Portland State University

PDXScholar

Engineering and Technology Management Faculty Publications and Presentations

Engineering and Technology Management

8-2-2015

Assessing the Effectiveness of Big Data Initiatives

Nayem Rahman
Portland State University

Fahad Aldhaban
Portland State University

Let us know how access to this document benefits you.

Follow this and additional works at: http://pdxscholar.library.pdx.edu/etm_fac

Part of the Operations Research, Systems Engineering and Industrial Engineering Commons

Citation Details

Assessing the Effectiveness of Big Data Initiatives

Nayem Rahman, Fahad Aldhaban

Portland State University, Department of Engineering and Technology Management, Portland, OR - USA

Abstract--There is a great enthusiasm with the prospect of big data among business and industry leaders, academia and researchers. A lot of big data tools and technologies have emerged recently to capture, store, process and analyze big data. One remarkable achievement is that a handful of open source technologies have been introduced by Apache Hadoop Foundation that allows any organization to undertake big data projects. Many big data projects have been implemented during the last few years. This paper explores the benefits achieved with those projects. Big data has been applied in diverse fields including scientific and medical discovery projects, studying social science phenomena, conducting meaningful observations of real-world phenomena and running analytics for healthcare and business. Many business organizations want to find new business opportunities such as fraud detection, customer sensitivity analysis, and new product offerings while some other business organizations are still pondering the long term value of big data investments. Business leaders and managers want to be sure that big data projects can deliver true value and provide long term benefits. This paper provides an account of how recent big data project initiatives have been successful in delivering business value and highlights what technology solutions are primarily used by those big data projects.

I. INTRODUCTION

Big Data is a new phenomenon that has recently emerged in data management discipline. While several organizations are enthusiastic to find new business opportunities with big data to drive the growth of their business many are not using big data because they do not fully understand what big data is or what are its benefits [37]. Wlodarczyk and Hacker [50] report that "although there is clear interest in big data, the adoption rate of analytical techniques fuelled by big data that can extract knowledge and value from these data is less understood". Kiron et al. [22] state that "the analytics revolution is still at an early point in its development. Many companies are still struggling to figure out how, where and when to use analytics." Research reports suggest that there are some barriers with implementing big data. One of the most prominent barriers is inadequate skills set [44] and there are issues of governance, lack of understanding as to how to leverage analytics for business value, lack of executive sponsorship [28] and compelling business cases.

Big data has become important with the emergence of data growth as a result of the Internet-generated data volume, machine generated (sensor) data and the evolution of social networking tools. Much of this new kind of data are unstructured when compared to relational database data which are structured. Big data has certain characteristics that make it distinct from 'normal data' stored in conventional relational database systems. It is not possible to capture, store, process and analyze them using the conventional database and business intelligence tools [20] which an organization's IT department has in its arsenal. This requires a new set of tools to process big data as 90% of such data are unstructured [41].

Rajpurohit [42] observed that business managers continue to struggle to capture the business value from big data initiatives. But he suggested that business managers need to understand big data processes. They also need to understand that big data analytics is not an IT solution to data problem. It is a new paradigm to manage data and transform it into valuable insights [9]. Business managers have responsibility to understand the big data processes and take ownership [1], [42].

Big data has not just suddenly appeared. Organizations started generating huge data volume since mid-90s with the advent of the Internet technology, the start of widespread use of personal computers [2] and computing technology-based automation [41]. Initially there was not much computer hardware and software capability to process large volume of complex (unstructured) data. Now the computing industry is more mature in terms of hardware and software capability and efficiency. Data storage and computer hardware have now become much cheaper. Large volumes of data can be processed with processing capability on commodity servers. Though there are software and hardware capability available out there to process, store, and analyze big data will requires a new set of tools and technologies and new skill sets to manage, analyze and generate analytical results.

This paper is organized as follows: Section 2 briefly discusses the use of big data and related work in this area. Section 3 provides an overview of big data characteristics. Section 4 provides success story big data projects. Section 5 talks about taking certain measures in initiating big data projects. Section 6 concludes the paper.

II. LITERATURE REVIEW

The rapid growth of data volumes is the byproduct of the use of the Internet, social media and smart mobile devices [36]. This huge volume of data which is known as Big Data requires a new set of tools and technologies to receive, store, process and analyze. Companies need to run their business based on IT infrastructure for strategic agility [48]. "Companies must employ information technology with a sophisticated understanding of the requirements for competitive advantage" [40]. To achieve the benefit of the big data companies need to drive action plans and get value out of data by applying algorithms and prescriptive models and technologies to solve specific business problems [3].

Research on big data suggests that business and other organizations can drive their growth by using new and additional data generated from big data. Nambiar et al. [34] suggest that big data can revolutionize the healthcare industry by improving operational efficiencies, helping to predict and plan responses to diseases, optimize healthcare spending and can improve overall quality in healthcare system.

Governments of different countries including the U.S. and European Union have taken proactive initiatives in big data research. The US government has recently announced new funding initiatives for different agencies to conduct research in big data applications [26], [45]. Governments are interested in using big data for predictive analysis applied to decision making and to improve national security [6]. Research [50] suggests that "almost half of the scientific publications using predictive analytics were in healthcare, smart services, the internet of things, and weather and environment."

A survey [39] of 211 senior marketers about big data effectiveness suggest that big data boosts marketing effectiveness. A leading practitioner of big data-based marketing said that her company had "achieved 8-23% sales, lift via its proprietary data analytics system compared to initiatives that did not use this resource" [37]. Bertino [6] observed that organizations are shifting reporting and decision support systems to predictive analytics as their next move in business intelligence. Rijmenam [35] suggested that the future of big data is to perform predictive analytics. With this he meant that "the objective of perspective analytics is not only to predict future outcome, but also to make recommendations based on those outcomes" [35]. Fan and

Bifet [14] report that existing data mining tools are not capable enough to handle big data due to the volume and complexity. For big data a new set of machine learning algorithms and tools are emerging. Leading commercial data warehouse vendor, Teradata® [27] has announced development of tools such as R programming languages and other analytical tools to enable business users and analytical community skilled in SQL queries to access Hadoop-based data.

III. CHARACTERISTICS OF BIG DATA

Big data has several distinct properties (Figure 1) which are identified by five V's - volume, velocity, variety, veracity and value [51]. Big data is big and complex in terms of volume, with characteristics that are beyond the capacity of conventional database systems and associated tools to capture, store, process and analyze. With the emergence of computer and the internet generated data the **volume** of data generation is much bigger [32]. Given round the clock generation of data in this digital age, it has become important to have the capability to receive, process and store data in real-time (**velocity**). In banking sectors most of the transactions occur electronically. In 2012, the numbers of credit and debit card transactions were 26 and 47 billion respectively [15]. This means that thousands to millions of transactions were created per minute. It has become increasingly important to integrate, process and analyze data flows from both internal and external sources in near real-time [13].
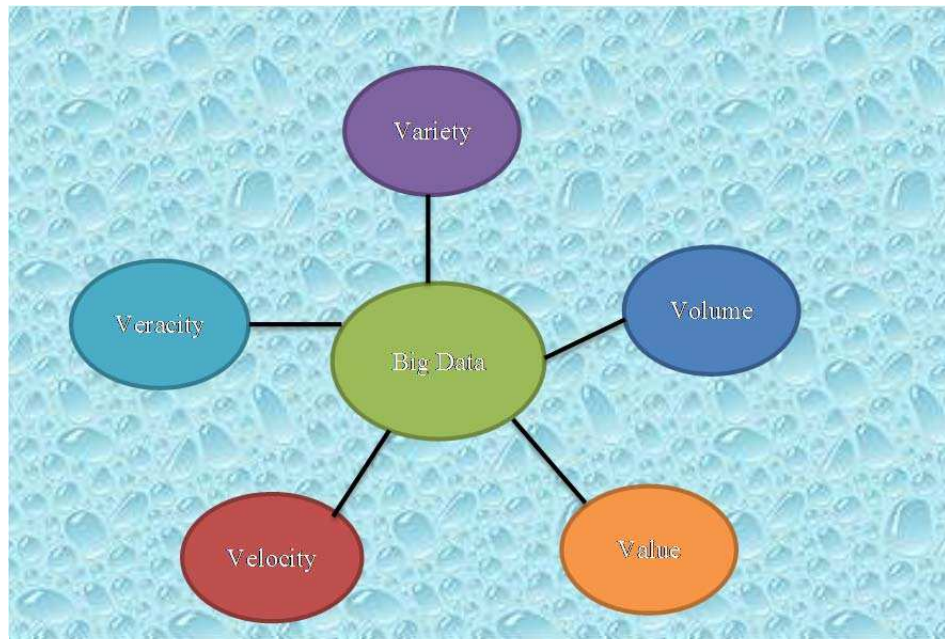

Figure 1: Characteristics of Big Data – The five V's.

Big data refers to data that comes from many sources and in different formats (**variety**). The data is unstructured (more than 90%), semi structured and structured. For example, in banking, customer records are structured data; email header information (sender, receiver, date, etc.) are semi structured data while email text messages and other bank documents are unstructured data. As big data is mostly unstructured, data consistency issues come into play (**veracity**). Big data technologies are used to extract value from unstructured data. But, extracting real value out of unstructured data poses a question of accuracy and reliability [13]. To make data reliable and meaningful, big data technologies need to make sure traceability is ensured. Any question or doubt about the validity of data must be addressed via some traceability mechanism with data processing that follows some standard procedures. With regards to the characteristic, **value**, business organizations must find new business opportunities to utilize big data. One example for bank to use machine learning algorithms for predictive analytics to detect and prevent fraud. Big data systems provide capabilities for analyzing a greater breadth and depth of data [43].

IV. BIG DATA USE-CASES, DEVELOPMENT AND IMPLEMENTATION

Industry and academic papers suggest that big data is used in different areas of research, businesses and government offices. For example, big data helps in agricultural sector with precision based agriculture yields. In agriculture, farmers collect massive amount of data through sensors "to predict the best time to plant, where to plant to improve yields, reduce operational costs and minimize environmental impacts" [53]. Here, big data technologies allow putting all data together and then running the proper analytics and making use of massively parallel computing architecture.

In an effort to augment the growth of big data analytics development Intel® has joined Hadoop community in offering enterprises the support of a big technology player in terms of its Hadoop distribution software [24] along with a new generation of processors to provide performance in big data processing and analyzing. Intel's proof of concept customer NextBio® confirmed software and hardware performance improvements on its computing resources using the Intel Hadoop distribution [24].

Case No.	Use Cases	User Benefits	References
UC-01	Electronics Records Archive	Online Public Access Systems	[51]
UC-02	Perimeter Intrusion Detection	Protect and monitor critical infrastructure and secure borders	[51]
UC-03	Traffic Pattern Analysis	Decrease traffic congestion and reduce traffic accident injury rates	[16], [51]
UC-04	Wind enegry and wind turbine placement and Maintenance	Identify optimal location for wind turbines to maximize power generation and reduce energy cost	[51]
UC-05	Medical Monitoring	Detecting infections in premature infants 24 hours before they exhibit symptoms	[51]
UC-06	Demand Response Targeting	How to target the proper customers for DR	[25]
UC-07	Battling the Flu	Helps CDC to combat influenza	[34]
UC-08	Monitorig Diabetic Patients' Insulin Intake	Allows healthcare professionals identify problems before they become severe.	[34]
UC-09	Realtime Integration and Trend Detection	Effective processing of data in real-time	[11]
UC-10	Real-Time Energy Management	Improvements in energy efficiency and cost	[4]
UC-11	Customer shopping Behavior	Predicting how long a customer will travel to buy a product or which shop she will choose	[38]
UC-12	Risk Adjustment of Patient Expenditures	Improving the accuracy of patient expenditures	[30]
UC-13	Taxi Drivers' Airport Pick-up Decisions	Maintain taxi demand and supply equilibrium at the airports	[52]
UC-14	Prediction of Energy Consumption	Improving energy consumptin prediction for efficiency and cost reduction	[46]
UC-15	Alarm Prediction in Railroad Networks	Monitor equipment conditions	[29]
UC-16	Load Shedding Problems	Fast solution of load shedding problems	[7]
UC-17	Risk-of-Readmission of Heart Patients	Big data driven solutions to predict the 30-day risk of readmission for congestive heart failrue patients	[54]
UC-18	Consumer Profiles from Social Media Data	Obtain valuable insights from consumer feedbacks	[19]
UC-19	Information Analytics for IT Incident Prediction	10-30% reduction in new incidents. Cost avoidance of $4 million over two years	[8]

Figure 2: Successful Big Data Use-Cases

Figure 2 provides a list of high profile big data projects that were implemented in different fields of studies. The use-cases were identified based on a study of recent articles that appeared at IEEE Big Data conference proceedings and a host of industry papers. All these case studies show significant benefits to customers. UC-01 (Figure 2) was conducted by the National Archive and Records Administration (NARA). It took the challenge of digitizing the huge of unstructured data from across the federal agency ecosystem, Congress and several presidential libraries [51]. The outcome of the use case showed faster ingestion of documents, and improved user experience and significant reduction of storage costs [51]. UC-02 was conducted by TerraEchos for premier intrusion detection [51]. By using big data tools and technologies such as streams analytics software and predictive analytics the project was able to provide a solution that helps organizations to protect and monitor critical infrastructure and secure borders.

UC-03 reports two use cases, one was conducted by the Royal Institute of Technology of Sweden to perform traffic pattern analysis [51] and the other was conducted by Garzo et al. [16] for location prediction over mobility traces. The traffic pattern analysis using big data technologies helps in handling a large traffic data streams to perform near real-time analytics on current traffic conditions. This has helped in decreasing traffic congestion and accidents in heavy traffic congestion cities. UC-04 big data case study was conducted by Vestas Wind Energy to build efficiency wind turbine placement and maintenance. The big data modeling solution has helped the Danish company to improve accuracy of wind turbine placement, reduce response time for wind forecasting by 97%, maximize power generation and reduce energy costs [51].

UC-05 was about medical monitoring, undertaken by the Institute of Technology at the University of Ontario. The implementation of this big data project has enabled optimal patient monitoring. Patients connected to the equipment are continuously monitored for vital signs including blood pressure and heart rates. This application also allows for detecting infections in premature infants about 24 hours before they show symptoms.

UC-06 is designed to determine how to target real customers for an air-conditioning disaster recovery (DR) program with help of big data analytics [25]. In this project a targeting algorithm is used based on high resolution data collection method using big data. Here big data technology makes high resolution information available (i.e., data accuracy) which in turn helps in identifying each consumers demand consumption and targeting customers according for DR program.

UC-07 and UC-08 provide capabilities and usefulness of big data in healthcare system. These use cases show how big data projects can help in preventing diseases and improving patient treatments conditions. UC-09 reports a success story of big data technology's high speed data stream capability in real-time. This project used an open-source data streaming software called Storm [11]. UC-10 is another example of utilizing predictive analytics capability with big data. This project used smart gird data with advanced forecasting and a big data analytics platform to perform analytics on large scale data to improve energy efficiency and reduce costs [4].

UC-11 uses a big data mining framework to accurately predict how far a customer will travel to buy a product and which shop will be chosen [38]. UC-12 uses an advanced machine learning algorithm by leveraging Hadoop distributed computing platform such as MapReduce [30]. This big data analytics framework was implemented for developing risk-adjustment model of patient expenditures. Big data technology could be applied in building a decision model [52] for airport taxi drivers' trip to and from the airport (UC-13). UC-14 applied an incremental clustering technique with big data to improve the prediction of time series data with polynomial time [46]. UC-15 is a case study in predicting alarm in large-scale sensor networks in railroad. Railroad censors generate massive amounts of data about equipment conditions. Using big data predictive analytics this project was able to predict alarms associated with catastrophic equipment failures several days ahead [29].

UC-16 provides a fast solution of load shedding problems using big data technology [7]. UC-17 provides a holistic predictive modeling solution for risk prediction in healthcare [54]. This large volume of healthcare predictions should be integrated with other heterogeneous sources used to improve quality of prediction. Big data technologies are used to store all data in one place (e.g., distributed file system), process them using paralleling processing system (e.g., MapReduce) and apply predictive analytics. This solution resulted in predicting the 30-day risk of admission for congestive heart failure. Big data helps in building consumer profiles based on social media activities.

UC-18 provides a story of successful construction of different consumer-segments using social media data [19]. This allows companies to offer the right products to the right customer segments which helps in increasing sales revenue. UC-19 is Intel's case study on IT log information analytics for incident prediction. This case study is able to reduce IT incidents by tracking data from previous incidents and predicting future behaviors. This use case reported that it was able to provide a 10 to 30% reduction in IT incidents which allowed a cost avoidance of $4 million over a two year period [8].

In big data, predictive models are heavily used and more new tools are emerging to open them up to new users to apply to a varieties of use cases. The uses cases listed in Figure 2 shows that many of them dominated by big data predictive analytics. In IBM, researchers are applying predictive analytics to design diagnostics and evaluate treatments of patients at risk of heart failure as much as two years ahead of time [17].

V. UNDERTAKING BIG DATA PROJECTS

Big data is a new paradigm [9] in terms of data characteristics as well as tools and technologies to be used. It requires financial involvement for acquiring new tools and technologies (many of them opens source though), training current work force to make them attain new skill sets and hiring new people (programmers and data scientists) who have strong knowledge in big data technologies. On the big data project implementation side, business leaders and project managers need to come up with some strategy that align big data with company business goals. They need to follow some criteria. They need to define a clear and measurable goals. Instead of worrying about collecting all tools and technologies they need to think about how far big data will drive business goals [5]. Will their data help in acquiring new customers or retaining current high profile customers? Business organization need to determine if their own data will be enough or they need external data. Answering these questions will help business managers in managing right data and acquiring right tools and technologies.

Business leaders and IT managers might think of running analytics test on small sets of data. This will help them to understand if big data will bring value to the organization and can be used to convince CEO or IT executive who do not see the value in spending time and resources on data and analytics. Bailey [5] suggests that one way to demonstrate "a clear connection between data and profitability and remove the intimidation factor is to test a small set". This will help to avoid getting bogged down in the tools and platforms at the very early stage of big data initiatives. Business and IT managers, project managers need to ensure that company CEO and CFO do not perceive that big data belongs to company's cost center rather than an investment for the entire company. Loshin [31] asserts that there needs to be a "clear processes for determining the value proposition". If not, the investment in the "technology were not grounded in expectations for clear business improvements."

Wheatley [49] reports seven best practices for big data retail analytics projects: (1) link the big data project to the business' strategies; (2) have clear project objectives; (3) walk in the customer's shoes; (4) start with the data you have; (5) look for quick wins and incremental gains; (6) get the team right; and (7) win executive sponsorship. He asserts that benefits outweigh the costs that leads to profitable gains in big data. In retail analytics space several big data analytics success stories have been reported. Many retailers use big data analytics to freshen up customer loyalty programs as well as product promotions (e.g., holiday promotions). One such example is that Walgreen's customer loyalty program is a key aspect of drugstore chain's revamp [12]. Having seen the benefits of big data analytics by online retail stores the brick-and-mortar retailers have started playing catchup on big data analytics strategies [12]. Starbucks executives [12] say that their store expansion plan to add hundreds of stores globally was driven by big data analytics results.

In order to deliver business value business managers have to ensure that big data provides quality information to users. This presents a challenge given the complex (mostly unstructured) nature of such data. Big data architecture must be designed in such a way that the data is traceable back to the source of origin. Smith [47] asserts that "there is no point on-boarding great big volumes of transactional data if it is not accurate, nor will ever be accurate." Smith [47] also points out properly organizing data and maintaining hierarchy in data which is one of the most critical parts of design to having simply an adequate solution versus a robust solution. Another key point is to maintain proper order in data integration effort. Smith [47] observes that for many companies data lives in the domains of multiple vendors. So an effective measurement and cohesive data management must be taken into consideration in integrating it into a big data landscape.

VI. CONCLUDING REMARKS

In this paper, an overview of big data characteristics was provided. We presented several successful big data initiatives that were implemented in diverse areas of business, engineering, research and government organizations. Highlighted, was how organizations are driving their growth by taking advantage of big data tools and technologies. We showed how big data started providing benefits to customers in different business areas including marketing analytics [10] and retail analytics [21], in healthcare [18], [23], [33] resulting in positive impacts to the lives of ordinary citizens in society.

To take advantage of big data business and IT managers need to collaborate and come up with some specific plan to generate business value to the organization. We have pointed out several criteria for consideration in starting big data projects. Big data is not just owned by IT department. It needs to be driven and owned by business managers as well. That will allow business organizations to maintain increasing agility, add business value, improve business performance and provide strategic directions. "Smart leaders across industries will see using big data for what it is: management revolution. The challenges of becoming a big data-enabled organization can be enormous and require hands-on or in some cases hands-off leadership" [32].

Further research can be conducted on how to replicate the success of these big data implementations on big data projects in small business enterprises. Activities of big data tools, technology and knowledge transfer need to be identified and articulated to promote them to small firms to allow them to drive growth and stay competitive.

ACKNOWLEDGMENTS

improving this paper. And many thanks to Joan Schnitzer for an excellent editing job.

REFERENCES

[1] S. Berinato, "With Big Data Comes Big Responsibility", Harvard Business Review, November 2014, pp. 101-104, 2014.

[2] S. Akhter, N. Rahman and M.N. Rahman, "Competitive Strategies in the Computer Industry," International Journal of Technology Diffusion (IJTD), Vol. 5, No. 1, pp. 73-88, January-March, 2014.

[3] J. Aldrete, "Big Data's Big Impact on Sales in 2015", Data Informed, December 2014. (http://data-informed.com/big-datas-big-impact-sales-2015/#sthash.GeVyCnKd.dpuf).

[4] N. Balac, T. Sipes, N. Wolter, K. Nunes, B. Sinkovits and H. Karimabadi, "Large Scale Predictive Analytics for Real-Time Energy Management", In Proceedings of the 2013 IEEE International Conference on Big Data (BigData 2013), pp. 657-664. October 6-9, 2013, Santa Clara, CA, USA.

[5] S. Bailey, "3 Steps to Take the Fear Factor Out of Big Data", Data Informed, July 2013. (http://data-informed.com/3-steps-to-take-the-fear-factor-out-of-big-data/#sthash.66oEUw9H.dpuf).

[6] E. Bertino, "Big Data - Opportunities and Challenges", In Proceedings of the 2013 IEEE 37th Annual Computer Software and Applications Conference, pp. 479-480. DOI: 10.1109/COMPSAC.2013.143.

[7] H.S. Bhat, G.J. Vaz and J.C. Meza, "Fast Solution of Load Shedding Problems via a Sequence of Linear Programs", In Proceedings of the 2013 IEEE International Conference on Big Data (BigData 2013), pp. 1-6. October 6-9, 2013, Santa Clara, CA, USA.

[8] A. Chandramouly, S. Sandeen, C. Yalla, Y. Goel, N. Ngo and D. Watson, "Intel IT Best Practices for Implementing Apache Hadoop Software", IT Intel White Paper, Available at: www.intel.com/it, 2013.

[9] R.M. Chang, R.J. Kauffman and Y. Kwon, "Understanding the paradigm shift to computational social science in the presence of big data", Decision Support Systems, Vol. 63, 2014, pp. 67–80.

[10] Data Informed Staff, "Focus On: Marketing Analytics", Data Informed, January 2014. (http://data-informed.com/focus-marketing-analytics/#sthash.XsIgbUvb.dpuf).

[11] T. Chardonnens, P. Cudre-Mauroux, M. Grund and B. Perroud, "Big Data Analytics on High Velocity Streams: A Case Study", In Proceedings of the 2013 IEEE International Conference on Big Data (BigData 2013), pp. 784-787. October 6-9, 2013, Santa Clara, CA, USA.

[12] Data Informed Staff, "Focus On: Predictive Analytics", Data Informed, Retrieved on 01/24/2015 from: http://data-informed.com/focus-predictive-analytics/#sthash.hlEQ1kGY.dpuf

[13] T.F. Dapp, "Big Data: The Untamed force", Deutsche Bank Research, 60262 Frankfurt am Main, Germany, May 2014, pp. 1-37. Print: ISSN 1612-314X / Internet/E-mail: ISSN 1612-3158

[14] W. Fan and A. Bifet, "Mining Big Data: Current Status, and Forecast to the Future", SIGKDD Explorations, Vol. 14, No. 2, pp. 1-5, 2014.

[15] Federal Reserve System, "The 2013 Federal Reserve Payments Study: Recent and Long-Term Payment Trends in the United States: 2003-2012", Federal Reserve System, USA, pp. 1-43.

[16] A. Garzo, A.A. Benczur, C.I. Sidlo, D. Tahara and E.F. Wyatt, "Real-time streaming mobility analytics", In Proceedings of the 2013 IEEE International Conference on Big Data (BigData 2013), pp. 697-702. October 6-9, 2013, Santa Clara, CA, USA.

[17] M. Goldberg, "IBM's Data Science Project to Build Analytics to Predict Heart Failure", Data Informed, October 2013. (http://data-informed.com/ibms-data-science-project-build-analytics-predict-heart-attacks/#sthash.VhWaYuiO.dpuf).

[18] P. Groves, B. Kayyali, D. Knott and S.V. Kuiken, "The 'Big Data' Revolution in Healthcare", McKinsey&Company, 2013.

[19] M. Hernandez, K. Hildrum, P. Jain and R. Wagle, "Constructing Consumer Profiles from Social Media Data", In Proceedings of the 2013 IEEE International Conference on Big Data (BigData 2013), pp. 710-716. October 6-9, 2013, Santa Clara, CA, USA.

[20] R.T. Herschel, "Principles and Applications of Business Intelligence Research", 1st Edition, 2012, pp. 286-300. IGI Global, Hershey, PA 17033, USA.

[21] Data Informed Staff, "Focus On: Retail Analytics", Data Informed, June 25, 2013. Retrieved on 01/24/2015 from: http://data-informed.com/focus-on-retail-analytics/#sthash.sl03j74B.dpuf

[22] D. Kiron, R.B. Ferguson and P.K. Prentice, "From Value to Vision: Reimagining the Possible with Data Analytics", Research Report, MIT Sloan Management Review, pp. 1-19. 2013.

[23] M.H. Kuo, "Health big data analytics: current perspectives, challenges and potential solutions", International Journal of Big Data Intelligence, Vol. 1, Nos. 1-2, pp. 114-126, 2014.

[24] M. Goldberg, "Intel Jumps into Big Data Pool with Hadoop Distribution", Data Informed, February 2013. (http://data-informed.com/intel-jumps-into-big-data-pool-with-hadoop-distribution/#sthash.vvOqDCYo.dpuf).

[25] J. Kwac and R. Rajagopal, "Demand Response Targeting Using Big Data Analytics", In Proceedings of the 2013 IEEE International Conference on Big Data (BigData 2013), pp. 683-690. October 6-9, 2013, Santa Clara, CA, USA.

[26] A. Labrinidis and H.V. Jagadish, "Challenges and Opportunities with Big Data", In Proceedings of the VLDB Endowment, Vol. 5, No. 12, August 2012, pp. 2032-2033.

[27] M. Goldberg, "Teradata Deepens Ties to Hadoop, Unveils In-Database R Analytics", Data Informed, September 2013. (http://data-informed.com/teradata-deepens-ties-hadoop-unveils-database-r-analytics/#sthash.fv6R46OV.dpuf).

[28] S. LaValle, E. Lesser, R. Shockley, M.S. Hopkins and N. Kruschwitz, "Big Data, Analytics and the Path from Insights to Value", MIT Sloan Management Review, Vol. 52, No. 2, Winter 2011.

[29] H. Li, B. Qian, D. Parikh and A. Hampapur, "Alarm Prediction in Large-Scale Sensor Networks - A Case Study in Railroad", In Proceedings of the 2013 IEEE International Conference on Big Data (BigData 2013), pp. 7-14. October 6-9, 2013, Santa Clara, CA, USA.

[30] L. Li, S. Bagheri, H. Goote, A. Hasan and G. Hazard, "Risk Adjustment of Patient Expenditures: A Big Data Analytics Approach", In Proceedings of the 2013 IEEE International Conference on Big Data (BigData 2013), pp. 12-14. October 6-9, 2013, Santa Clara, CA, USA.

[31] D. Loshin, "Business Problems Suited to Big Data Analytics", Data Informed, October 2012. (http://data-informed.com/business-problems-suited-to-big-data-analytics/#sthash.p3QjT2SD.dpuf).

[32] A. McAfee and E. Brynjolfsson, "Big Data: The Management Revolution", Harvard Business Review (HBR), October 2012, pp. 61-68.

[33] T.B. Murdoch and A.S. Detsky, "The Inevitable Application of Big Data to Health Care", The Journal of the American Medical Association (JAMA), 2013, Vol. 309, No. 13, pp. 1351-1352. DOI: 10.1001/jama.2013.393.

[34] R. Nambiar, A. Sethi, R. Bhardwaj and R. Vargheese, "A Look at Challenges and Opportunities of Big Data Analytics in Healthcare", In Proceedings of the 2013 IEEE International Conference on Big Data (BigData 2013), pp. 17-22. October 6-9, 2013, Santa Clara, CA, USA.

[35] M.V. Rijmenam, "The Future of Big Data: Prescriptive Analytics Changes the Game", Data Informed, June 2014. (http://data-informed.com/future-big-data-prescriptive-analytics-changes-game/#sthash.Ry9tZa59.dpuf).

[36] K. Pantelis and L. Aija, "Understanding the value of (big) data", In Proceedings of the 2013 IEEE International Conference on Big Data (BigData 2013), pp. 38-42. October 6-9, 2013, Santa Clara, CA, USA.

[37] J. Rubin, "Survey Demonstrates The Benefits Of Big Data", Forbes Insights, November 2013. (http://www.forbes.com/sites/forbesinsights/2013/11/15/survey-demonstrates-the-benefits-of-big-data/).

[38] D. Pennacchioli, M. Coscia, S. Rinzivillo, D. Pedreschi and F. Giannotti, "Explaining the Product Range Effect in Purchase Data", In Proceedings of the 2013 IEEE International Conference on Big Data (BigData 2013), pp. 648-656. October 6-9, 2013, Santa Clara, CA, USA.

[39] E.L. Porres, "The Big Potential of Big Data: A Field Guide for CMOS", Forbes Insights, pp. 1-22, 2013. (www.forbes.com/forbesinsights).

[40] M.E. Porter and V.E. Millar, "How Information Gives You Competitive Advantage", Harvard Business Review, July-August 1985, pp. 101-109.

[41] N. Rahman, D. Rutz, S. Akhter and F. Aldhaban, "Emerging Technologies in Business Intelligence and Advanced Analytics," ULAB Journal of Science and Engineering (JSE), Vol. 5, No. 1, November 2014, pp. 7-17.

[42] A. Rajpurohit, "Big Data for Business Managers - Bridging the gap between Potential and Value", In Proceedings of the 2013 IEEE International Conference on Big Data (BigData 2013), pp. 29-31. October 6-9, 2013, Santa Clara, CA, USA.

[43] M. Ricknäs, "Big data not just about the analytics," Retrieved on 10/18/2014: http://www.computerworld.com/s/article/9224986/Big_data_not_just_a bout_the_analytics_says_Amazon_CTO, 2012.

[44] P. Russom, "Managing Big Data", TDWI Best Practices Report, TDWI Research, pp. 1-40, 2013.

[45] S. Sharma, "5 Business Areas Poised for Big Data Opportunities in 2015", Data Informed, December 2014. (http://data-informed.com/5-business-areas-poised-big-data-opportunities-2015/#sthash.Mxjab7la.dpuf).

[46] Y. Simmhan and M.U. Noor, "Scalable Prediction of Energy Consumption using Incremental Time Series Clustering", In Proceedings of the 2013 IEEE International Conference on Big Data (BigData 2013), pp. 29-36-44. October 6-9, 2013, Santa Clara, CA, USA.

[47] M. Smith, "For Marketing, Big Data Has No Value Without 'Big Measurement'", Data Informed, October 2012. (http://data-informed.com/for-marketing-big-data-has-no-value-without-big-measurement/).

[48] W. Weill, M. Subramani and M. Broadbent, Building IT Infrastructure for Strategic Agility. MIT Sloan Management Review, Fall 2002.

[49] M. Wheatley, "7 Best Practices for Retail Analytics Projects", Data Informed, October 2102. (http://data-informed.com/7-best-practices-for-retail-analytics-projects/).

[50] T.W. Wlodarczyk and T.J. Hacker, "Current Trends in Predictive Analytics of Big Data", International Journal of Big Data Intelligence, Vol. 1, No 3, pp. 172.

[51] TechAmerican Foundation, "Demystifying Big Data: A Practical Guide to Transforming the Business of Government." Retrieved on 10/18/2014 from: http://www.techamerica.org/Docs/fileManager.cfm?f=techamerica-bigdatareport-final.pdf, 2012.

[52] M.A. Yazici, C. Kamga and A. Singhal, "A Big Data Driven Model for Taxi Drivers' Airport Pick-up Decisions in New York City", In Proceedings of the 2013 IEEE International Conference on Big Data (BigData 2013), pp. 37-44. October 6-9, 2013, Santa Clara, CA, USA.

[53] C. Waxer, "Precision Agriculture Yields Big Data Challenges" Data Informed, September 2014. (http://data-informed.com/precision-agriculture-yields-big-data-challenges/#sthash.2fc3MkeH.dpuf).

[54] K. Zolfaghar, N. Meadem, A. Teredesai, S.B. Roy, S.C. Chin and B. Muckian, "Big Data Solutions for Predicting Risk-of-Readmission for Congestive Heart Failure Patients", In Proceedings of the 2013 IEEE International Conference on Big Data (BigData 2013), pp. 64-71. October 6-9, 2013, Santa Clara, CA, USA.