Available online at www.sciencedirect.com

ScienceDirect

www.compseconline.com/publications/prodclaw.htm

Computer Law & Security Review

Big Data and security policies: Towards a framework for regulating the phases of analytics and use of Big Data

Dennis Broeders [a,b,*], *Erik Schrijvers* [a], *Bart van der Sloot* [c], *Rosamunde van Brakel* [d], *Josta de Hoog* [a], *Ernst Hirsch Ballin* [a,e,f]

[a] *The Netherlands Scientific Council for Government Policy, The Hague, The Netherlands*
[b] *Department of Public Administration and Sociology, Erasmus University Rotterdam, Rotterdam, The Netherlands*
[c] *Tilburg Institute for Law, Technology and Society, Tilburg University, Tilburg, The Netherlands*
[d] *Law, Science, Technology & Society Research Group, Vrije Universiteit Brussel, Brussels, Belgium*
[e] *Department for Public Law, Jurisprudence and Legal History, Tilburg University, Tilburg, The Netherlands*
[f] *Asser Institute for International and European Law, The Hague, on behalf of the University of Amsterdam, The Netherlands*

Keywords:
Big Data
Security
Data protection
Privacy
Regulation
Fraud
Policing
Surveillance
Algorithmic accountability
the Netherlands

ABSTRACT

Big Data analytics in national security, law enforcement and the fight against fraud have the potential to reap great benefits for states, citizens and society but require extra safeguards to protect citizens' fundamental rights. This involves a crucial shift in emphasis from regulating Big Data *collection* to regulating the phases of *analysis* and *use*. In order to benefit from the use of Big Data analytics in the field of security, a framework has to be developed that adds new layers of protection for fundamental rights and safeguards against erroneous and malicious use. Additional regulation is needed at the levels of analysis and use, and the oversight regime is in need of strengthening. At the level of analysis – the algorithmic heart of Big Data processes – a duty of care should be introduced that is part of an internal audit and external review procedure. Big Data projects should also be subject to a sunset clause. At the level of use, profiles and (semi-) automated decision-making should be regulated more tightly. Moreover, the responsibility of the data processing party for accuracy of analysis – and decisions taken on its basis – should be anchored in legislation. The general and security-specific oversight functions should be strengthened in terms of technological expertise, access and resources. The possibilities for judicial review should be expanded to stimulate the development of case law.

[*] *Corresponding author.* The Netherlands Scientific Council for Government Policy, The Hague, The Netherlands and Department of Public Administration and Sociology, Erasmus University Rotterdam, Mandeville Building, Room T17-12, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands.

E-mail address: broeders@fsw.eur.nl (D. Broeders).

http://dx.doi.org/10.1016/j.clsr.2017.03.002

Table 1 – Frame of reference for Big Data.

Data	• Amount of data: large amounts of data are involved. • Organisation of data: Big Data analytics can deal with both structured and unstructured data. • Variety of data: there is a combination of various data sources and data formats (text, sound, video).
Analysis	• Method of analysis: the analysis is *data-driven*, so patterns are sought in the data without pre-established hypotheses. It favours correlations over causality. • Orientation of the analysis: although Big Data analyses also give information about the past (retrospective analyses), it is particularly the analyses of the present (*real-time analyses/nowcasting*) and the future (*predictive analyses/forecasting*) that draw attention.
Use	• Decompartmentalisation of domains: data from one domain are used for decisions in another domain. • *Actionable knowledge*: conclusions at aggregated level can be applied to decisions at group or individual level (person or object).

1. The promise and perils of Big Data in security policies

Big Data is a catchword that promises radical change. Expectations are high when it comes to increasing sales, targeted advertising, optimising processes and generating unforeseen, unexpected and unprecedented insights. According to some, Big Data will revolutionise the way we live, work and think.[1] Governments are keen to make sure that the benefits of these new technologies will be integrated into public policies as well. In the policy domain of security – broadly interpreted as ranging from national security, via law enforcement to the combat and prevention of fraud – the number of programmes that involve large-scale data collection, linking and analyses is on the rise. Most of those are not on the scale of Big Data 'proper' yet, but the trends indicate that this may change in the coming years.

The opportunities and benefits (both potential and realised) of applying Big Data analytics in the security domain are many, including greater operational efficiency and speed, more precise risk analyses and the discovery of unexpected correlations, all of which feed into risk profiles, better targeted inspections and more efficient use of scarce resources. Big Data analyses help in reconstructing past events (immediately after an attack, for example) and are useful in monitoring developments in real time. This is of great value, for example, in traffic management, organising information and aid following a disaster, or for crowd control at events. Most of all, however, there is the promise that Big Data analytics will deliver insights into the future and may provide the foundation for effective preventive policies. However, these potential gains in security might come at a price in terms of individual and collective freedoms and fundamental rights. Just as the state is responsible for the security of its citizens, it is also – and equally – tasked to protect their personal freedom.

This paper aims to lay the groundwork for a regulatory framework for the use of Big Data in security policies that respects and protects fundamental rights. Most crucially, this requires a shift from regulating data *collection* to regulating the *analysis* and *use* of Big Data.

2. A working definition of Big Data

Big Data is still very much a moving target. Technological developments and new applications continue to feed into the debate about what defines Big Data and sets it apart from earlier forms of data analysis. There is no real consensus regarding its key characteristics, although most definitions of Big Data refer to the ubiquitous three Vs.[2] The first of these three stands for *Volume* (the use of large amounts of data), the second V is for *Variety* (the use of diverse data sources that are stored in diverse structures or even in an unstructured way) and the third stands for *Velocity*, or the speed of data processing (data is often analysed in real time). Over time, a number of authors have added additional Vs to this threesome, such as Veracity[3], Variability[4], Value[5] and Virtual[6]. The various definitions do not amount to a broad consensus on the issue but do demarcate the corners of Big Data as a field of study. In this paper, we will not add our own definition, but rather collect a number of important elements from the definitions of others to construct a frame of reference for the use of Big Data in the context of public administration, especially in the security domain.[7] This frame of reference is grouped around three main aspects of Big Data processes: data (collection), analysis (techniques) and the use of Big Data results (see Table 1).

Big Data is seen here as the interplay between these characteristics rather than as a well-defined and definable object. This leaves room to discuss the use of data analysis in public policy making that includes some of these characteristics to a degree but does not tick all the boxes. Most current policy programmes analysed in the Netherlands do not cover the full range of this frame of reference. It is often the potential to grow into full Big Data systems that makes it important to scrutinise policy initiatives now.

[1] Mayer-Schönberger and Cukier (2013); see also Greengard (2015).

[2] Laney (2001).

[3] IBM (2015); Klous (2016).

[4] Hopkins and Evelson (2011); Tech America Foundation (2012).

[5] Dijcks (2012); Dumbill (2013).

[6] Zikopoulos and Eaton (2011); Akerkar et al. (2015).

[7] See WRR (2016): 33–34.

Please cite this article in press as: Dennis Broeders, et al., Big Data and security policies: Towards a framework for regulating the phases of analytics and use of Big Data, Computer Law & Security Review: The International Journal of Technology Law and Practice (2017), doi: 10.1016/j.clsr.2017.03.002