



Big Data in Logistics

To make intelligent logistics solutions.

Table of Content

Contents

Table of Content.....	1
Abstract	2
1. Introduction - What is Big Data?	3
2. Application of big data to develop intelligent logistics solutions	5
2.1. Transportation solutions	6
2.2. Forecasting and logistics planning	6
2.3. Warehouse Management.....	7
3. Impact of Big-data systems for logistics.....	9
4. Risk associated with using big-data in logistics solutions.	12
4.1. Macro Risks	12
4.2. Cyber-Security Risks.....	12
4.3. Risk of bad data and wrong interpretations	13
5. Conclusion	14
6. References	15

Abstract

This report discusses the utilization of big data in order to create effective logistics solutions. Certain characteristics such as high volume and variety, help logistics managers in different parts of supply chains make well-informed decisions. These areas include and are not restricted to transportation, logistics planning, forecasting and warehouse management.

Smooth functionality of IT (Information technology) systems such as cloud computing and point of sale (POS) registries are essential for the success of a big data solution in logistics.

Even though a big data system may yield many benefits for a supply chain, it has its own inherent risks. Out-dated data may lead logistics managers to make wrong decisions and cyber security threats may lead companies to give away sensitive information. These risks and potential benefits are comprehensively discussed in the lecture.

1. Introduction - What is Big Data?

Big data is a concept that is being evolved since late 1990s, where Cox and Ellsworth (1997) defined it as large data sets that are hard to be stored with-in internal memories. Back then, big-data was perceived as a problem. Today big data analytics have evolved into a value adding technique for organisations, which provide them a competitive edge over other organisations.

With the increased use of data transfer over the internet, organisations are left with a large amount of data at their disposal which provide insights into market trends, purchasing patterns, customer behaviour and maintenance cycles. To utilize this opportunity, new academic disciplines such as big data business analytics (BDBA) and supply chain analytics (SCA) have emerged (Wang et al., 2016).

Improving on the initial concept by (Manyika et al., 2011) which stated that there are three dimensions that characterize big data, IBM Big Data & Analytics Hub (2016) explains that there are five dimensions to it; also known as the 5 V's. They are as follows.

- i. Volume – Substantially large amount of data being processed than in the case of a usual database.
- ii. Velocity – The data is processed real time or nearly real time in order to keep up with the large quantity of data.
- iii. Variety – The data is generated from different sources, most of being unstructured- such as tweets, photos or customer purchase history.
- iv. Veracity – The certainty of data is high. Depending on the origin, methods used to mine and the technology used, data has a high statistical reliability.
- v. Value – The insights provided via big data analytics add value to entities' operations.

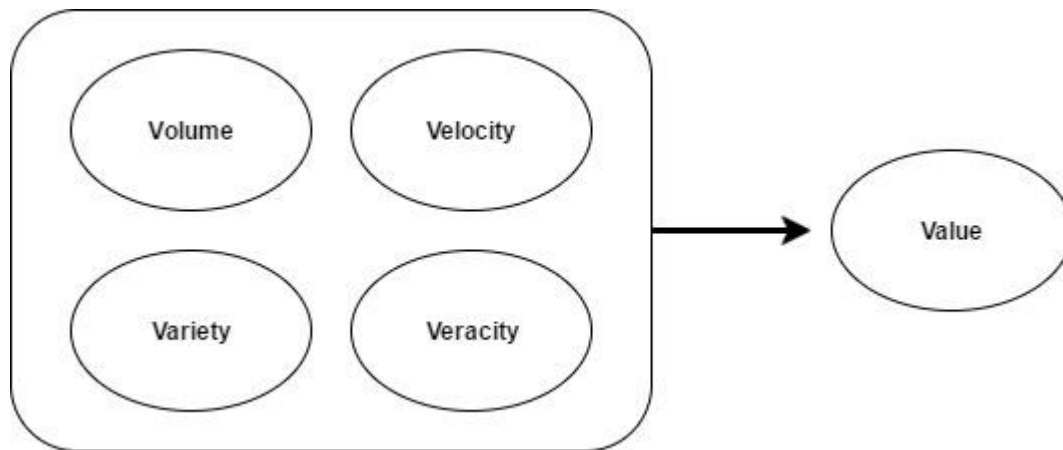


Figure 1 - Big data model (Author made based on IBM Big data & Analytics)

According to Foster et al. (2016), since late 1990s, leading retailers in USA such as Wal-Mart, have been using data driven supply chain optimization mechanisms to improve their competitiveness. By the present day, it has evolved into comprehensive analysing of customer data to make strategic decisions across different stages in supply chains; inventory management, distribution and supplier negotiations.

2. Application of big data to develop intelligent logistics solutions

Waller and Fawcett (2013) explain that there are three users of big data in logistics business; namely, carriers, manufacturers and retailers. They use the analytics for forecasting, inventory management, transportation and human resource management.

User	Forecasting	Inventory management	Transportation management	Human resources
Carrier	Time of delivery, factoring in weather, driver characteristics, time of day and date	Real time capacity availability	Optimal routing, taking into account weather, traffic congestion, and driver characteristics	Reduction in driver turnover, driver assignment, using sentiment data analysis
Manufacturer	Early response to extremely negative or positive customer sentiment	Reduction in shrink, efficient consumer response, quick response, and vendor managed inventory	Improved notification of delivery time, and availability; surveillance data for improved yard management	More effective monitoring of productivity; medical sensors for safety of labor in factories
Retailer	Customer sentiment data and use of mobile devices in stores	Improvement in perpetual inventory system accuracy	Linking local traffic congestion and weather to store traffic	Reduction in labor due to reduction in misplaced inventory

Figure 2 - Possible uses of Big data in logistics (Source - Waller and Fawcett (2013))

In this chapter, case studies are analysed to highlight the use of big data to introduce intelligent solutions in these different logistics components.

2.1. Transportation solutions

Cavanillas, Curry and Wahlster (2016) showcase that the data sources in the transportation sector are increasing at an exponential rate. This growth rate is majorly driven by factors such as the deployment of sensors at logistics hubs, ports and train/bus stations, EOBR (Electrogonic on-board recorders) in trucks and other carriers, smart cards and loyalty schemes such as frequent flyer programs.

These big data sources are effectively used by both the private and public sector to provide intelligent logistics solutions for their customers.

For instance, Dublin city council partnered up with IBM's smarter city program to analyse city's traffic updates and related data. The real-time GPS based analytics data was used to plan the city's bus fleet effectively and the outcome was the reduction of average journey time by 10-15% (Tabbitt, 2014).

In the private sector, according to McAfee and Brynjolfsson (2012), a major U.S airline used big-data analytics powered by PASSUR aerospace; a decision-support technology service provider, to improve the accuracy of estimated time of arrival (ETA) of their planes. The big-data project replaced the historical practice of using the ETAs provided by the pilots with big-data-backed intelligent forecasts. This helped the airline to eliminate the gap between estimate and the actual arrival times, saving millions of USDs at each airport.

2.2. Forecasting and logistics planning

DHL (2013) explains that big-data can be used as a 360-degree solution in a 3PL/4PL operation. Some of the key features of this suggested solution are as follows.

At tactical level, real-time traffic data and location information is used to make dynamic 'travelling salesman problem' related decisions.

At an operational level, detection of ad-hock demand/supply variances derived from external sources to adjust capacity of warehouses, and analysing data on local developments in macro factors such as economic, political and natural to identify, evaluate and to mitigate the impact of the risk of possible supply chain disruptions.

This helps the company to collaborate with their customers effectively and to gain their trust, as reliability and accuracy of real-time information is the key to effective collaboration (McClellan, 2003).

Higher volume and the variety big-data provides, is used for strategic network planning activities to make accurate strategy level decisions.

2.3. Warehouse Management

Wang et al. (2016) points out that intelligence generated via big-data in the forms of aggregate production demand, plant capacities and operational cost at different locations can be utilized to create a profitable supply chain network design. This leads organisations to configure the optimum number, location and the size of warehousing facilities.

Along similar lines, Cohen (2015) explains that it is viable to map all the product interactions such as orders, product examinations and transactions using internet and cloud technology. These big-data sources can be used in material-management systems to control the inventory levels across an organisation's extended supply chain, eliminating capacity planning issues. The way this system works is conceptualized in the figure 3 below.

This research further suggests that machine learning and optimization algorithms can be used in the monitoring stage shown in the diagram, to make intelligent business decisions.

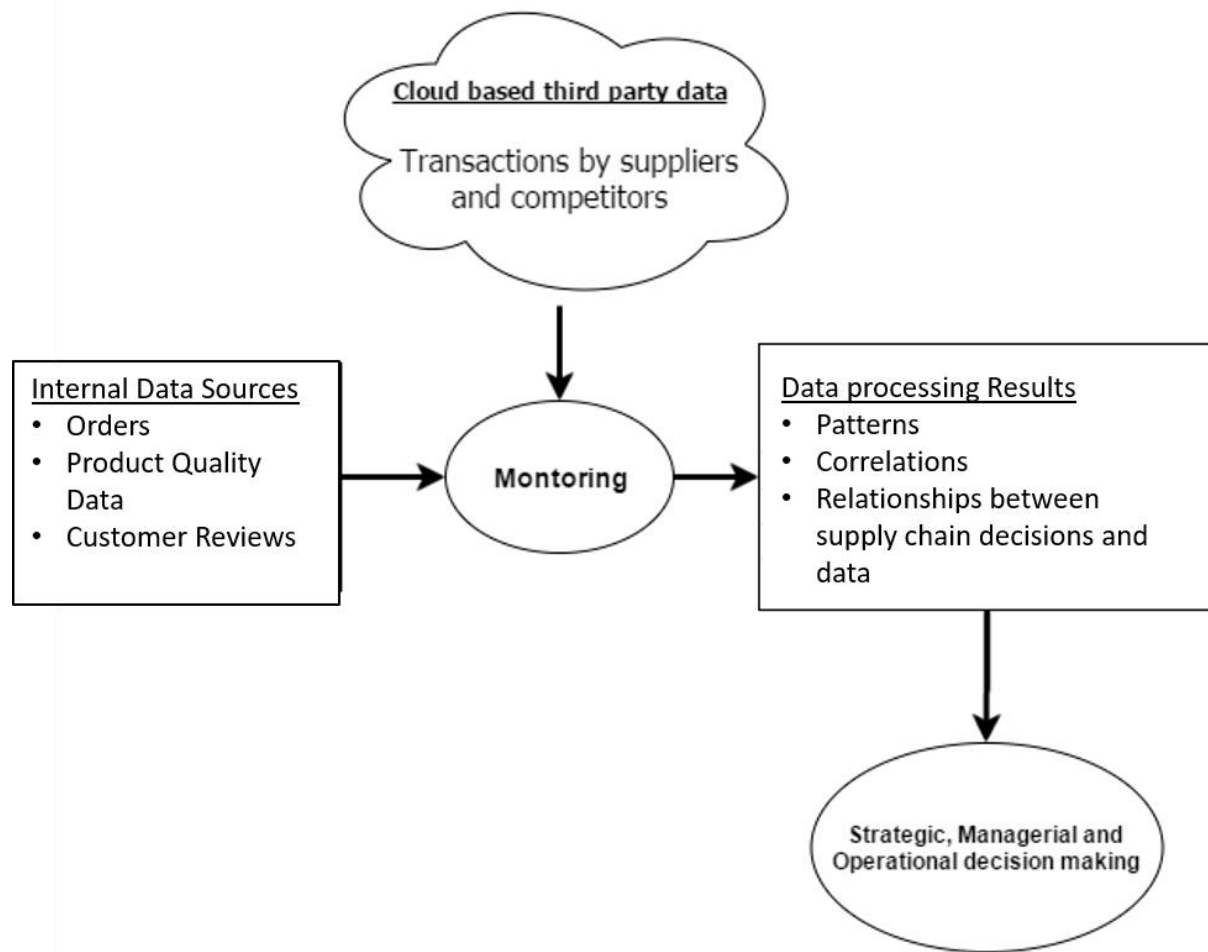


Figure 3 - Big-data analytics usage in decision making (Author made based on Cohen, 2015)

3. Impact of Big-data systems for logistics

Big-data analytics could be used as a global platform that can enhance the global logistics performance. However, the success of the proposed big-data logistics business platform depends on two assumptions. First, the authorities, shippers, retailers and other parties should be willing to collaborate. Secondly, the planning mechanisms should be integrated into a single platform via cloud technology, Neaga et al. (2015).

Cloud infrastructure bridges the communication and decision making gaps between the partners real-time, increasing the usefulness of big-data (Ferkoun, 2014).

Myerson (2007) supplements this argument by pointing out IBM's use of hybrid data bases to arrange large quantities of structured and unstructured data which helps businesses to overcome limitations in searching and sorting issues that may arise in a big-data solution. The figure 4 below demonstrates the function of a hybrid system; local big-data sources, data storage and a cloud solution collaborating with one another. Such a system comprehends multiple supply chain data sources, identifies patterns as shown in figure 3 above and helps managers to make effective logistics solutions.

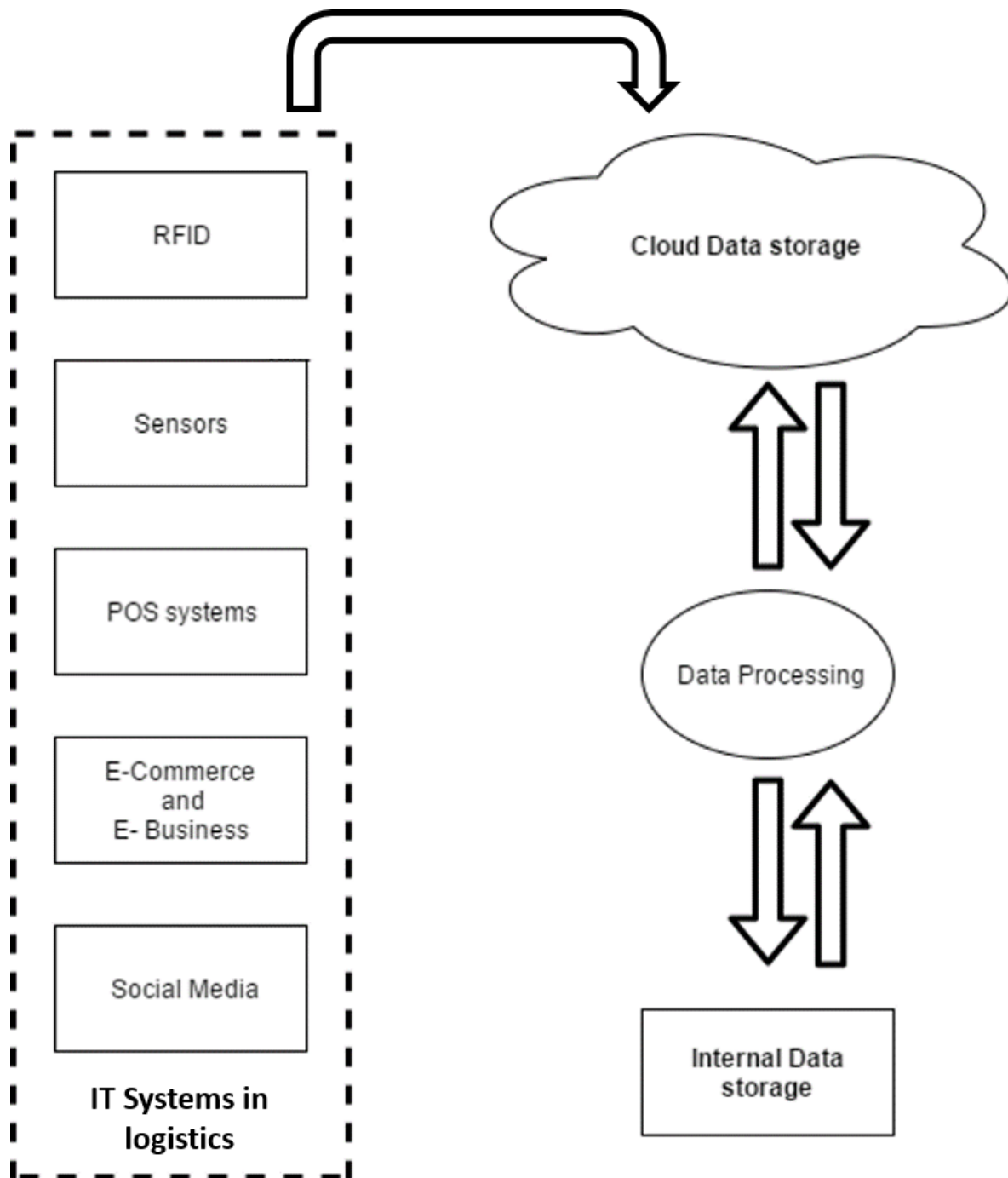


Figure 4 - IT systems in a big-data solution (Author made)

Note – Arrows indicate data flows

Following are a few real-life supply chain case studies that benefited out of a big-data solution.

Medium sized department store; Globus implemented a SAP HANA solution. The real-time analytics helped the company to generate a comprehensive sales report for the full product line in just 17 seconds; an operation which took more than 22 minutes, enabling company to make real-time operational decisions (Forbes, 2013).

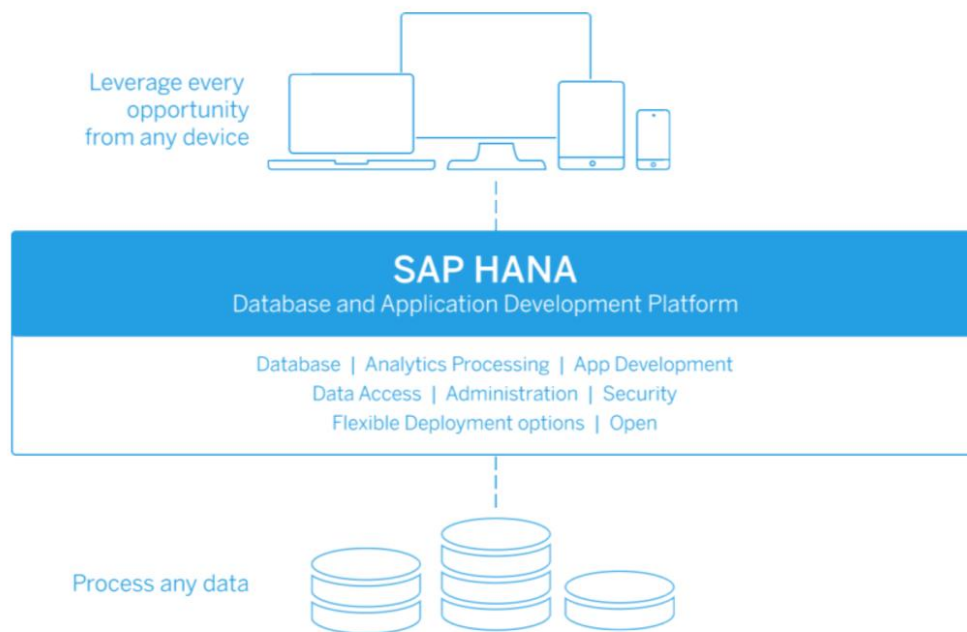


Figure 5 - SAP HANA solution (Source - SAP, 2017)

One of the leading players in returned merchandise management sector – Appriss retail (merged entity of the retail equation and Sysrepublic), claims to have saved retailers over USD 1 Billion by fraudulent return prevention. This is a result of accurately predicting customer behaviour via big-data solution. Their solutions promise to deliver further benefits such as employee fraud reduction, prevention of organized retail crimes and minimizing profit shrink (Appriss retail, 2017).

4. Risk associated with using big-data in logistics solutions.

Irrespective of the usefulness of big-data analytics for intelligent logistics solutions, there are risks associated.

4.1. Macro Risks

The current global political situation, where policy decisions such as Brexit, are taking place, has created challenges for the political and economic integrations among countries (Oliver, 2016). On that grounds, it is hard to agree with the practicality of Neaga et al. (2015)'s first assumption; a global data interchange. This diminishes the possibility of an introduction of cloud based big-data platform as a global intelligent logistics solution.

Furthermore, existing investments on global big-data platforms will become obsolete because of such policy decisions. As Carter (2017) explains, the end result; unused data lakes, will incur prohibitively high costs for the particular organisations.

4.2. Cyber-Security Risks

Saraladevi et al. (2015) identify security issues as one of the primary concerns of a big-data project. They explain that the threats come in various forms including engineered and unintentional data leakages, lack of encryption and public/private databases being vulnerable. Furthermore, data hackers may engineer variety of attacks including brute force (Bernstein, n.d.) and denial of service.

These cyber-security threats may lead organisations to leak sensitive data. Retail giant, Target's customer database breach in 2013 which ended up leaking sensitive financial details of over 40 million customers, is one of the largest ever big-data attacks (The New York Times, 2015).

To make matters worse, competitors might use it as leverage against the organisations facing cyber-security breaches.

4.3. Risk of bad data and wrong interpretations

The usage of big-data analytics can be simplified as shown in the figure 5 below. If the data is bad or the data is wrongly interpreted, the decisions made based on them are not optimal.

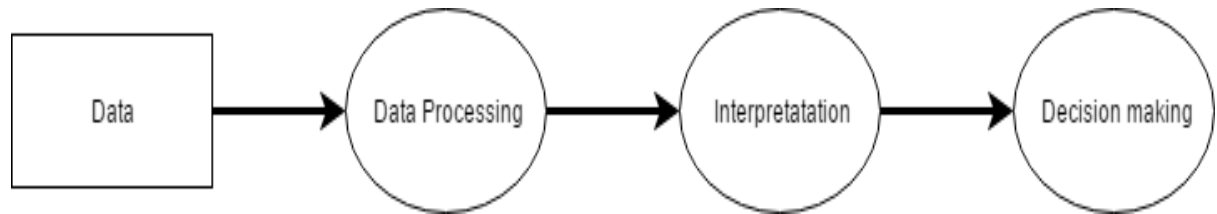


Figure 6 - Big-data analytics process in decision making (Author made based on Zins, 2007)

Pries and Dunnigan (2015) explain that quality of the data can be tarnished due to many reasons including ambiguity of data, data aging, missing variables and inconsistent use of units and terminology.

The data interpretation is done using advanced algorithms, and they are not error proof. Google's flu trend – one of the famous big-data projects was proven to be erroneous, when it was identified that the algorithms used were unable to identify anomalies (Lohr, 2014).

If a similar data misinterpretation or unhealthy data usage is to take place in a commercial application which utilizes big-data as a logistics solution, supply chain disruptions with adverse financial and operational implications are inevitable.

5. Conclusion

Big-data is an effective way to provide intelligent logistics solutions due to three key reasons. Primarily, the richness in volume and variety of data increases the accuracy of analytics. Secondly, real-time data processing enhances the relevancy of analytics. Finally, big-data solutions backed by a cloud network connect the extended supply chain of an organisation eliminating disruptions. Collectively, they help a supply chain operation to run smoothly.

However, there are inherited risks in big-data systems such as macro risks, misinterpretation of data and cyber security threats. By using risk mitigation mechanisms such as firewalls, enhanced encryptions, implementation of hybrid security measures; both digital and physical, the intended outcome of a big-data project can be harvested.

6. References

- Apprissretail. (2017). *Overview – Appriss Retail*. [online] Available at: <https://apprissretail.com/benefits/prevent-fraud/> [Accessed 14 Mar. 2017].
- Bernstein, D. (n.d.). *Understanding brute force*. 1st ed. [ebook] Chicago: The University of Illinois at Chicago. Available at: <https://cr.yp.to/snuffle/bruteforce-20050425.pdf> [Accessed 28 Feb. 2017].
- Carter, J. (2017). *Is big-data a big failure?* [online] TechRadar. Available at: <http://www.techradar.com/news/world-of-tech/is-big-data-a-big-failure-1313340> [Accessed 28 Feb. 2017].
- Cavanillas, J., Curry, E. and Wahlster, W. (2016). *New Horizons for a Data-Driven Economy*. 1st ed. Cham: Springer International Publishing, pp.225-244.
- Cohen, M. (2015). *Inventory Management in the Age of Big Data*. [online] Harvard Business Review. Available at: <https://hbr.org/2015/06/inventory-management-in-the-age-of-big-data> [Accessed 27 Feb. 2017].
- Cox, M. and Ellsworth, D. (1997). Application-Controlled Demand Paging for Out-of-Core Visualisation. *Proceedings of the 8th conference on Visualization '97*, [online] p.235-ff. Available at: <http://dl.acm.org/citation.cfm?id=266989.267068> [Accessed 21 Feb. 2017].
- DHL, (2013). *Big Data in Logistics*. Germany: DHL Customer Solutions & Innovation.
- Ferkoun, M. (2014). Cloud computing and big data: An ideal combination. [Blog] *IBM Cloud computing news*. Available at: <https://www.ibm.com/blogs/cloud-computing/2014/02/cloud-computing-and-big-data-an-ideal-combination/> [Accessed 27 Feb. 2017].
- Forbes. (2013). *SAPVoice: How Retailers Use Big Data to Spot Hot Shopping Trends*. [online] Available at: <https://www.forbes.com/sites/sap/2013/11/25/how-retailers-use-big-data-to-spot-hot-shopping-trends> [Accessed 14 Mar. 2017].

Foster, I., Ghani, R., Jarmin, R., Kreuter, F. and Lane, J. (2016). *Big data and social science*. 1st ed.

IBM Big Data & Analytics Hub. (2016). *Extracting business value from the 4 V's of big data*. [online] Available at:

<http://www.ibmbigdatahub.com/infographic/extracting-business-value-4-vs-big-data> [Accessed 23 Feb. 2017].

Lohr, S. (2014). *Google Flu Trends: The Limits of Big Data*. [online] Bits Blog.

Available at: <https://bits.blogs.nytimes.com/2014/03/28/google-flu-trends-the-limits-of-big-data/> [Accessed 1 Mar. 2017].

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C. and Byers, A. (2011). *Big Data: The next frontier for innovation, competition and productivity*. 1st ed. [ebook] McKinsey Global Institute. Available at:

<http://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/big-data-the-next-frontier-for-innovation> [Accessed 23 Feb. 2017].

McAfee, A. and Brynjolfsson, E. (2012). Big Data: The Management

Revolution. *Harvard Business Review*, [online] pp.59-68. Available at:

<https://hbr.org/2012/10/big-data-the-management-revolution> [Accessed 26 Feb. 2017].

McClellan, M. (2003). *Collaborative manufacturing*. 1st ed. Boca Raton, FL: St. Lucie Press, pp.109-132.

Myerson, J. (2007). *RFID in the supply chain*. 1st ed. Boca Raton (Florida):

Auerbach, p.117.

Neaga, I., Liu, S., Xu, L., Chen, H. and Hao, Y. (2015). Cloud Enabled Big Data Business Platform for Logistics Services: A Research and Development Agenda. *Decision Support Systems V – Big Data Analytics for Decision Making*.

Oliver, T. (2016). European and international views of Brexit. *Journal of European Public Policy*, 23(9), pp.1321-1328.

Pries, K. and Dunnigan, R. (2015). *Big data analytics*. 1st ed. Florida: Taylor & Francis Group, pp.379-395.

SAP. (2017). *What is SAP HANA | In Memory Computing and Real Time Analytics*. [online] Available at: <https://www.sap.com/product/technology-platform/hana.html> [Accessed 14 Mar. 2017].

Saraladevi, B., Pazhaniraja, N., Paul, P., Basha, M. and Dhavachelvan, P. (2015). Big Data and Hadoop-a Study in Security Perspective. *Procedia Computer Science*, 50, pp.596-601.

Tabbitt, S. (2014). *Big data analytics keeps Dublin moving*. [online] The Telegraph. Available at: <http://www.telegraph.co.uk/sponsored/sport/rugby-trytracker/10630406/ibm-big-data-analytics-dublin.html> [Accessed 26 Feb. 2017].

The New York Times, (2015). \$10 million settlement in target data breach gets preliminary approval. [online] Available at: <https://www.nytimes.com/2015/03/20/business/target-settlement-on-data-breach.html> [Accessed 28 Feb. 2017].

Waller, M. and Fawcett, S. (2013). Data Science, Predictive Analytics, and Big Data: A Revolution That Will Transform Supply Chain Design and Management. *Journal of Business Logistics*, 34(2), pp.77-84. i

Wang, G., Gunasekaran, A., Ngai, E. and Papadopoulos, T. (2016). Big data analytics in logistics and supply chain management: Certain investigations for research and applications. *International Journal of Production Economics*, 176, pp.98-110.

Zins, C. (2007). Conceptual approaches for defining data, information, and knowledge. *Journal of the American Society for Information Science and Technology*, 58(4), pp.479-493.