Business Ethics and Big Data

Big Data is one of the more frequent topics in current business discussions. Defined by some as the ‘new oil’ and regarded as a dangerous threat by others, Big Data presents a number of grey areas that might lead to potential ethical lapses. Whilst protection of privacy is seen as the main concern by many, it is certainly not the only one, and estimates predict that 50 percent of business ethics violations will occur through improper use of big data analytics by 2018.

This more than usual in depth IBE Briefing seeks to bring some clarity to the ethical issues related to Big Data, defining why they are relevant to companies and where the main ethical risks might lie. It also provides a set of questions that can help ethics practitioners liaise with their colleagues and make sure that their organisation is living up to its values when dealing with Big Data.

What is Big Data?

Whilst the term Big Data has become widespread only recently, the acquisition and use of aggregated information is not a new phenomenon. Data collection and analysis have been common practice for centuries and have played a crucial role in the development of society, since the first censuses were taken to determine population-wide characteristics and needs.

However, the increase in processing power and the decrease of computation and storage costs have changed today’s data landscape. In addition, the growing number of sensor technologies, embedded in devices of all kinds, and the widespread use of the internet in daily life, has expanded exponentially the amount of information available and the different ways in which it is collected (Figure 1).

Even though there might be no universally accepted definition of Big Data, the term generally refers to the increased complexities in the use of data. More precisely, it can be described as the product of the following elements.

1. THE ‘3Vs’

Volume: a vast amount of data is generated every second. It has been estimated that every minute of the day, Google receives over 4,000,000 search queries, Facebook users share 2,460,000 pieces of content and YouTube users upload 72 hours of new videos.

 Variety: many different types of data are now available. New technologies can analyse data from different sources such as messages, photos or video recordings, social media or sensor data. Data collected through all these sources can either be categorised as ‘structured’

---

1 A Pols Data is the New Oil, Privacy is the New Green (12/06/2015)
2 TechWeek Gartner: Big Data Could Put Your Business At Risk (07/10/2015)
3 The Wall Street Journal Big Data’s Big Problem: Little Talent (29/04/2012). Despite the great popularity of ‘Big Data’ in the business world, the author of this article points out that a widespread lack of understanding of the issue still persists.
4 Aci The Data Explosion in 2014 Minute by Minute – Infographic (12/07/2014)
5 B Marr Big Data: Using Smart Big Data, Analytics and Metrics to Make Better Decisions and Improve Performance Wiley (01/02/15)
Big Data for Business: potential and concerns

"Making the world more intelligent, identifying patterns undetected before, taking decisions not based on the limited experts’ knowledge but on the huge mass of data from the inscrutable reality. This is the promise of Big Data."9

In a time of digital revolution, information represents a valuable asset. Thanks to rapid developments in technologies, companies worldwide have increasingly come to rely on Big Data in order to solve problems, develop new strategies and target their messages and products. Recognising the importance of this, some companies have created a specific role to link Big Data with their organisation's strategy.10

Moreover, it is increasingly acknowledged that the availability and use of such datasets can have benefits for both customers and society as a whole. Box 1 gives some examples of the potential positive impact of Big Data.

On the other hand, the public has developed a greater awareness and sensitivity towards the topic, driven by the increasingly prominent role of the IoT in everyone's lives. The IoT, in particular, has made it possible for companies to collect data in ways that might not be fully understood by users (e.g. from mobile phone calls or public transport travel passes). As a result, some feel constantly under the scrutiny of a 'Big Brother' that serves the economic interests of businesses and over which they have little or no control.11 This perception has produced what has been labelled as a ‘data trust deficit’: research shows that the public trust in companies to use data appropriately is lower than trust generally. This can negatively affect the reputation of companies or whole industries, with media, internet, telecommunication and insurance companies being particularly affected.12

6 ‘Reinventing society in the wake of Big Data. A conversation with Alex (Sandy) Pentland’ Edge (30/08/2012).
7 The Guardian Defining the internet of things – time to focus on the data (06/11/2014)
8 U Shafaque et al Algorithm and Approaches to Handle Big Data International Journal of Computer Applications (2014)
9 R Smolan and C Kucklick Der vermessene Mensch GEO (2013) p.85
10 As an example, Vodafone has advertised the position of Head of Big Data for Business Development, whose purpose is to “Develop, commercialise, and continuously improve a competitively advantaged Big Data strategy, driving adoption, innovation, and business outcomes throughout the operating countries. This includes the internal value realisation and external monetisation Big Data across Commercial and Enterprise businesses.”
11 BBC When Big Data becomes Big Brother (05/06/2015)
12 Royal Statistical Society Public attitudes to the use and sharing of their data: research for the Royal Statistical Society by Ipsos MORI (2014)
Box 1 **Big Data used to promote innovation and sustainable development.**

**Big Data as a source of innovation** — An example of innovation driven by ‘Big Data sharing’ is provided by the pharmaceutical company GlaxoSmithKline. Building on its commitment to trial transparency, GSK has developed an online portal where researchers can request access to the underlying data of these trials to help further their own experiments.13

The company’s medical policy director, Rob Frost, who oversees the ethical conduct of the company’s research, acknowledged that there were some issues in the process. Legitimate questions around patient privacy and data protection were raised, while the need of ensuring that the data is put to appropriate use for good science emerged. “It is a very real issue. It’s about sharing in an appropriate way — we are sharing anonymised data in a secure portal through which researchers can access the data, but not download the results. Explaining those safeguards to people within the company and outside is an important part of convincing people why this is the right way to do it.” He added that all requests are reviewed by an independent panel.

“Hopefully the winners will be the patients,” Frost said. “Good science demands data scrutiny; we can’t possibly do this on our own and if any part of that network works on their own, it’s not going to lead to the best science. Working together and sharing data is in the best interests of the patients — if you can provide the appropriate safeguards.”14

**Big Data as an enabler of development** – In the past few years, there has been an exponential increase in the use of digital communication technologies in low and middle-income countries, from mobile phones to the internet. The digital traces from this use have been identified by policymakers and researchers as a potential solution to the lack of reliable local statistical data and could help inform policy and interventions.15

A relevant example is the ‘Data for Development (D4D) Challenge’, a project run by the mobile operator Orange. It involved the controlled release of call records from Orange’s subsidiaries in the Ivory Coast and Senegal to researchers, allowing them to use the data to address development issues in an innovative way.16

These concerns were recognised by the UN High-Level Panel, appointed in 2012 to advise on the global development agenda after 2015. The Panel called for a ‘Data Revolution’, expressing the need to change the current situation to improve how data is produced and used around the world - particularly regarding analysis capacity, the closure of data gaps to prevent discrimination and privacy violations, the promotion of transparency and accountability and the development of new targets and indicators.17

The Leadership Conference on Civil and Human Rights also indentified some critical areas that require particular attention (see Box 2).

## Some ethical issues

Discussions the IBE held with ethics practitioners and policy experts highlighted a number of new challenges that companies need to consider in living up to their core ethical values and building public trust. Companies’ commitment to respect personal information, to be open and transparent in their communications, to behave with integrity in all their dealings, to be trustworthy and to treat their customers fairly are all values that can find new applications in the era of Big Data. Some specific issues that responsible businesses need to address to fulfil their duty of care towards customers and other stakeholders are now considered.

### RESPECT FOR PERSONAL INFORMATION

One of the main challenges that businesses have to face is related to the ethical dilemma between their desire to collect and use data to improve performance and customer services, and their commitment to respect the privacy of stakeholders.

---

13 The online portal developed by GSK is available at the following address: [DataRequest.com](#).
14 Rob Frost presented the project at the conference [Data for Humanity](#), held on 11 May 2015 by The Crowd, a platform for the business community to share ideas and innovations.
15 **L Taylor, R Schroeder** *Is bigger better? The emergence of Big Data as a tool for international development policy* Springer Science (2014)
16 For more information, see [Challenge 4 Development](#) and IBE [Data for Development Senegal: Report of the External Review Panel](#) (April 2015).
Box 2 How to protect human rights in the ‘Era of Big Data’.18

Stop High-Tech Profiling. New surveillance tools and data gathering techniques that can assemble detailed information about any person or group create a heightened risk of profiling and discrimination. Clear limitations and robust audit mechanisms are necessary to make sure that if these tools are used it is in a responsible and equitable way.

Ensure fairness in Automated Decisions. Computerised decision-making in areas such as employment, health, education, and lending must be judged by its impact on real people, must operate fairly for all communities, and in particular must protect the interests of those that are disadvantaged. Independent review and other remedies may be necessary to assure that a system works fairly.

Respect the Law. Laws and regulations need to be complied with always. Big Data misuse must be prevented and core legal protections, including those of privacy and freedom of association, have to be promoted.

Enhance Individual Control of Personal Information. Personal information that is known to a corporation, such as the moment-to-moment record of a person’s movements or communications, is sensitive and can potentially damage some groups of people. Individuals should have meaningful control over how a corporation gathers data from them, and how it uses and shares that data.

Protect People from Inaccurate Data. Government and corporate databases must allow everyone to appropriately ensure the accuracy of personal information that is used to make important decisions about them. This requires disclosure of the underlying data, and the right to correct it when inaccurate.

The right to privacy, which allows people to limit who has access to their personal information, is a fundamental human right and, as such, companies must be committed to respect it. However, its application poses some concerns.19 When customers use particular services provided by a company, they are required to trust the organisation with their data, but often they have little insight into how information about them is being collected, analysed and used.

Once a dataset is shared, for instance, it is very hard to control its diffusion and to predict how it will be used. Different datasets could be merged and linked together, for example, providing extremely detailed information about individuals or groups. This introduces ethical grey areas around privacy.

Another important issue is data anonymisation. This is an important tool to protect privacy as it involves either encrypting or removing personally identifiable information from datasets. While there are efforts underway to ensure anonymisation is effective, there is still room for improvement and often the anonymisation cannot be guaranteed.20

Privacy issues can be grouped in three main clusters.

Customer profiling
The information collected by some organisations allows them to create profiles of their customers. While this is predominantly used for marketing purposes, it can also be used in ways that determine personal attributes about a consumer such as their ability to pay for a certain good/service, or influence their opportunities to find housing, forecast job security or influence access to credit.21 Often, individuals have little recourse to understand or contest the information that has been gathered about them or what that data, after analysis, suggests.

One such example involved the retailing company Target in the US. Thanks to Big Data analysis, the company was able to predict specific events in the life of its consumers, such as the birth of a child, based on changing spending habits. They were then able to target certain products to certain audiences, such as expectant mothers. One person targeted in this way was a teenage girl in Minneapolis, whose family were...

18 The Leadership Conference on Civil and Human Rights Civil Rights Principles for the Era of Big Data
19 Royal Statistical Society Public attitudes to the use and sharing of their data: research for the Royal Statistical Society by Ipsos MORI (2014). The Royal Statistical Society sees real potential benefits from data-sharing, however there are also public concerns about the use of data.
21 Frank Pasquale The Black Box Society: The Secret Algorithm Behind Money and Information Harvard University Press (2014)
unaware of her pregnancy and who found out as a consequence of Target’s approach. The company declined to comment on the specific situation, but numerous questions were raised about Target’s conduct.22

**Group Privacy**

The issue of group privacy is also of concern. When used to analyse large groups of people, the information that Big Data can reveal may be hugely beneficial. Examples include the possibility of tracking the spread of a disease more quickly, or bringing relief to a disaster zone more effectively.

However, there can also be downsides which require consideration, especially when operating in countries with limited regulation and potentially weak government. Datasets could easily be acquired by companies with ethically questionable marketing strategies, or political groups wanting to use the information to target specific sets of people.23

These privacy issues can only be magnified by the spread of the IoT particularly in low and middle income countries, which are generally less technologically advanced and might have less reliable privacy protection systems. This may particularly be the case in Africa, which has seen an exponential rise in the use of digital communication technologies and especially of mobile phones as users have embraced mobile communications to overcome a weak or non-existent landline infrastructure.24

**Data security**

A critical issue closely linked with privacy is the security of personal data and how companies make sure that their databases are protected from unauthorised users. Appropriate security mechanisms are essential to promote trust in business: customers and other stakeholder groups need to be assured that the information they provide is safely and confidentially stored. It is worth noting that threats to data security might be both external and internal, and the risk of misuse by employees of the company’s information should not be underestimated.

In the past few years, this topic has come to public attention with some well publicised cases of data security violation which have shown the significant impact of corporate data breaches on individuals. For example, in July 2015, a group called ‘The Impact Team’ hacked the database of Ashley Madison, a dating website for extramarital affairs. The group copied personal information about the site’s user base, including real names, home addresses, search history and credit card transaction records, and threatened to release users’ names and personally identifying information if Ashley Madison was not immediately shut down. Although this cyber attack was aimed at preventing what were considered ethically questionable activities, it was a violation of people’s right to privacy and the company was accused of not taking data protection seriously.25

**INFORMED CONSENT AND OPENNESS WITH INFORMATION**

How informed consent to process personal information is obtained from users is another critical issue. Traditional methods of data collection require the explicit consent of respondents, stating clearly the purpose and objectives of the data collection. The advent of the IoT has challenged this approach, blurring the borders of what can be considered informed consent to the use of personal data.

In the UK the ability of organisations and researchers to use such data is limited by the Data Protection Act 1998. This states that consent must be obtained from individuals before their data can be used for research or commercial purposes.26 The primary method for obtaining consent, especially on social media platforms, is by asking users to agree to terms and conditions when they register to use the service. However, research shows that “Signing social media platforms’

---

22 The New York Times *[How Companies Learn Your Secrets]* (16/02/2012)

23 This issue emerged as a particularly material one within the project “Data for Development”, promoted by the telecommunication company Orange. An example that emerged from the project is related to the risk that armed groups might use the information available to plan their military strategy. For further information, see IBE *[Data for Development Senegal: Report of the External Review Panel]* (April 2015).

24 The Guardian *[Internet use on mobile phones in Africa predicted to increase 20-fold]* (05/06/2014)

25 Business Insider *[Extramarital affair website Ashley Madison has been hacked and attackers are threatening to leak data online]* (20/07/2015).

According to the BBC, the release of such sensitive information had strong consequences for some of the people involved and at least two people are believed to have committed suicide as a result.

26 Data Protection Act (1998)
terms and conditions does not necessarily correlate to informed consent, as research has shown that users sign these complicated documents without reading them in order to open their accounts".  

A recent example of this issue around consent is provided by the social experiment that Facebook undertook on 700,000 of its users to determine whether the company could alter their emotional state. The experiment was carried out without informing these users and was designed to assess whether more positive or negative comments in a Facebook newsfeed would impact how they updated their own page. Facebook used an algorithm to filter content. Researchers found those shown more negative comments responded by posting more negative comments and vice versa. The company stated that the experiment was legitimate as all users had to tick a box agreeing to their terms and conditions, including consenting to "internal operations, including troubleshooting, data analysis, testing, research and service improvement". However, the criticism sparked by the initiative forced Facebook to apologise and to produce more transparent guidelines on the matter.  

FAIR TREATMENT OF STAKEHOLDERS AND INTEGRITY OF BIG DATA  
The reliability of Big Data and the ability of algorithms to generate valid conclusions are matters of debate. Whereas traditional statistical methodologies rely on samples that are chosen to be representative of the whole population being analysed, the new datasets produced by Big Data might not be statistically accurate and therefore could produce flawed results. For this reason, a fourth ‘V’ of Big Data is often added to Volume, Variety and Velocity, listed at the beginning of this Briefing: it is Veracity, which refers to the trustworthy and integrity of data. When the veracity of a dataset can’t be guaranteed, significant issues might arise. Some individuals or groups might accidentally be accorded more visibility and thus be favoured, or discriminated against, at the expense of those less visible. The access and the ability to use new information and technologies vary between individuals, communities and countries. Similar disparities, often known as ‘digital divide’, might produce inequalities in opportunities and outcomes.  

The city of Boston faced this type of issue when implementing ‘Street Bump’, a mobile application that used a smartphone accelerometer and GPS feed to collect data about road conditions such as potholes. The public were encouraged to download and use the app to report potholes to the city’s Public Works Department. However, as some groups of people (e.g. individuals on low income and the elderly) were less likely to own a smartphone or download the app, information from these groups was not being recorded. Repair services were therefore being concentrated in wealthier neighbourhoods. The city of Boston solved the problem by completing the dataset using other sources not subject to this bias, such as reports from city-roads inspectors and other more traditional channels.  

The Ethics Test  
QUESTIONS FOR ETHICS AND COMPLIANCE PRACTITIONERS  

Given the increasing importance these issues have for business, more structured forms of governance of Big Data appear necessary. During a workshop organised by the Royal Statistical Society in November 2015 to discuss the opportunities and ethical risks of Big Data participants stressed the need for data governance to minimise harm and maximise benefits from the use of Big Data. They also emphasised the inclusion of considerations of risk and risk management. Additionally, it was also pointed out that legal terms and conditions are more geared toward resolving corporate liability than addressing public understanding.  

These observations highlight potential new opportunities for the Ethics Function in companies, which could hold an oversight responsibility on the ethical aspects of Big Data collection and use.

27 Beninger et al Research using Social Media; Users’ Views NatCen Social Research (February 2014)  
28 The Guardian Facebook sorry – almost – for secret psychological experiment on users (02/10/15)  
30 See the website that the city of Boston launched to promote the initiative.  
31 Royal Statistical Society The Opportunities and Ethics of Big Data, Workshop Report February 2016  


In order to see what this role might involve, this Briefing now provides some questions which Ethics Professionals can usefully ask themselves.

*Do we know how the company uses Big Data and to what extent it is integrated into strategic planning?*
Knowing clearly for what purpose the data will be used is important both to make the most of this resource and to identify the critical issues that may arise. Moreover, research has found that public support increases if the context for data use is explained and people are able to deliberate on it. As it represents a particularly sensitive area, Ethics Officers should make sure they are aware and up to date with what is happening on Big Data within their organisations.

*Do we send a privacy notice when we collect personal data? Is it written in a clear and accessible language which allows users to give a truly informed consent?*
When customers or other stakeholders are required to provide personal information, having terms and conditions that state clearly how and to what extent the data will be used is an important first step in protecting privacy. In particular, it is advisable to be careful with the small print and/or sensitive information. A customer’s perception of being ‘blackmailed’ and forced to agree to conditions they do not fully understand can have a strongly negative impact on trust and reputation. The Privacy Notice Code of Practice issued by the Information Commissioner, the regulator in the UK, provides guidance on the matter.

*Does my organisation assess the risks linked to Big Data?*
It is important that companies develop a structured and methodical approach to assessing the risks associated with Big Data. Identifying any possible outstanding negative impact that the use of Big Data might have on some groups of people, who may be the most vulnerable amongst the company’s stakeholders and what might happen if the datasets become public, can increase awareness of the potential damage that may result from a data breach. Consequently, appropriate mechanisms can be put in place that could help prevent such potential outcomes. Before sharing data, it could be useful to map a company’s suppliers and other stakeholders in order to identify the most vulnerable links. Moreover, a privacy impact assessment may be advisable, as highlighted by the Information Commissioner’s Office (ICO) in its Code of Practice on this subject.

*Does my organisation have any safeguard mechanisms in place to mitigate these risks?*
Evidence shows that having preventative processes in place to enhance data security and protection is an effective way to promote trust. Customers are particularly sensitive about anonymity, and companies should adopt a method of anonymisation which, while allowing the company to make information derived from personal data available in a form that is rich and usable, clearly protects individual data subjects. Again, the ICO has issued a Code of Practice to help companies on this point. Providing people with the ability to opt-out, harsh penalties on data misuse and control on data access can also make a difference.

*Do we make sure that the tools to manage these risks are effective and measure outcome?*
Audit has a key role to play in helping companies deal with these issues. The ICO can undertake a consensual audit across the public and private sector to assess the processing of personal information, and provides practical advice on how organisations can improve the way they deal with it.

*Do we conduct appropriate due diligence when sharing or acquiring data from third parties?*
People expect organisations to share their personal data where it’s necessary to provide them with the services they want and to prevent certain risks. To do so, companies rely on different types of disclosure, involving very complex information chains that cross organisational and even national boundaries. Often companies rely on third parties to collect and acquire the data they need. It is important that due diligence procedures are in place when buying information in the same way as they are for the other kinds of goods and

---

32 Economic and Social Research Council [Public dialogues on using administrative data](https://www.esrc.ac.uk/research-overview/research-impact/public-dialogues-on-using-administrative-data/) (2014)
33 Information Commissioner’s Office (ICO) [Privacy Notice Code of Practice](https://www.ico.org.uk/privacy/advice-and-code-of-practice/privacy-notice-code-of-practice/)
37 Information Commissioner’s Office (ICO) [Auditing data protection a guide to ICO data protection audits](https://www.ico.org.uk/privacy/advice-and-code-of-practice/data-protection-audit/)
services. Companies should ensure that their suppliers uphold similar ethical standards, and guarantee the transparency and accountability of these practices. Moreover, in circumstances where companies are required to share data with business partners or other organisations, it is important that the company lives up to its commitment to integrity. Some useful guidance on this aspect is provided by the Data Sharing Code of Practice issued by the ICO.  

**Conclusion**

It is important that companies realise that they have a responsibility to promote transparency and prevent misuse of personal data.

The consequences and repercussions of questionable ethical conduct when dealing with Big Data can be significant and affect a company’s reputation, customer relationships and ultimately revenues. Even the perception of unethical data handling has the power to undermine both internal and external trust.

In a fast growing and fairly new regulatory area, it can be difficult for business to determine the right approach and define responsibilities. Some internationally recognised standards do exist (namely the ISO/IEC 27001 on Information Security Management), and can provide some guidelines and assistance to organisations seeking to deal with these issues in their code of ethics or internal policies. It is also worth noting that the European Commission intends to strengthen and unify data protection for individuals within the European Union (EU) through the General Data Protection Regulation (GDPR). This regulation, which also addresses export of personal data outside the EU, was formally adopted by the EU Council and Parliament in April 2016 and will take effect after a two-year transition period. Nevertheless, each company is encouraged to articulate its own specific approach, based on their corporate values. Open dialogue and a joint effort of companies and public bodies can help promote effective action and ensure stakeholders are fully aware of the real risks that they face.

The threats and opportunities posed by Big Data represent issues that go beyond transparency or privacy. Responsible businesses are encouraged to take the initiative and tackle the main issues they face when using Big Data, to maintain a consistent alignment between values and behaviour and to mitigate the risks.

---

**The IBE would like to thank all those who contributed to this Briefing. A special thanks goes to Natasha Le Sellier from L’Oreal for the important input to the discussion and to Claudia Natanson – Security Practitioners, Gareth Tipton – BT, and Roeland Beerten and Olivia Varley-Winter – Royal Statistical Society for reviewing the briefing and providing useful comments and suggestions. We are grateful to all who contributed, but the IBE remains solely responsible for its content.**

---

**This and other Business Ethics Briefings are available to download free of charge from the IBE website:**


If there is a topic you would like to see covered, please get in touch with us on +44 (0) 20 7798 6040 or email: research@ibe.org.uk