Extended Abstract

Big Data – It Is Not the Silver Bullet That Some People Think It Is?

Paul Taylo

Trydown Limited, UK

Introduction:

There is certainly a large amount of noise at the moment regarding Big Data especially around what it can do, its challenges and how it could change the world for a better place.

However like most new concepts and ideas, one has to maintain a certain amount of suspicion around any new technology idea. This is because (a) new ideas often have a large amount of hype and therefore under-deliver (b) people cannot see anything wrong with a new idea and tend to overlook its shortfalls and (c) people often jump on the bang wagon and 're-badge' other ideas as the new idea typically for commercial reasons.

Therefore in this article, we investigate what big data is, what it can be used for, and the challenges with its implementation.

What is Big Data?

If one was to search the internet then you are likely to find hundreds, if not thousands, of different definitions of Big Data. However, the following three trends seem to underpin most definitions

• There is a massive volume of data. While size and volume is often relative to your circumstances, we are talking in the range of millions of data items with often hundreds of data variables within each data item.

• The data is constantly changing often at a rapid pace. New items are being added, updated and removed quickly.

• Finally the data is stored in a variety of different formats. This will cover the more 'traditional' pre-defined structured database formats but also a wide range of unstructured formats such as videos, audio recordings, free format text, images, social media comments, etc.

Once this data is collected then it is possible to undertake various forms of analysis. This analysis will find patterns, trends, themes and correlation between variables. This analysis can then be used to explain historical behaviours as well as to predict and shape future behaviours. A few simple examples are listed below is illustrate this point:

· Governments obtain insights to help them with healthcare analysis

• Meteorologists can use Big Data to predict and understand weather conditions

• Political parties can Big Data to understand voting intentions

• Medics can try and understand the cause and spread of diseases. This will allow preventative measures to be implemented. (This is very topical at the time of writing with the COVID-19 pandemic).

• Finally Big Data can help with the 'normal' functions of a business. For example cost/profit management, marketing/product management, improving the clients' experience and internal process efficiencies

In fact, Big Data can be used in most areas of life.

However, what are the challenges and issues with using Big Data?

Big Data is not a silver bullet and there are challenges with implementing it successfully. A poorly implementation Big Data project will cause more problems than it solves.

Before an organisation attempts to implement or use Big Data then (like any change) it needs to have a clear business reason which is linked to the organisation's strategy. This will ensure senior management buy-in and a clear focus on what needs to be implemented.

It would also be advisable to perform some sort of cost/benefits analysis to understand whether the benefits outweigh the costs, stress and challenges of implementation. Six of the main implementation challenges are detailed below:

• The first challenge of Big Data is that the term is often misunderstood and misused. While it is often very easy to be sceptical but firms selling business will often use Big Data to cover a wide range of data analysis techniques because they feel using the 'more trendy' term will generate more business for them.

Therefore the first rule of thumb for Big Data is to ensure that you are actually using Big Data.

• The second item relates to the sheer challenge of processing a vast amount of constantly changing data across many differing and incompatible formats. A complex (and no doubt expensive) stack of technology will be required to continuingly retrieve the data, interpret it, store it and then analyse it.

Therefore before an organisation embarks on to implement a Big Data project then it is important the firm fully understands the costs, overheads and complexity of this technology. This should be covered in the cost/benefits analysis mentioned earlier.

• Thirdly data security is a massive challenge. Data security and its set of legal rules is a complex issue at the best of times. However, its complexity will increase dramatically when used with Big Data especially if data is gathered and processed across international boundaries.

Therefore organisations need to be aware of the rules and ensure they have policies and processes in place to comply with them.

• Like any new discipline or speciality, there is a large shortage of genuinely skilled and experienced individuals in Big Data. There are many people who will pass themselves off as data scientists, data miners or Big Data specialists but care needs to be taken when employing people to ensure they have the skills and experiences required.

Therefore it is important the firms clearly define what skills, capabilities and experiences are required when trying to recruit Big Data 'experts'.

• (Like all data analysis or research techniques) there is the risk of inaccurate data. This could be due to (a) the data sources being separate and not linked together properly (such as purchasing habits not being linked to geographical locations) (b) the data being of poor quality (c) the data is gathered over a poor sample size which means the results could be bias and/or (d) the data being gathered is misunderstood by the data analysis team.

These problems are exaggerated by the size of the data, its constantly changing nature and the differing formats.

Therefore, like any data analysis or research project, it is important the organisation is fully aware of any data inaccuracies so assumptions, warnings or even disclaimers can be noted against any analysis produced.

Journal of Materials Science and Nanomaterials

Extended Abstract

• Finally there could also be issues when processing or analysing the data. There could be errors in the algorithms employed, the wrong variables could be measured or people simply misinterpret the outcomes provided. Again this will be exaggerated by the size of the data, its constantly changing nature and the differing formats.

Therefore when performing Big Data analysis then organisations need to fully analyse the data across multiple algorithms so the data is assessed through multiple lenses to try and obtain a much-rounded view. Also, any material issues with the analysis should also be clearly stated.

The Dark Side of Big Data

Finally, there is a dark side of big data. As mentioned earlier, Big

Data techniques allows one to predict and change people's behaviours. While this is not necessarily a bad thing (because it could help with disease prevention) but this technique could be used to change people's behaviours for somebody else's own personal needs. For example, there have been various documented examples where Big Data techniques have been used to change people's voting intentions.

What is the future of big data?

Big data definitely has a massive future going forward and will no doubt provide a great benefit to society. However, it is important that one does not underestimate both the implementation challenges posed and the dark side of Big Data.