COMPUTER LAW & SECURITY REVIEW ■■ (2017) ■■–■■

Available online at www.sciencedirect.com

ScienceDirect

www.compseconline.com/publications/prodclaw.htm

Computer Law & Security Review

The use of Big Data: A Russian perspective of personal data security

Anna Konstantinovna Zharova [a,*], *Vladimir Mikhailovich Elin* [b]

[a] *Department of Innovations and Business in IT, Faculty of Business and Management, National Research University Higher School of Economics, Moscow, Russia*

[b] *Department of Information Security, National Research University, Higher School of Economics, Moscow, Russia*

Keywords:
Personal data
Privacy
Identifying information
Russia
Personal data security
Russian public authorities
Big Data

A B S T R A C T

This article examines the impact of Big Data technology on Russian citizens' constitutional rights to a private life. There are several laws in the Russian Federation covering data privacy and protection, but these are proving inadequate to protect the citizens' rights in the face of the ever-increasing use of massive data sets and their analysis by Big Data tools. One particular problem in this regard is that datasets of anonymised records currently not covered under personal data laws (because they do not identify individuals) can, in fact, be used to identify data subjects (the individuals to whom the data refers) when combined and analysed using Big Data tools. Furthermore, existing sanctions for misuse of personal data are minor, and often fail to act as a deterrent when the commercial benefits of exploiting user data (e.g. through targeted advertising) are so much greater. From the point of view of companies handling Big Data, a general confusion over definitions and responsibilities is making compliance with the law difficult, leaving most to come up with their own forms of best practice, rather than being able to follow clear industry recommendations. The article examines existing laws and oversight bodies, discusses how the current provisions are inadequate to deal with new developments in Big Data, and proposes recommendations for amending and updating existing laws and policies.

1. Introduction

The development of information technology (IT) has increased the possibilities for personal data to be used in ways that are damaging to the data subject. However, the creation of strong organisational measures and legal regulation can potentially reduce the level of threats and risks. This article focuses on the processing of personal data using what is generally referred to as Big Data technology. We understand "Big Data" as a set of information technology, tools, data and processes that allows the analysis of large sets of structured and unstructured data. A particular difficulty lies in the separation of information into identifying information that relates to specific people, and anonymised data that, in theory at least, gives a general picture without identifying individuals. It has been

* *Corresponding author.* 33 Kirpichnaya Str., Moscow, Russia. Fax: (495) 771-32-38.
E-mail address: ajarova@hse.ru (A.K. Zharova).

http://dx.doi.org/10.1016/j.clsr.2017.03.025

noted that information, which is initially presented as anonymised, can still be used to identify individuals, especially when two or more anonymised databases are combined for analysis with Big Data tools. For example, research of Massachusetts Institute of Technology specified that *"just four fairly vague pieces of information – the dates and locations of four purchases – are enough to identify 90% of the people in a data set recording three months of credit-card transactions by 1.1 million users"*.[1]

This article explores the privacy, security and consumer welfare issues linked with the collection, storage, analysis, processing, reuse and sharing of data within the Russian Federation, with a particular focus on issues that arise with Big Data. It discusses legal issues relating to the management and security of personal data processed by Big Data technology and looks at the role of the state in regulating the industry, exploring existing measures such as legislation and industry guidelines, as well as the role of the courts and of governmental oversight body in enforcing standards. There is a particular focus on the use of data gleaned from the Internet about the activities of users, which can potentially be used to profile individuals for commercial gain (e.g. targeted advertising) to identify good or bad candidates for credit, or to set insurance premiums.

The discussion section explores the effectiveness of Russian legislation in providing clear, usable guidelines for business, adequate protection for individual privacy and appropriate sanctions for organisations that fail to meet these requirements. It examines how in many cases, legislation is contradictory or simply out-of-date when it comes to protecting personal data, especially when datasets created for different purposes are combined for analysis by Big Data tools. The authors furthermore note that administrative sanctions, consisting of minor fines, applicable in the case of abuse of personal data, often fail to deter certain businesses from mining personal data, as the financial benefits of doing so often outweigh the fines.

The article concludes with a set of recommendations for adaptation of current legislation to provide more protection for data subjects, clearer guidelines for companies to help them comply with the legal requirements, and firmer, criminal sanctions in the case of intentional misuse of personal data.

This article seeks to contribute to the literature in two respects. Firstly, it offers insights into how various characteristics of Big Data are linked to privacy, security and consumer welfare issues. Secondly, it shows how the privacy, security and consumer welfare aspects of Big Data are linked to the interrelated issues of information collection, storage, sharing and accessibility.

The following research questions guided this article:

1. Is Big Data a unique information technology that requires the development of new legal and practical approaches to the security and management of personal data?
2. Is the Russian legislative system ready to regulate the security of personal data in Big Data?
3. Can businesses adapt their activities to existing legal principles?
4. Do the principles of the legal requirements satisfy the business environment?
5. How can recommendations be formulated to address the problems identified?

2. Legal background

The following section explores the legal definitions pertaining to data privacy and Big Data, the state bodies that have responsibility for ensuring data security and individual rights to privacy, and the existing laws in the Russian Federation that govern the responsibilities of data handlers. One particular point to bear in mind is that the existing Russian Federation laws described below were designed to establish and uphold principles of privacy in the handling of personal data, but that these laws[2] are not fully equipped to ensure these principles are upheld in the case of Big Data analysis. These shortcomings are explored further in the Discussion section, and recommendations for updating existing laws are presented in the Conclusions.

2.1. *The concept of Big Data*

Big Data technology allows the collection and processing of large amounts of data, including personal information or information that can identify an individual. In this regard, Russia faces a number of challenges in terms of protecting confidentiality in the provision of services reliant upon personal data such as online purchases, social networks and banking. These challenges include ensuring personal data is secure from theft or leaks, balancing companies' ability to access to personal data with individuals' right to privacy, and ensuring that the protection for anonymity provided by law is extended to data processed using Big Data tools.

No widely accepted definition of Big Data technology exists. In Russia, at the conference "Big Data and Business Intelligence 2012", research manager Alexander Prokhorov pointed out that four Vs determine Big Data: volume, variety, velocity and the numerical values of infrastructure,[3] as does the European Union Agency for Network and Information Security (ENISA).[4] Gartner 2013 defined Big Data in terms of three Vs: volume, variety and velocity. However, by 2015 Gartner did not include Big Data in his report "Gartner's 2015 Hype Cycle for Emerging Technologies Identifies the Computing Innovations that Organizations should Monitor",[5] explaining that the concept

[1] Privacy challenges. Analysis: It's surprisingly easy to identify individuals from credit-card metadata. Available from: http://news.mit.edu/2015/identify-from-credit-card-metadata-0129 [Accessed 17 February 2017].

[2] Federal Law № 149 "On information, information technologies and protection of information", Federal Law № 152-FZ "On Personal Data", etc.

[3] Naidich A, "Bolshiye dannyie: naskolko oni bolshiye? [Big Data: How big it is]" (2012) 12 *Kompyuter Press*. Available from http://compress.ru/article.aspx?id=23469. [Accessed 1 June 2016].

[4] Naydenov R, Liveri D, Dupre L, Chalvatzi E and Skouloudi Ch, "Big Data Security" (European Union Agency for Network and Information Security 2015). doi: 10.2824/13094.

[5] "Gartner's 2015 Hype Cycle for Emerging Technologies identifies the computing innovations that organizations should monitor". Available from http://www.gartner.com/newsroom/id/3114217. [Accessed 1 June 2016.

Please cite this article in press as: Anna Konstantinovna Zharova, Vladimir Mikhailovich Elin, The use of Big Data: A Russian perspective of personal data security, Computer Law & Security Review: The International Journal of Technology Law and Practice (2017), doi: 10.1016/j.clsr.2017.03.025