Big data and the production of anticipatory transparency in international development: A governmentality analysis

Mikkel Flyverbom, Copenhagen Business School (mfl.ikl@cbs.dk)
Anders Koed Madsen, Aalborg University (akma@learning.aau.dk)
Andreas Rasche, Copenhagen Business School (ara.ikl@cbs.dk)

Submitted to: Sub-theme 51 – "Seeing, Knowing and Governing: The Organization, Institutionalization and Politics of Transparency"

The masses of digital data produced by Internet traffic (e.g. Google searches, tweets, Facebook posts) and various forms of tracking and navigation (e.g. GPS devices) offer new insights into human practices and hidden societal trends. The term 'big data' has been used to describe such large datasets requiring new ways of data storage, analysis, and visualization technologies (Chen, Chiang & Storey, 2012). Much of the excitement surrounding big data is less about the mere existence of larger volumes of real-time data, but more about the ability to aggregate, search, and cross-reference big sets of data (boyd & Crawford, 2012). Of course, corporations and governments have collected and processed large amounts of data for a long time (Scott, 1998; Stephens & Lubar, 1986). This data is usually well organized as well as highly structured and hence can be processed without much difficulty according to established and transparent procedures (e.g. survey-based data). By contrast, the swift rise and availability of big data has pushed many organizations into coping with unstructured data sources: Twitter produces around 500 million tweets per day, Facebook contains more than 3.2 billion likes and comments every day, and 72 hours of video are uploaded to YouTube every minute (all information as of 2012; see Waters, 2012).

Whereas the literature has primarily discussed big data as an opportunity for private enterprises to increase their market share and competitiveness (e.g. McAfee & Brynjolfsson, 2012; Bughin, Livingston & Marwaha, 2011), this paper explores the ways in which big data contributes to improving international development efforts (for another example of this line of research see Hilbert, 2013). Some examples illustrate the use of big data for development (hereafter BD4D). First, Google has found that there is a close relationship between search queries for dengue-related topics and the number of people affected by the disease (Ginsberg

et al., 2009). This allows for a more precise and earlier detection of where and when dengue outbreaks are likely to occur. Second, Churana et al. (2012) found that trends in volume of data from social and news media correlated in time with official cholera case patterns during the 2010 Haitian outbreak, but were available up to two weeks earlier. Third, the UN Global Pulse (2012) initiative analyzed large amounts of tweets commenting on the price of rice in Indonesia. The analysis showed that the quantity of tweets on the topic followed the official inflation for the food basket in the country, indicating that social media data could be used as a predictor of price trends on local markets (UN Global Pulse & Crimson Hexagon, 2011). Policymakers are currently realizing the potential of these BD4D applications to produce actionable information that can be used to improve development – e.g. by identifying needs, providing services, and predicting and preventing crises (World Economic Forum, 2012).

The paper starts from the assumption that such new data practices afford new ways of seeing and knowing in organizations engaged with development efforts, and that big data analyses produce particular types of (anticipatory) transparency that deserve more scrutiny. This raises important questions regarding the relationship between big data technologies and the norms and forms of transparency at work in the production of knowledge. The aim of this paper is to unpack how big data analysis leads to particular ways of making international development legible and transparent, and how related visualizations shape governance in this area. We are interested in pinpointing the mechanisms by which big data analysis renders certain areas of international development knowable and governable, and in analyzing the implications of this emerging 'regime of governance' for transparency. More precisely, we aim to undertake this analysis through an exploration of the practices and rationales of government that allow aspirations of reform, such as BD4D, to be constituted.

The theoretical framework used to address this question is based on Michel Foucault's (1978, 1991) notion of governmentality and, in particular, its elaboration through the work of Mitchell Dean (2009, 1996, 1995). Governmentality aims "to uncover and examine the often invisible rationality which is behind an assemblage of actions and mechanisms that are in place to govern certain actions." (Gouldson & Bebbington, 2007: 12) Although scholarly work has used the governmentality lens to explore the rationales, practices and power structures underlying international development (see e.g. Murray Li, 2007), we know very

little about how recent changes in sourcing, processing and communicating relevant data affect the way development problems are depicted and addressed. This is a surprising omission since data, and the analytical techniques attached to it, shape how governance problems are "re-presented in the place where decisions are to be made about them." (Miller & Rose, 1990: 7) The governmentality lens fits well with our research aim. It offers a theoretical framework to discuss the necessary conditions for big data to have an effect on the way international development is practiced and thereby also on the way organizations in this field work with transparency. The conceptualization of the kinds of governance made possible by big data underpins our discussion of the particular norms and forms of transparency that result from these. Thus, we seek to contribute to current research on the workings of transparency in governance (Florini, 2007; Flyverbom, forthcoming; Fung et al., 2007; Hood & Heald, 2006).

The governmentality lens allows us to explore the linkages between new technologies, transparency, and new modes of governance. To embed this lens into our analysis, we structure the discussion around Dean's (2009) four dimensions of an analytics of government: the fields of *visibility* surrounding regimes of practices, the instruments and techniques (*techne*) that enable and constrain these regimes, the forms of knowledge (*episteme*) attached to certain regimes, and the forms of *identity* that belong to them. Together these four dimensions provide a framework to examine the specific conditions under which BD4D emerges as a way of producing anticipatory transparency when addressing international development problems. Using Dean's four dimensions as a yardstick for our analysis, we argue that the effective uptake of BD4D is conditioned upon the following four dimensions that each raise questions about the relation between big data technologies and norms and forms of transparency involved in the governance of international development.

(1) Fields of Visibility: BD4D is conditioned upon the utilization of new ways of visualizing development problems and hence making unacknowledged dimensions of these problems transparent to decision makers. For instance, the granularity of big data sources enables organizations to move away from the kind of aggregates that have been central to statistical inquiry. Platforms like Twitter are providing data that makes individual feelings and behaviors of governed subjects visible in real-time. In other

words, big data makes new aspects of the developmental world transparent to legislators and decision makers in a much faster way than, for instance, survey data.

(2) Techne: BD4D is conditioned upon the technical integration of data provided by private companies (e.g. Twitter) with the synthesizing algorithms developed by public and nonprofit organizations. This means that no single professional or organization is responsible for the whole data-production process. Compared to data produced in-house, the provenance of big data is therefore less transparent for organizations at the end of the production chain. For instance, an organization like the United Ntaions must give up a certain degree of transparency and control over data in exchange for fast 'digital smoke signals'. This, in turn, influences accountability mechanisms, which are important in the context of addressing development challenges.

(3) Episteme: BD4D is conditioned upon the acceptance of new epistemic foundations for governing development problems (e.g. when developing policies). Standardized guidelines for the production and analysis of statistical data (e.g. conventions on sample sizes and p-values) are hard to transfer to the new and fluctuating data landscape offered by big data. This means that the criteria for legitimate and actionable data become less transparent than they were in statistics based on sampling methods. There are no agreed-upon procedures that one can use to assess the validity of this kind of data, creating challenges for organizations like the United Nations and The World Bank, both which have been working with household surveys as a source of data for a long time.

(4) Formation of Identities: BD4D is conditioned upon the acknowledgment that working with big data challenges professional and organizational identities (e.g. when "traditional" development analysts need to turn into data-savvy managers). This makes the locus of expertise and the competencies needed to be a good analyst less transparent than they are in situations where the identity of established professions is more stable. Furthermore, the data used within BD4D applications is often re-purposed from other contexts (e.g. when using Google searches). This means that people producing data 'on the ground' have very limited possibilities for knowing what their data is being used

for. Compared to a traditional respondent in a sample exercise, big data practices are not driven by an interest in giving subjects insight into the life of the data they produce.

We suggest that, taken together, these conditions problematize selected aspects of established data processing practices used in international development. Furthermore, they raise important questions regarding the relation between BD4D and transparency as a form of governance in the field of development. Accordingly, the paper illustrates that BD4D is a regime of practices that is double-headed in its relation to the way it fits with existing norms of transparency. In other words: Using big data to address development problems involves making selected development challenges more transparent (e.g. when Google Flu trends allows predicting flu outbreaks for specific cities instead of entire regions). However, our discussion also shows that BD4D's own transparency can be challenged in numerous ways.

With this in mind, the contribution of this paper is twofold. First, we extend the emerging scholarly discourse around the societal relevance of big data (Hilbert, 2013; Meier, 2012) by theorizing BD4D as a particular, yet subtle, form of power shaping how problems and opportunities in international development are made visible, framed and acted upon. Second, we contribute to discussions on how transparency constitutes a mode of governance shaping areas such as social, environmental, and economic affairs in the field of development. More specifically, we provide insights into how big data practices offer important opportunities and challenges for governance in this field. Our discussions revolve around the particular types and orientations of international development that big data affords, such as anticipatory, pre-emptive pattern recognition, the correlation of disparate signals and real-time aggregations of information. But we also stress that the reliance on algorithmic forms of knowledge production may result in information being packaged so densely – and at an alarming distance from the processes and phenomena they seek to grasp – that we gain little or false insights.

Such worries that big data may produce opacity rather than transparency and accountability echo the concerns of Mayer-Schoenberger & Cukier (2013) when they stress that big data analyses may end up as black boxes, because the data and algorithmic operations involved are too intricate for anyone to understand and scrutinize. Such opacities are also

central to Kallinikos' (2013) concern that big data analyses may produce signals and predictions that are dangerous to trust in governance. By way of example, he stresses that in part we can understand the financial crisis as a result of an overwhelming trust in machine intelligence at the expense of human judgment and expertise. Our discussions of the ramifications of big data in terms of transparency and opacity contribute to current attempts to understand the role of transparency in governance. While it is widely recognized that the international development agenda is shifting (e.g. towards stronger collaboration with private actors when reducing poverty; Ilcan & Phillips, 2010; United Nations, 2013), the governmentality effects of working with large-scale data remain unacknowledged to date.

The paper will be structured as follows. We start by describing the context and background of BD4D, showing first what characterizes big data more generally to then discuss how big data is used in different ways to make development problems knowable and governable. Next, we introduce the governmentality lens in more detail, mostly drawing on Mitchell Dean's four dimensions of an analytics of government. Based on this, we analyze the conditions under which BD4D emerges as a regime of knowledge production and foundation for governance in international development, using this analysis as a springboard to discuss in which ways BD4D produces transparency and is, at the same time, undercutting it. The paper concludes by discussing the implications of our findings, acknowledging the potentials and limitations of big data when used in the field of development and outlining an agenda for future research at the intersection of big data, governance and transparency.

Word Count: 2770

REFERENCES

boyd, d. and Crawford, K. (2012). Critical questions for Big Data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, 15(5): 662-679.

Bughin, J., Livingston, J., & Marwaha, S. (2011). Seizing the potential of "big data." *McKinsey Quarterly*, (4): 103–109.

Cheng, C., C., Chiang, R. H. L., & Storey, V. C. (2012). Business intelligence and analytics: from big data to big impact. *MIS Quarterly*, 36(4): 1165-1188.

Chunara, R., Andrews, J. R., & Brownstein, J. S. (2012). Social and news media enable estimation of epidemiological patterns early in the 2010 Haitian cholera outbreak. *The American journal of tropical medicine and hygiene*, *86*(1): 39–45.

Dean, M. M. (2009). *Governmentality: Power and Rule in Modern Society (2nd edition)*. London et al: SAGE.

Dean, M. (1996). Putting the technological into government. *History of the Human Sciences*, *9*(3): 47–68.

Dean, M. (1995). Governing the unemployed self in an active society. *Economy and Society*, *24*(4): 559–583.

Florini, A. (2005). *The Coming Democracy: New Rules for Running a New World*, Washington DC: Brookings Institution Press

Flyverbom, M. (forthcoming). Sunlight in Cyberspace? On Transparency as a form of Governance, *European Journal of Social Theory*

Foucault, M. (1991). Governmentality. In G. Burchell, C. Gordon, & D. Murphy (Eds.), *The Foucault Effect: Studies in Governmentality* (pp. 87–104). London: Harvester.

Foucault, M. (1978). *Discipline and Punish: The Birth of the Prison*. New York: Vintage.

Fung, A., Graham, M., & Weil, D. (2008). *Full Disclosure: The Perils and Promise of Transparency*, Cambridge: Cambridge University Press.

Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S., & Brilliant, L. (2009). Detecting influenza epidemics using search engine query data. *Nature*, *457*(7232): 1012–4.

Gouldson, A., & Bebbington, J. (2007). Corporations and the governance of environmental risk. *Environment and Planning C: Government and Policy*, *25*(1), 4–20.

Hilbert, M. (2013). Big Data for Development: From Information- to Knowledge Societies. Available at: http://ssrn.com/abstract=2205145 (Accessed 15 June 2013).

Hood, C. & Heald, D. (2006) *Transparency – the Key to Better Governance?* Oxford: Oxford University Press.

Ilcan, S., & Phillips, L. (2010). Developmentalities and Calculative Practices: The Millennium Development Goals. *Antipode*, *42*(4): 844–874.

Kallinikos, J. (2013). The allure of big data. *Mercury*, 2(3): 41-43.

Mayer-Schönberger, V. & Cukier, K. (2013). *Big data: A Revolution that will transform how we live, work and think*. London: John Murray.

McAfee, A., & Brynjolfsson, E. (2012). Big Data: The Management Revolution. *Harvard Business Review*, *90*(10): 60–68.

Meier, P. (2012). Crisis Mapping in Action: How Open Source Software and Global Volunteer Networks Are Changing the World, One Map at a Time. *Journal of Map & Geography Libraries*, *8*(2): 89–100.

Miller, P. & Rose, N. (1990). Governing Economic Life. *Economy and Society,* 19(1): 1-31.

Murray Li, T. (2007). *The will to improve: governmentality, development, and the practice of politics*. Durham, NC: Duke University Press.

Scott, J. (1998). *Seeing like a state*. New Haven: Yale University Press.

Stephens, C. & Lubar, S. (1986). A Place for Public Business: The Material Culture of the Nineteenth-Century Federal Office, *Business and Economic History,* 15: 165-179.

United Nations. (2013). *A new global partnership: Eradicate poverty and transform economies through sustainable development (The report of the high-level panel of eminent persons on the post-2015 Development agenda)*. New York, NY: United Nations.

United Nations Global Pulse. (2012). *Big Data for Development: Challenges and Opportunities*. New York, NY: United Nations.

United Nations Global Pulse and Crimson Hexagon (2011). *Twitter and perceptions of crisis-related stress: Methodological white paper*. New York, NY: United Nations.

Waters, R. (2012). Push to exploit an ocean of information. *Financial Times*, 10 December, p. 17.

World Economic Forum (2012). *Big Data, Big Impact: New Possibilities for International Development.* Geneva: World Economic Forum.