Available online at www.sciencedirect.com

ScienceDirect

www.compseconline.com/publications/prodclaw.htm

Computer Law & Security Review

Legal aspects of managing Big Data

Richard Kemp *

Kemp IT Law, London, UK

Keywords:
Big data
Information architecture
Data protection
Intellectual property rights
Contract law
Data governance

ABSTRACT

Big Data is shorthand for the currently rapidly evolving techniques of gathering and analysing for competitive advantage vast unstructured and structured sets of digital data. Big Data is currently at an early stage of development, but many organisations will be embarking on Big Data projects in the next couple of years in order to be in a position to know more about their customers than their competitors. Central to the success of these projects will be four critical factors: (i) understanding the legal framework for Big Data and how it applies to the organisation concerned; (ii) effectively bringing together the organisation's IT and legal functions in the Big Data project; (iii) a clear understanding of the organisation's objectives for its Big Data operations; and (iv) a structured approach to the strategy, policy and process aspects of Big Data governance.

1. Introduction

1.1. *'Big Data is everywhere'*

'If you haven't heard' trumpeted the *Financial Times*' Lex column of 27 June 2014, 'Big Data is everywhere'.[1] Over the past twenty years, the bow wave in IT has moved on from hardware and software to the data they process, and in an increasingly competitive and data-centric world, harnessing the tides of the Big Data ocean will confer competitive advantage in enabling a company to know more about its customers and market place than its competitors.

Commenting that the business intelligence and analytics ('**BIA**') software market is worth $16bn a year and growing at 8% a year, the FT Lex column called out research from consultancy Gartner Inc.[2] who showed that the BIA market is currently undergoing an 'accelerated transformation' from retrospective BIA used mainly for measurement and reporting to prospective BIA software used for prediction, forecasting and modelling. This is fuelling a race as the BIA software majors – Oracle, SAP, IBM and SAS, whose combined BIA software turnover totals $10bn – vie with smaller, faster growing BIA specialists like QlikTech, Splunk and Tableau to bridge the gap between the oceans of available Big Data and BIA software's ability to harness Big Data for competitive advantage in a structured, legally compliant way.

It is this race for competitive advantage – knowing more than your competitor not so much about what your customer has just done as about what he or she is likely to do next – that is at the commercial epicentre of Big Data. Yet this is a race that is just beginning: Gartner also points out[3] that only 15% of Fortune 500 companies will be able to exploit Big Data

* Kemp IT Law, 21 Napier Avenue, London SW6 3PS.
E-mail address: richard.kemp@kempitlaw.com.
URL: http://www.kempitlaw.com

[1] http://www.ft.com/cms/s/3/525236ca-fd4f-11e3-bc93-00144feab7de.html?siteedition=uk#axzz35vtpzx2A.
[2] http://www.gartner.com/technology/reprints.do?id=1-1QHKSEP&ct=140206&st=sb.
[3] http://www.gartner.com/technology/topics/big-data.jsp.

http://dx.doi.org/10.1016/j.clsr.2014.07.006

for competitive advantage by the end of next year and that only 8% of companies are currently using Big Data analytics at all.

1.2. What is 'Big Data'?

Commenting that there was no one generally accepted definition of Big Data, the White House's Executive Office of the President in a report dated 1 May 2014[4] nevertheless gave a useful description:

> *Most definitions reflect the growing technological ability to capture, aggregate, and process an ever-greater volume, velocity, and variety of data. In other words, "data is now available faster, has greater coverage and scope, and includes new types of observations and measurements that previously were not available."*[5] *More precisely, big datasets are "large, diverse, complex, longitudinal, and/or distributed datasets generated from instruments, sensors, Internet transactions, email, video, click streams, and/or all other digital sources available today and in the future".*[6]

As used in this White Paper, 'Big Data' is therefore shorthand for the collation, processing, analysis and use of vast exploitable datasets of unstructured and structured digital information. Along with Cloud, mobile[7] and social computing, Big Data is one of the four main drivers of change in information technology as it moves into new areas whose features currently include machine learning, 3D printing, virtual reality, the Internet of Things and nanotechnology.

1.3. The US NIC's December 2012 report

Big Data's direction of travel is well signposted in the December 2012 long range report of the US National Intelligence Council's '*Global Trends 2030: Alternative Worlds*'[8] where it articulates a focus on data solutions and Big Data as a key IT driver over the next two decades:

> *Information technology is entering the big data era. Process power and data storage are becoming almost free; networks and the cloud will provide global access and pervasive services; social media and cybersecurity will be large new markets.*[9]

Opportunities arising through Big Data are not without their challenges and issues however:

> *Since modern data solutions have emerged, big datasets have grown exponentially in size. At the same time, the various building blocks of knowledge discovery, as well as the software tools and best practices available to organizations that handle big datasets, have not kept pace with such growth. As a result, a large — and very rapidly growing — gap exists between the amount of data that organizations can accumulate and organizations' abilities to leverage those data in a way that is useful. Ideally, artificial intelligence, data visualization technologies and organizational best practices will evolve to the point where data solutions ensure that people who need the information get access to the right information at the right time — and don't become overloaded with confusing or irrelevant information.*[10]

It is these challenges and issues that the fast growing BIA software market is seeking to address.

1.4. Scope and aims of this white paper

The main purpose of this paper is to provide a practical, overview of the legal aspects of Big Data management and governance projects. In order to illustrate how Big Data and BIA software are beginning to have real impact and provide context for the discussion that follows, Section 2 briefly overviews Big Data initiatives and potential in a number of different vertical sectors (financial services, insurance, healthcare, air travel, music and public sector). The focus is then on providing three 'views' of Big Data from the legal perspective:

- Section 3 offers a common legal analytical framework for Big Data, centred on intellectual property rights in relation to data, contracting for data and data regulation;
- Section 4 considers Big Data within the organisation from the standpoint of input, processing and output operations; and
- Section 5 overviews the key aspects of Big Data management projects from the perspective of governance, addressing risk assessment, strategy, policy and processes/procedures.

The Legal and the IT Groups are likely to be the two business functions most closely associated with an organisation's Big Data management project. This paper addresses primarily the issues that will be relevant for the Legal Group rather than the IT group, but data modelling is addressed in outline at Sections 2 and 3 in view of its central importance. Detailed discussion of the technical aspects of data law and the detail of Big Data governance is outside the scope of this paper, but

[4] '*Big Data: Seizing Opportunities, Preserving Value*', http://www.whitehouse.gov/issues/technology/big-data-review. The report focuses on 'how big data will transform the way we live and work and alter the relationships between government, citizens, businesses, and consumers'.

[5] Liran Einav and Jonathan Levin, "The Data Revolution and Economic Analysis," Working Paper, No. 19035, *National Bureau of Economic Research*, 2013, http://www.nber.org/papers/w19035; Viktor Mayer-Schonberger and Kenneth Cukier, *Big Data: A Revolution That Will Transform How We Live, Work, and Think*, (Houghton Mifflin Harcourt, 2013).

[6] National Science Foundation, Solicitation 12–499: *Core Techniques and Technologies for Advancing Big Data Science & Engineering (BIGDATA)*, 2012, http://www.nsf.gov/pubs/2012/nsf12499/nsf12499.pdf.

[7] See Kemp, 'Mobile payments: current and emerging regulatory and contracting issues' (29 CLSR [2], pp. 175–179).

[8] http://globaltrends2030.files.wordpress.com/2012/11/global-trends-2030-november2012.pdf.

[9] At page ix.

[10] At page 85.