

# Market research & the ethics of big data

Daniel Nunan\*

*Henley Business School, University of Reading, UK.*

Maria Di Domenico

*Surrey Business School, University of Surrey, UK.*

## **Abstract**

The term ‘big data’ has recently emerged to describe a range of technological and commercial trends enabling the storage and analysis of huge amounts of customer data, such as that generated by social networks and mobile devices. Much of the commercial promise of big data is in the ability to generate valuable insights from collecting new types and volumes of data in ways that were not previously economically viable. At the same time a number of questions have been raised about the implications for individual privacy. This paper explores key perspectives underlying the emergence of big data and considers both the opportunities and ethical challenges raised for market research.

## Introduction

“We never, ever in the history of mankind have had access to so much information so quickly and so easily.” – Vint Cerf, ‘father of the internet’ (quoted in Silva, 2009)

Concerns over privacy and the collection and use of personal information have been closely associated with the growing influence of technology in society (Garson, 1988; Zuboff, 1988). As long ago as 1890 legal scholars raised concerns over the commercial application of new photographic technologies in the newspaper industry. In their now famous paper on the potential impacts of future use of information technologies in commerce, Warren and Brandeis made the prescient observation: “*what is whispered in the closet shall be proclaimed from the house-tops*” (Warren & Brandeis, 1890). These concerns have grown significantly since commercial use of the internet first became widespread in the mid-1990s (Bush et al. 2000). For many people, the question of how personal data is used for marketing purposes has become a defining social feature of the internet (Nissenbaum, 2004). Yet the same technology has also created significant new opportunities for market researchers to collect and analyse information to generate more timely and relevant insights (Christiansen, 2011). To date, these two perspectives have existed side-by-side, albeit sometimes uneasily, with market researchers able to leverage the internet as an important research tool within the framework of existing ethical approaches. However, the trend towards ‘big data’ presents a number of challenges in terms of both the ways that personal information is collected and consumer relationships with this information.

Because of its key role in collecting, analyzing and interpreting data many of the problems, and opportunities, of big data are also those of market research. For market research to prosper it requires the continuing cooperation of respondents (Bednall et al. 2010) both in terms of providing data for research studies and in giving permission for this data to be analysed. In an environment where there are issues around increasing non-cooperation by respondents (Jarvis, 2002; Curtin et al. 2005) it is essential for market

researchers to be at the forefront of understanding emergent ethical and privacy issues. This is even more critical where regulatory change poses a potential threat to market researchers' ability to collect data in the future. On the other hand, progressions offered by big data present significant opportunities for generating new insights into consumer behaviour. The ability to triangulate multiple data sources and perform analyses of these massive data sets in real-time enables market researchers to gather a range of insights that may not be possible using existing market research techniques.

This paper has four parts. Firstly, we present a discussion of the key characteristics of 'big data', highlighting emergent technological, commercial and ethical perspectives. In the second part we consider the ways that big data may impact market research. Thirdly, we discuss the key role that privacy guidelines play in market research. We conclude with future directions for market research ethics based on the changing environment brought about by big data.

### **Defining Big Data**

Big data is a term that has quickly achieved widespread use amongst technologists, researchers, the media and politicians. Perhaps due to the speed of dissemination the use of the term has been rather nebulous in nature. In order to fully explore the role of big data in market research it is first necessary to unpack the meaning of the term. The concept of big data can be framed by one of three perspectives. The first is a response to the technology problems associated with storing, securing and analysing the ever-increasing volumes of data being gathered by organizations. This includes a range of technical innovations, such as new types of database and 'cloud' storage, that enable forms of analysis of that would not have previously been cost effective. The second perspective focuses on the commercial value that can be added to organizations through generating more effective insights from this data. This has emerged through a combination of better technology and greater willingness by consumers to share personal information through web services. The third perspective considers the wider societal impacts of big data, particularly the implications for individual privacy and the effect on

regulation and guidelines for ethical commercial use of this data. We now consider each of these perspectives on big data in more detail.

### *Big Data & Technology Innovation*

In its original form, big data referred to technical issues relating to the large volumes of data being created (Jacobs, 2009). Whilst the rate at which data has been generated by information technology has always been increasing, recent growth produces some startling statistics. Take the following two examples: (1) 90% of all the data in the world has been produced in the last two years (IBM, 2011) and (2) The sum of all information ever produced by humans by 1999, estimated at 16 exabytes (16 *trillion* megabytes) in 1999 will be the same as generated every nine weeks by the world's largest telescope, the Square Kilometre Array, when it opens later this decade (Redfern, 2011). At the same time as data volumes have increased, the cost of storing this information has reduced drastically. For example, in 2011 \$600 would buy a disk drive with the capacity to store all the worlds recorded music (Kelly, 2011). In providing these statistics we seek to highlight that the big data problem is not the volume of data itself, but the issues arising with analysing and storing this data in a way that can be easily accessed. The costs of storing large volumes of data mean that, until recently, it has been common practice to discard information not strictly required for legal, regulatory or immediate business use. For example, hospitals and health care providers discard more than 90% of the data they generate, including nearly all real time video data generated from operations (Gantz, et al. 2007).

In addition, two other factors, *velocity* and *variety*, are significant in big data, (IBM, 2011). Velocity refers to the challenges in accessing stored data quickly enough for it to be useful. For most real world uses, data needs to be accessible in something close to real time. Offering fast access to massive amounts of data at a reasonable cost is a key limitation of existing technologies, both in terms of commonly used relational database software and the use of cheaper 'offline' tape storage devices. Variety refers to the type of information being stored. Previously, data stored tended to be highly structured in nature. By contrast, the types of data that tend to dominate modern data stores are

unstructured, such as streams of data gathered from social media sites, audio, video, organizational memoranda, internal documents, email, organizational Web pages, and comments from customers. (Kuechler, 2007).

From a technology perspective the solution to the big data problem has occurred through the intersection of several innovations. These include flash based disc drives which allow much faster access to high volumes of information and a new generation of non-relational database technologies that make it practical to store and access massive amounts of unstructured data. Fittingly, much of this new database technology has emerged from inside companies who run social networks including Google, Facebook, LinkedIn and Twitter.

### *Big Data & Commercial Value*

Whilst technology has served as the enabler of big data services the broader interest in big data has been driven by thoughts of the potential commercial value that it may bring. This is derived from the ability to generate value from data in ways which were not previously possible. For example, financial services are using high performance computing to identify complex patterns of fraud within unstructured data that were not previously apparent (Economist, 2012). This has enabled the cost-effective provision of financial services in areas that would have previously been regarded too risky to be sustainable. Another example is the use of personal location data, gained from a combination of smartphones, cell-tower tracking and GPS navigation data within vehicles. Such information is already being used to calculate fuel efficient smart-routing, report diagnostic information back to manufacturers or tracking applications to locate family members (Manyika, 2011). In practical terms this gives organizations the ability to generate insights in minutes that might have once taken days or weeks, for example by using diagnostic information to predict quality issues or develop new understanding of consumer behaviour (Taylor, 2012).

The need to derive commercial value, and insight, from data is not new. Indeed, providing information to help support management insights can be considered a

foundation of the market research sector. However, the key difference with big data strategies is not simply the provision of high quality and more timely data into the decision making process but the enablement of continuous autonomous decision-making via the use of automation (Yulinsky, 2012). For example, the use of remote monitors for health conditions such as heart disease or diabetes or ‘chip-on-pill’ technologies could enable the automation of health decisions (Manyika et al., 2011).

### *Big Data & Privacy*

As the collection of unstructured data becomes more economically viable, and shifts in consumer usage of technology make a much wider range of data available, there is an incentive for organisations to collect as much data as possible. Yet, just because consumers are willing to provide data it does not mean that its use is free from privacy implications (Boyd, 2010). Four examples of these privacy challenges follow.

The first arises from different sets of data that would not have previously been considered as having privacy implications concerns being combined in ways that threatens privacy. One example, albeit experimental, was discovered by researchers who used publicly available information and photographs from Facebook and, through application of facial recognition software, matched this information to identify previously anonymous individuals on a major dating site (Acquisti et al., 2011). In another example, anonymous ‘de-identified’ health information distributed between US health providers was found to be traceable back to individuals when modern analytical tools were applied (Ohm, 2009). This creates an unintended use paradox. How can consumers trust organisations with information when the organisation does not yet know how the information might be used in the future?

The second challenge comes from security, specifically the issue around hacking or other forms of unauthorised access. Despite increasing awareness of the need to maintain physical security computer systems are only as strong as their weakest point, and for databases the weakest point is usually human. For all the advanced technical security used to protect the US diplomatic network the Wikileaks scandal was caused by a low

level employee copying data onto a fake 'Lady Gaga' CD (Leigh, 2010). For big data stores to be useful there needs to be a certain amount of regular access, often by a range of employees in different locations. Whilst treating data like gold bullion and storing it in a vault may guarantee security this is not a practical solution for most use cases. But what of security breaches? When a credit card is stolen it is relatively straightforward, if time consuming, to cancel the card and be issued a new one. Yet a comprehensive set of information about one's online activities, friends or any other type of big data set is more difficult to replace. In a sense, these are not simply items of data but a comprehensive picture of a person's identity over time.

The third privacy challenge is that data is increasingly being collected autonomously, independent of human activity. Previously, there was a natural limit on the volume of data collected related to the number of humans on the planet, and the number of variables we are interested in on each individual is considerably less than the number of people on the planet (Jacobs, 2009). The emergence of network-enabled sensors on everything from electricity and water supplies through to airplanes or cars changes this dimension. Combining these sensors with nano-technology it becomes possible to embed large numbers in new buildings to provide early warnings to dangers relating to structural integrity of the building (Saafi, 2010). The volume of data, and the speed with which the data must be analysed means that there is the requirement for data to be collected, and autonomously analysed without an individual providing specific consent. This raises ethical concerns relating to the extent to which organizations can control the collection and analysis of data when there is limited human involvement.

The final privacy challenge relates to the contextual significance of the data. Currently the ability of organizations to collect and store data runs far ahead of their ability to make use of it (Jacobs, 2009). As a function of storing any, and all, unstructured data regardless of potential use cases this means that combinations of data for which there are currently no capabilities to analyse could become subject to privacy breaches in the future.

### **Big Data: Issues for Market Research**

Having presented an analysis of what big data is we now suggest five areas in which big data may have an impact. Each of these represents an opportunity as a source of new insight for market researchers but also has the potential to create significant challenges in terms of privacy and the use of personal data.

### *1. The Social Graph*

Much of the growth in data is driven by the voluntary sharing of information between members of social networks. Rather than focussing on individual responses, big data allows a picture to be built of group level interactions and the nature of the bonds that bring these people together, a concept that has been labelled the 'social graph' (Berners-Lee, 2007). The relationship is symbiotic; in order to create value in their social graph users need to contribute information about their lives, but in doing so they also increase the digital exhaust of information that is available about them. Yet, the boundaries of this social graph are imprecise. The challenge of continuously identifying and labelling 'friends', particularly those where there are weak social ties creates the potential for social uncertainty. Furthermore, the labels for these virtual world connections, such as 'followers' or 'friends', may not be analogous to their physical world meanings. It is this source of ambiguity that presents ethical challenges. Understanding how an individual's online social graph relates to real world meaning is thus likely to be essential in effectively leveraging it.

### *2. Ownership of Data*

With big data the nature of the organizations that collect the largest stores of personal information is changing. In general, it is not central governments or traditional large corporations who are storing information but rather a breed of smaller high-technology firms such as Facebook, Twitter, LinkedIn, Google and others. On the one hand this provides researchers access to sources of data that may not have been previously available. Whilst there is little incentive for governments to monetize their data commercially, the business models of the majority of consumer-facing web services are built around, to put it simply, driving commercial value from customer data. For example, Twitter will now make available a feed of several years' historical content



available to anyone wishing to use it for research or analysis purposes. This raises the questions of long-term ownership of personal data that consumers make available online. Even those companies that do not currently sell access to their data stores could themselves be potentially sold in the future, and policies for use of data change.

### *3. Big Data has memory*

The capability for big data technology to enable the storage, and recall, of large volumes of information gives a temporal dimension to the storage of personal information. Information recorded today, even if not public now, can be recalled instantly in decade's time. For example, the emerging focus of Facebook on a 'timeline' has creates challenges that activities people partake in whilst at college may reflect badly on them when they enter the world of work. Whilst analysing data and building effective models of consumer behaviour has always been a part of market research, big data provides the promise of more accurate and far-reaching models. Thus big data enables the ability to rewind and fast-forward people's lives, but in doing so may remove the ability for individuals to forget and be forgotten.

### *4. Passive Data Collection*

Much information collection is now automatic and passive. Existing approaches to market research are typically reliant on some form of active opt-in. Big data makes use of passive technologies, such as location-based information from mobile phones, data from autonomous sensors, or facial recognition technology in retail stores. This creates the potential for powerful new variables to be included in consumer research. At the same time the individual may no longer have specific knowledge and awareness that data is currently being collected about them. Even if permission has been given initially, these services are not asking for permission every time such contextual data is gathered.

### *5. Respecting privacy in a public world*

Whilst privacy concerns have been raised over the use and creation of big data these have been outpaced by individuals' use of social networks. The value inherent in the social graph provides some form of counterbalance to the potential privacy issues. Put another

way, for all the privacy implications people derive great benefit from services such as mobile applications and social networks – many of which are available at no charge. Beyond this, for many social groups contributing to big data stores becomes socially necessary form of communication in a world where avoiding social networking sites serves the potential to exclude people from their communities. This creates a paradox in that whilst individuals can opt-out of having their personal data collected, to do so may result in increasing their exclusion from the digitally connected world in which they reside.

### **Personal data & privacy in market research**

For many sectors the ability to collect data and turn it into insight has a key role in developing more innovative and successful products and services. However, for market research the importance of access is instrumental to the ability to deliver the product. The history of marketing activity provides us with many examples of situations where regulators have responded reactively to public perceptions of over-zealous, or unethical, marketing activity. From the promotion of ineffective ‘patent’ medicines in the 19<sup>th</sup> Century through to tobacco and alcohol in the 20<sup>th</sup> Century, in sectors that generate negative externalities regulatory pressure is never far behind. Given the criticality of online data collection to market research, and the potential for personal data to become a similarly hot topic of the 21<sup>st</sup> Century, for the successful realization of the potential of big data in market research it is also necessary to be proactive in responding to potential privacy issues, even if these issues have yet to reach the public imagination. A central plank in this is the development and maintenance of an effective self-regulation strategy.

Market research has a strong tradition of proactive development of ethical standards with the first ESOMAR code of Marketing & Research Practice being published in 1948 (ESOMAR, 2008) and the MRS publishing its first self-regulatory code in 1954 (MRS, 2006). These guidelines have become widely adopted by local market research associations and professionals around the world, and have been frequently updated to take account of the changing social and technological environments in which market research operates.

Intertwined with the development of these standards, the argument for ethical market research has been supported by recognition that good ethical practice and data quality are interlinked (Tybout & Zaltman, 1974). Table 1. outlines the key section of the current ESOMAR code relating to privacy.

*Table 1 goes about here.*

There is a theme in the market research literature around understanding the potential ethical and privacy challenges of new technology, such as those around virtual ethnography (Hair & Clarke, 2007) or the growth of Web 2.0 research (Cooke, 2008). However, developments in technology highlight the challenge of maintaining appropriate ethical codes in an environment where the ways consumers use technology are quickly changing. One example of this can be seen in the question of how to determine informed consent in an environment where the collection of data may be both passive and autonomous. Recent ESOMAR guidelines (ESOMAR, 2011) on passive data collection have attempted to answer this question for scenarios around video collection, such as from CCTV.

*“There may be instances in public places where informed consent from individuals is impossible to achieve. In those cases public notice should be given about the data collection.” (ibid:4)*

However, given that big data often disintermediates the collection from analysis of data the challenge is raised over what consent is being sought, given that the purpose of data collection may not be known. Additionally, given the increase and widespread use of autonomous data collection, such as through sensors, there are many situations where it is simply not practical to gather informed consent – potentially hundreds of times a day. One limited example of this can be seen in the recent EU legislative changes around cookies, requiring some form of implied informed consent for the setting of website

cookies. This is an issue covered by both the ESOMAR Code of Conduct (ESOMAR, 2008) and the MRS Code (MRS, 2006). ESOMAR guidelines on “Conducting Market and opinion research using the internet” confirm the basic principles of voluntary cooperation, disclosure of researcher identity and safeguarding of respondent’s anonymity. More specifically, in relation to cookies:

*“Respondents must always be told when cookies (or small text files that will ensure that they won’t be interviewed again, for example) or other covert software is being used to collect information about them are being used and that they can turn them off or remove them.” (ibid: 13)*

Similarly, the MRS guidelines (MRS, 2006) emphasizes the need to be clear about the purposes behind gathering information via cookies and gain consent for their use:

*“Cookies & invisible processing – In accordance with the Privacy and Electronic Communications Regulations, cookies or similar devices shall not be used unless the subscriber or user of the relevant terminal equipment is:*

- a. provided with clear and comprehensive information about the purposes of the storage of, or access to, that information; and*
- b. given the opportunity to refuse the storage of, or access to, that information.*

The guidelines are similar, in spirit at least, to the legislative changes that they predate. However, there have also been considerable confusion and challenges in implementing these changes in a way that does not limit the ability of firms to carry out commercial activities online (ICO, 2011). For example, for many sites dependent on advertising the specific cookie that is set is not-known until the page is loaded (Arthur, 2012). Related to this is the question of whether, even when individuals provide consent to online services, they are actually aware of the type and range of data that is being collected (Traung, 2010).

## **Future Directions for Market Research Ethics Guidelines**

The potential for big data to change the landscape of market research is encapsulated by Lewis (2012:11):

“Corporate market researchers believe that the leading agency in 2020 is just as likely to be Google, Facebook or a company from outside the industry as it is to be one of the ‘old guard’.”

Such trends, particularly the growth in passive and remote data collection, call into question some of the implicit safeguards present when someone has given permission for market research to be undertaken and is aware of the boundaries under which data collection is taking place. Furthermore big data is built upon the use of unstructured data which, by definition, is collected without necessarily having knowledge of the purpose to which it will be put in the future. We therefore suggest three ethical practices that would enable individuals to maintain control over their privacy, whilst still enabling firms to provide the services from which consumers gain so much benefit. For each of these we also highlight some of the current challenges with implementation.

### **1. The right to be forgotten.**

Individuals can request that data held about them on social networking sites that might be used for market research purposes in the future can be deleted. This question of long-term ownership of personal data currently forms part of a number of European legal challenges (Falkenrath, 2012), raising questions over where the line over personal data can be drawn. For example, a case in Spain centering on individuals requesting that material about them not show up in search engines raises questions about both privacy and freedom of speech.

### **2. The right to data expiry**

In addition to a general ‘right to be forgotten’, unstructured data held about individuals can be expired after a set period of time if it is of no commercial use. This creates questions over defining both commercial use and unstructured data.

However, if big data involves the commercial collection of data without regard to its potential use, or potentially without knowledge that it is being collected in the case of autonomous sensors, then there must also be a safeguard that data not used will be destroyed.

### **3. Ownership of a social graph**

Much of the value of big data is help in the social graph where information can be collated about an individual as part of their social graph without their knowledge. For example, Person A can take a photo of Person B, which is then tagged by Person C on a social networking site. Allowing individuals ownership of their social graph – information that references them and their relationships with others – prevents wider and potentially inadvertent misuse of personal data. Challenges exist in identifying and verifying individuals as well as those relating to the questions of ownership of data by third parties, for example when dealing with children.

We outline these suggestions as much as a means of creating discussion as of providing prescriptive solutions. We recognize that these suggestions could be interpreted as threatening the commercial potential of many services that generate big data. Yet this is only the case if we view big data as an exercise in unrestrained collection and analysis of data by organizations. Above all, for the commercial promise of big data to be delivered it relies on trust. Without this trust in place, organizations face a pushback from both consumers and regulators.

### **Conclusion**

In this paper we have introduced the concept of big data and its impact on notions of privacy. We have also identified the potential opportunities for market research together with the challenges for privacy, challenges which we do not believe are fully encapsulated within existing ethics codes. Ethics and market research remain grounded in a model of the technology that is at odds with the multi-device, socially-based adoption that is likely to form the basis of future internet growth. For example, 82% of internet

users over the age of 15 now use social networking sites (Comscore, 2011). The combination of smartphones and mobile internet have enabled consumers to access services nearly everywhere they go, leaving behind them a 'digital exhaust' of personal data. Likewise, the growth in autonomous networked sensors in cars or smart-meters in the home means that huge volumes of data are increasingly generated independent of human action.

Big data presents a convergence of both technical and strategic capabilities that provides significant potential for organizations of all sizes to generate value from the data they store. There is the risk that market research will be 'left out of the loop' as organizations strive for the commercial benefits brought by big data without consideration of the needs for appropriate consideration of personal privacy. However, with appropriate engagement there are significant opportunities for market researchers, particularly as one of the key stumbling blocks for the adoption of big data strategies is the lack of sufficiently qualified staff with the necessary analytical and research skills (Manyika et al., 2011; Woods, 2011). Yet big data is more than just a technical phenomenon. As we have outlined in this paper, with big data comes the possibility of significantly changing the relationship that individuals have with data collected about them. A failure to appreciate this changing relationship risks a political and regulatory environment that limit opportunities for almost any kind of online data collection and analysis, with knock-on effects for market researchers. To realise the potential of big data, and maintain influence, we believe that it is essential for market researchers to engage with and debate these important ethical questions.

## References

Acquisti, A., Gross, R. and Stutzman. (2011) Faces of Facebook: Privacy in the Age of Augmented Reality. *Presented at BlackHat Conference Las Vegas*, August 4, 2011.

Arthur, C. (2012). Cookies law changed at 11th hour to introduce 'implied consent'. *The Guardian*, May 26, 2012.

Bednall, D., Adam, S. and Plocinski, K. (2010) Using compliance techniques to boost telephone response rates. *International Journal of Market Research*, 52(2), pp. 155-168.

Berners-Lee, T. (2007) Giant Global Graph. *Available at:*  
<http://dig.csail.mit.edu/breadcrumbs/node/215> (Last accessed 15th February 2012).

Bush, V., Venable, B. and Bush, A. (2000) Ethics and Marketing on the Internet: Practitioners' Perceptions of Societal, Industry and Company Concerns. *Journal of Business Ethics* 23, pp. 237–248.

Boyd, D. (2010) Privacy and Publicity in the Context of Big Data. *Presented at WWW conference* Raleigh, North Carolina, April 29 2010

Cooke, M and Buckley, N (2008) Web 2.0, social networks and the future of market research. *International Journal of Market Research* 50(2), 267-292

Christiansen, L. (2011). Personal privacy and Internet marketing: An impossible conflict or a marriage made in heaven? *Business Horizons*, November-December, pp. 509-514.

Comscore. (2011) It's a Social World: Top 10 Need-to-Knows About Social Networking and Where It's Headed. *Available at:* [http://www.comscore.com/it\\_is\\_a\\_social\\_world](http://www.comscore.com/it_is_a_social_world) (last accessed 13th February 2012).



Curtin, R., Presser, S. & Singer, E. (2005) Changes in Telephone Survey Nonresponse over the Past Quarter Century. *Public Opinion Quarterly*, **69**, pp.87-98.

Economist. (2012) Crunching the numbers; Big data. *The Economist*, 19 May, 2012.

ESOMAR. (2008) ICC/ESOMAR International Code on Market & Social Research. Available at: <http://www.esomar.org/knowledge-and-standards/codes-and-guidelines.php>

ESOMAR. (2011) Passive Data Collection, Observation and Recording. Available at: <http://www.esomar.org/knowledge-and-standards/codes-and-guidelines.php>

Falkenrath, R. (2012) Google must remember our right to be forgotten. *Financial Times*, February 15.

Gantz, J., Reinsel, D., Chute, C., Schlichting, W., McArthur, J., Minton, S. Xheneti, I., Toncheva, A. and Manfrediz, A. (2007) The expanding digital universe. *IDC white paper, sponsored by EMC*, March.

Garson, B. (1988) *The electronic sweatshop: How computers are transforming the office of the future into the factory of the past*. Simon & Schuster, New York.

Hair, N. and Clarke, M. (2008). Ethical Dilemmas and Challenges of Ethnographic Research in Electronic Communities. *International Journal of Market Research*, 49(6).

IBM. (2011) Bringing big data to the enterprise: what is big data? Available at: <http://www.ibm.com/software/data/bigdata/> (accessed 15<sup>th</sup> February 2012).

ICO. (2011) Cookies Regulations and the New EU Cookie Law. Information Commissioners Office, (available at:

[http://www.ico.gov.uk/for\\_organisations/privacy\\_and\\_electronic\\_communications/the\\_guide/cookies.aspx](http://www.ico.gov.uk/for_organisations/privacy_and_electronic_communications/the_guide/cookies.aspx)) accessed 20 Feb 2012.

Jacobs, A. (2009) Pathologies of Big Data. *ACM Queue*, 7(6), July.

Jarvis, S. (2002) CMOR finds survey refusal rate still rising. *Marketing News*, February 4, p. 4.

Kelly, K. (2011) Everything. *Web 2.0 Expo and Conference*, March 29, 2011. Available at: [www.web2expo.com/webexsf2011/public/schedule/proceedings](http://www.web2expo.com/webexsf2011/public/schedule/proceedings).

Kuechler, W. (2007) Business applications of unstructured text, *Communications of the ACM*, 50(10), October.

Leigh, D. (2010) How 250,000 US embassy cables were leaked. *The Guardian*, 28 November 2010.

Lewis, I. (2012). The future of market research. *International Journal of Market Research*, 54(1), pp. 11-13.

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R, and Hung Byers, A. (2011) Big data: The next frontier for innovation, competition and productivity. *McKinsey Global Institute Report*, (available at: [http://www.mckinsey.com/Insights/MGI/Research/Technology\\_and\\_Innovation/Big\\_data\\_The\\_next\\_frontier\\_for\\_innovation](http://www.mckinsey.com/Insights/MGI/Research/Technology_and_Innovation/Big_data_The_next_frontier_for_innovation)).

MRS. (2006) Market Research Society Code of Conduct. Available at: <http://www.mrs.org.uk/standards/guidelines.htm>.

Nissenbaum, H. (2004) Privacy as Contextual Integrity. *Washington Law Review*, 79(1).

Ohm, P. (2010) Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization. *UCLA Law Review*, 57, pp. 1701-1777.

Redfern, M. (2011) World's biggest radio telescope, Square Kilometre Array, *BBC News Online*, (available at <http://www.bbc.co.uk/news/science-environment-12891215>), last accessed 23<sup>rd</sup> February 2012.

Saafi, M. (2010) Wireless and embedded nanotechnology-based systems for structural integrity monitoring of civil structures: a feasibility study. *International Journal of Materials and Structural Integrity*, 4 (1).

Silva, D. (2009) The internet has only just begun, say founders. *PhysOrg*, (available at <http://phys.org/news159644537.html>), last accessed 14<sup>th</sup> July 2012.

Taylor, P. (2012) Crunch time for big data. *Financial Times*, 19<sup>th</sup> June 2012.

Traung, P. (2010) EU Law on Spyware, Web Bugs, Cookies, etc., Revisited: Article 5 of the Directive on Privacy and Electronic Communications. *Business Law Review*, 31(10), pp. 216-218.

Tybout, A. and Zaltman, G. (1974) Ethics in Marketing Research: Their Practical Relevance. *Journal of Marketing Research*, 11(4), pp. 357-358.

Warren, S. and Brandeis, L. (1890) The right to privacy. *Harvard Law Review*, IV(5).

Woods, D. (2011) Curing the Big Data Storage Fetish. *Forbes*, 12 July 2011, (available at <http://www.forbes.com/sites/danwoods/2011/12/07/curing-the-big-data-storage-fetish> ), last accessed 20th February 2012.

Yulinsky, C. (2012) Decisions, decisions...will 'Big Data' have 'Big' impact? *Financial Times*, 24 January, 2012.

Zuboff, S. (1988) *In the Age of the Smart Machine: Machine: The Future of Work and Power*. Basic Books, NY.