accountability, it also creates the power to represent data subjects. Individuals cannot control how they are represented, however, where their data flows through private sector systems and is then sold on or given out pro bono to third parties. Big-data representation occurs on an aggregate scale and the subject has no control over – or often knowledge of – their visibility. Similar problems apply on the state level: representing countries using big data has implications for state sovereignty, given that big data can make visible economic and social dynamics which are relevant to state concerns but which are not collected using state mechanisms or necessarily shared with states. In PPPs or under conditions where corporations act as development donors, states are not – or are only partly – in charge of collecting and disseminating these representations of themselves.

A further risk posed by the new data visibilities is that of what Morozov (2013) has termed 'solutionism': the application of engineering solutions to problems that are long-term and structural in nature. The problems the new data are being applied to are old and

deeply embedded, and it is fair to ask whether corporations can offer anything new in terms of the far-reaching structural change that is necessary to 'solve' problems such as disease, poverty and inadequate sanitation – or whether the language of 'solutions' is even relevant in this domain. Responsible data science is being conducted using commercial data (e.g. Bengtsson et al., 2011, which used mobile data to map cholera transmission in Haiti in the aftermath of the 2011 earthquake), but despite the positive potential of the new shadow-maps and data doubles, so far findings have yet to connect with action (Talbot, 2014). If the authorities with the resources and capacity to use better data do not act on it, the data simply constitutes, as Nathaniel Raymond of the Harvard Humanitarian Initiative has said, 'a post-mortem' on problems in LMICs.

Last, perhaps the most serious risk is that informational capitalism becomes the new norm for the field of international development. States and traditional development donors have until now 'seen like states', using legibility through conventional statistics as a basis for intervention. Yet today, if people are most visible to institutions with the most data and analytical power, LMIC citizens are now seen most clearly by multinational technology corporations. Meanwhile, the possession of unprecedentedly extensive data is empowering the private sector as primary actors, rather than contractors, in development planning and interventions. The shift from legibility to visibility – and from public-sector reading to privatised seeing – requires both new theory and a new ethical approach. In theoretical terms, development interventions may increasingly become a byproduct of informational capitalism, rather than a way for states to achieve domestic or foreign policy aims, and would therefore need to be assessed according to different criteria and with a different institutional focus. Ethically, if LMICs are becoming the new laboratories of data science, there is a case to be made for closely monitoring the boundary between care and control; for determining responsible bodies to regulate and monitor the data science being conducted on LMIC subjects; and particularly for reevaluating the nature and priorities of privacy in an LMIC context, where – as we have shown – accountability for data use and intervention may be minimal.

Acknowledgement

This work was supported by the European Union's Marie Curie postdoctoral program, Grant #624583.

References

- ACLU, 2011. Protecting Privacy from Aerial Surveillance: Recommendations for Government Use of Drone Aircraft. Online publication, accessed 15.08.14 at <<https://www.aclu.org/files/assets/protectingprivacyfromaerialsurveillance.pdf>>.
- Amoore, L., 2011. Data derivatives. On the emergence of a security risk calculus for our times. *Theory Cult. Soc.* 28 (6), 24–43.
- Bauman, Z., Bigo, D., Esteves, P., Guild, E., Jabri, V., Lyon, D., Walker, R.B.J., 2014. After Snowden: rethinking the impact of surveillance. *Int. Political Sociol.* 8 (2), 121–144.
- BBC News, 2012. Matternet: swapping roads for flying drones. Accessed 15.08.14 at <<http://www.bbc.com/future/story/20120209-i-say-to-you-today-i-hover-dream>>.
- Bengtsson, L., Lu, X., Thorson, A., Garfield, R., von Schreeb, J., 2011. Improved response to disasters and outbreaks by tracking population movements with mobile phone network data: a post-earthquake geospatial study in Haiti. *PLoS Med.* 8 (8), e1001083.
- Berlingerio, M., Calabrese, F., Di Lorenzo, G., Nair, R., Pinelli, F., Sbodio, M.L., 2013. AllAboard: a system for exploring urban mobility and optimizing public transport using cellphone data. *Mobile Phone Data for Development: analysis of mobile phone datasets for the development of Ivory Coast*. Netmob conference, 1–3 May 2013, MIT, Boston (also published in *Machine Learning and Knowledge Discovery in Databases* (pp. 663–666). Springer, Berlin Heidelberg).
- Bowker, G., Star, S., 2000. *Sorting Things Out. Classification and Its Consequences*. The MIT Press, Cambridge, Massachusetts.
- Burns, R., 2014. Rethinking big data in digital humanitarianism: practices, epistemologies, and social relations. *GeoJournal*, 1–14.

- Business Standard News, 2013. 'UIDAI has issued 400 mn Aadhaar numbers till date'. Accessed 06.10.14 at <http://www.business-standard.com/article/economy-policy/uidai-has-issued-400-mn-aadhaar-numbers-till-date-113081701225_1.html>.
- Castells, M., 1996. *The Rise of the Network Society. The Information Age: Economy, Society and Culture*, vol. 1. Blackwell, Oxford, UK.
- Center for Global Development, 2014. Delivering on the Data Revolution in Sub-Saharan Africa. Accessed online 17.9.2014 at <<http://www.cgdev.org/sites/default/files/CGD14-01%20complete%20for%20web%200710.pdf>>.
- Cherlet, J., 2014. Epistemic and technological determinism in development aid. *Sci. Technol. Human Values* 39 (6), 773–794.
- Churchill, E., 2014. Philippines: drones and spreadsheets to the rescue. <<http://blogs.worldbank.org/eastasiapacific/philippines-drones-and-spreadsheets-rescue>>.
- Cohen, J., 2013. What privacy is for. *Harv. L. Rev.* 126, 1904.
- Crampton, J.W., Roberts, S.M., Poorthuis, A., 2014. The new political economy of geographical intelligence. *Ann. Assoc. Am. Geogr.* 104 (1), 196–214.
- Crowe, A., 2013. United Nations' drones: A sign of what's to come? Accessed 15.08.14 at <<http://www.privacyinternational.org/blog/united-nations-drones-a-sign-of-whats-to-come>>.
- Degli Esposito, S., 2014. When big data meets dataveillance: the hidden side of analytics. *Surveill. Soc.* 12 (2), 209–225.
- Dreze, J., 2010. Unique facility, or recipe for trouble? *The Hindu*, November 25, 2010. <<http://www.thehindu.com/todays-paper/tp-opinion/article911931.ece>> (accessed 16.09.14).
- EFF, 2015. Internet.org Is Not Neutral, Not Secure, and Not the Internet. Electronic Frontier Foundation, May 18, 2015. Accessed 15.6.2015 at <<https://www.eff.org/deeplinks/2015/05/internetorg-not-neutral-not-secure-and-not-internet>>.
- Global Pulse, 2012. Big Data for Development: Opportunities & Challenges. White Paper. Accessed 22.08.14 at <<http://www.unglobalpulse.org/projects/BigDataforDevelopment>>.
- Global Pulse, 2013. United Nations Global Pulse Information Sheet. <www.unglobalpulse.org>.
- Global Pulse, 2014. Mining Indonesian tweets to understand food price crises. Methods paper. Accessed 07.08.14 at <<http://www.slideshare.net/unglobalpulse/global-pulse-mining-indonesian-tweetsfoodpricecrises-copy>>.
- González-Bailón, S., Wang, N., Rivero, A., Borge-Holthoefer, J., Moreno, Y., 2012. Assessing the bias in communication networks sampled from twitter. Available at SSRN 2185134.
- Graham, M., 2009. Neogeography and the palimpsests of place: web 2.0 and the construction of a virtual earth. *Tijdschrift voor Economische en Sociale Geografie* 101 (4), 422–436.
- Green, D., 2014. Big data and development: upsides, downsides and a lot of questions. Accessed 04.08.14 at: <<http://oxfamblogs.org/fp2p/what-is-the-future-impact-of-big-data/>>.
- Greenleaf, G., 2013. Sheherezade and the 101 data privacy laws: origins, significance and global trajectories. *J. Law Inf. Sci.* 23 (1).
- Grumbach, S., Frénot, S., 2013. Les données, puissance du futur. *Le Monde*. Accessed 5.7.2016 at <http://www.lemonde.fr/idees/article/2013/01/07/les-donnees-puissance-du-futur_1813693_3232.html>.
- GSMA (2013b), 2013: the year in research. Accessed 22.08.14 at <<https://gsmaintelligence.com/files/analysis/?file=2013-12-12-2013-the-year-in-research.pdf>>.
- Haggerty, K.D., Ericson, R.V., 2000. The surveillant assemblage. *Br. J. Sociol.* 51 (4), 605–622.
- Harley, J., 1989. Deconstructing the map. *Cartographica* 26 (2), 1–20.
- Hildebrandt, M., 2013. Slaves to Big Data. Or Are We? *Idp. Revista De Internet, Derecho y Política* 16.
- Hosein, G., Nyst, C., 2013. Aiding Surveillance: An Exploration of How Development and Humanitarian Aid Initiatives are Enabling Surveillance in Developing Countries. Available at SSRN 2326229.
- IBM, 2014. Project Lucy: How IBM Research is engaging in Africa. White paper, IBM, New York. Accessed 25.08.14 at <<http://asmarterplanet.com/files/2014/03/Project-Lucy-High-Res.pdf>>.
- IndiaToday, 2013. Aadhaar scheme: How safe and secure is your identity? Accessed 12.08.14 at <<http://indiatoday.intoday.in/story/aadhaar-scheme-how-safe-and-secure-is-your-identity/1/313616.html>>.
- Internet.org, 2014. Announcing the Connectivity Lab at Facebook. Accessed 07.08.14 at <<http://www.internet.org/press/announcing-the-connectivity-lab-at-facebook>>.
- Jayaram, M., 2014. India's Big Brother Project. *Boston Review*, May 19th 2014. Accessed online 06.10.14 at <<http://www.bostonreview.net/world/malavika-jayaram-india-unique-identification-biometrics>>.
- Jerven, M., 2013. *Poor Numbers: How We Are Misled by African Development Statistics and What to Do about It*. Cornell University Press, Ithaca.
- Kirkpatrick, R., 2011. Data philanthropy: public and private sector data sharing for global resilience <<http://www.unglobalpulse.org/blog/data-philanthropy-public-private-sector-data-sharing-global-resilience>>.
- Kitchin, R., 2014. The real-time city? Big data and smart urbanism. *Geojournal* 79, 1–14.
- Krishnan, R., 2012. Aadhaar: A National Innovation Platform? Accessed 06.10.14 at <<http://jugaadtoinnovation.blogspot.nl/2012/06/aadhaar-national-innovation-platform.html>>.
- LiveMint, 2013. Nasscom, UIDAI to push Aadhaar app development. Accessed online 25.08.14 at <http://www.livemint.com/Politics/T4Wj3cVQ1qiVfUq5ZJ99N/Nasscom-UIDAI-to-push-Aadhaar-app-development.html?utm_source=copy>.
- Livingstone, S., Walter-Drop, G., 2014. *Bits and Atoms: Information and Communication Technology in Areas of Limited Statehood*. OUP, Oxford.
- Lyon, D., 2007. *Surveillance Studies. An Overview*. Polity Press, Cambridge.
- Lyon, D., 2014. Surveillance, Snowden, and Big Data: Capacities, consequences, critique. *Big Data Soc.* July–September, 1–13.
- Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., Hung Byers, A., 2011. *Big Data: The Next Frontier for Innovation, Competition and Productivity*. McKinsey Global Institute, Washington DC.
- Mao, H., Shuai, X., Ahn, Y.Y., Bollen, J., 2013. Mobile communications reveal the regional economy in Côte d'Ivoire. In: *Proceedings of the 3rd Conference on the Analysis of Mobile Phone Datasets (NetMob)*.
- Mayer-Schönberger, V., Cukier, K., 2013. *Big data: a revolution that will transform how we live, work, and think*. Houghton Mifflin Harcourt.
- Mirafab, F., 2004. Public-private partnerships the Trojan horse of neoliberal development? *J. Plan. Educ. Res.* 24 (1), 89–101.
- Morozov, E., 2013. To save everything, click here: Technology, solutionism, and the urge to fix problems that don't exist. Penguin, UK.
- MTN, 2010. MTN launches Facebook® ZERO, a FREE way to access Facebook® on your Mobile. Accessed 07.08.14 at <<http://www.mtn.co.ug/About-MTN/News-Room/2010/May/MTN-launches-Facebook%20AE-ZERO,-a.aspx>>.
- Netmob, 2013. In: *Proceedings of the 3rd Conference on the Analysis of Mobile Phone Datasets*, May 1–3, MIT, Boston.
- Nilekani, N., 2013. Technology to Leapfrog Development: The Aadhaar Experience. Richard H. Sabot lecture, Center for Global Development, April 22 2013. <<http://www.cgdev.org/sites/default/files/nandan-nilekani-sabot-lecture-transcript-technology-leapfrog-development.pdf>> (accessed 16.09.14).
- Orange, 2012. D4D project: <<http://www.d4d.orange.com/learn-more>>.
- Pentland, A., 2011. Society's nervous system: building effective government, energy, and public health systems. *Pervasive Mobile Comput.* 7 (6), 643–659.
- Polzer, T., Hammond, L., 2008. Invisible displacement. *J. Refug. Stud.* 21 (4), 417–431.
- Ramanathan, U., 2013. Who Owns The UID Database? Medianama, <<http://www.medianama.com/2013/05/223-who-owns-the-uid-database-usa-ramanathan>> (accessed 16.09.14).
- Reid-Henry, S., 2011. Spaces of security and development an alternative mapping of the security-development nexus. *Secur. Dialogue* 42 (1), 97–104.
- Ruppert, E., 2012. The governmental topologies of database devices. *Theory Cult. Soc.* 29 (4–5), 116–136.
- Schneier, B., 2015. *Data and Goliath. The Hidden Battles to Collect Your Data and Control Your World*. W.W. Norton and Company, New York.
- Scott, J.C., 1998. *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed*. Yale University Press.
- Shearmur, R., 2015. Dazzled by data: big data, the census and urban geography. *Urban geography*. <http://dx.doi.org/10.1080/02723638.2015.1050922>.
- Talbot, D., 2014. Cell-Phone Data Might Help Predict Ebola's Spread. MIT Technology Review. Accessed 17.09.14 at <http://www.technologyreview.com/news/530296/cell-phone-data-might-help-predict-ebolas-spread/?utm_content=buffer05d6d&utm_medium=social&utm_source=twitter.com&utm_campaign=buffer>.
- Taylor, L., 2015a. No place to hide? The ethics and analytics of tracking mobility using mobile phone data. Accessed 30.06.15 at <https://www.academia.edu/7502204/No_place_to_hide_The_ethics_and_analytics_of_tracking_mobility_using_mobile_phone_data>.
- Taylor, L., 2015b. A god's eye view? Big data, ground truth and the D4D challenge. Accessed 30.06.15 at <https://www.academia.edu/8079377/A_god_s_eye_view_Big_data_ground_truth_and_the_D4D_challenge>.
- Taylor, P.J., Johnston, R.J., 1995. Geographic information systems and geography. In: Pickles, J. (Ed.), *Ground Truth*.
- Taylor, L., Schroeder, R., 2014. Is bigger better? The emergence of big data as a tool for international development policy. *Geojournal*, 1–16.
- Thomas, O.D., 2014. Foucaultian dispositifs as methodology: the case of anonymous exclusions by unique identification in India. *Int. Political Sociol.* 8 (2), 164–181.
- UIDAI.gov.in, 2014. About UIDAI. Accessed 12.08.14 at <<http://uidai.gov.in/organization-details/uidai-headquarter.html>>.
- Volkman, W., 2014. Small Drones for Progress in International Development Projects. Post on LinkedIn, April 24, 2014. Accessed 15.08.14 at <<https://www.linkedin.com/post/article/20140424005308-38421723-small-drones-for-progress-in-international-development-projects>>.
- World Economic Forum, 2012. Big Data, Big Impact: New Possibilities for International Development. Accessed 22.08.14 at <<http://www.weforum.org/reports/big-data-big-impact-new-possibilities-international-development>>.
- Yadav, V., 2013. Unique Identification Project for 1.2 billion People in India. SAI Working Paper, Harvard University.
- Zelazny, F., 2012. *The Evolution of India's UID Program. Lessons Learned and Implications for Other Developing Countries*. Center for Global Development Policy Paper.