Int. J. Production Economics 165 (2015) 293–306

Contents lists available at ScienceDirect

Int. J. Production Economics

journal homepage: www.elsevier.com/locate/ijpe

# Managing a Big Data project: The case of Ramco Cements Limited

Debprotim Dutta [a,b], Indranil Bose [a,*]

[a] Indian Institute of Management Calcutta, Diamond Harbor Road, Kolkata 700104, India
[b] Amazon India, Malleshwaram, Bangalore 560055

ARTICLE INFO

*Article history:*
Received 15 January 2014
Accepted 27 December 2014
Available online 7 January 2015

*Keywords:*
Big Data analytics
Cement industry
Data driven decision making
Manufacturing
Outbound logistics
Visualization

ABSTRACT

Currently many organizations are in the process of implementing Big Data related projects in order to extract meaningful insights from their data for better decision making. Though there are various frameworks postulating the best practices which should be adopted while implementing analytics projects, they do not cater to the complexities associated with a Big Data project. In this paper our goal is two-fold: to develop a new framework that can provide organizations a holistic roadmap in conceptualizing, planning and successfully implementing Big Data projects and to validate this framework through our observation of a descriptive case study of an organization that has implemented such a project. Although the manufacturing sector has been slow in incorporating analytics in their strategic decision making, the situation is changing with increasing use of analytics for product development, operations and logistics. We explore the Big Data project at a manufacturing company, Ramco Cements Limited, India, describe the system developed by them and highlight the benefits accrued from it. We investigate the entire process by which the project is implemented using the lens of our proposed framework. Our results reveal that a clear understanding of the business problem, a detailed and well planned step-by-step project map, a cross functional project team, adoption of innovative visualization techniques, patronage and active involvement of top management and a culture of data driven decision making are essential for the success of a Big Data project.



## 1. Introduction

Big Data technologies have made it possible for companies to obtain newer and deeper insights into customer behavior trends and preferences compared to what traditional analytics provided. Consequently, companies are increasingly looking at adopting Big Data analytics in order to better understand their customers and design better products or provide them more customized services. With increase in the volume, variety and velocity of data being stored, collectively known as 3Vs, a large number of new data science techniques for carrying out predictive analytics on them have been developed. This overlap or cusp of Data science, Predictive analytics and Big Data, addressed by the term DPB in the academic literature (Waller and Fawcett, 2013), has generated significant research interest. Big Data analytics is however still at a nascent stage with many companies presently developing a strategy and roadmap around Big Data or doing a pilot project (Miele and Shockley, 2013) to better understand the technology and its benefits. Though traditional database systems are able to handle the volume associated with Big Data, both relational database management systems (RDBMS) based dimensional modeling (Jacobs, 2009) and cube based online analytical processing (OLAP) fare badly when dealing with the "too fast" and "too hard" fronts of Big Data (Madden, 2012). Hence, new technologies are needed for Big Data and this makes implementation of Big Data projects quite complex (Biesdorf et al., 2013). Among all sectors, manufacturing has been particularly slow in embracing digital technologies (Ebner and Bechtold, 2012). A research by Oxford Economics in 2011 revealed that only 25% of industry executives believed that the manufacturing sector would be highly impacted by digital transformation in the next five years, the second lowest among all sectors being considered (Oxford Economics, 2011). Yet, it has been observed that "Every manufacturer has an unbelievable amount of data that is never put to use. They are literally drowning in it, and when they begin to gather it, analyze it and tie it to business outcomes, they are amazed by what comes out" (Records and Fisher, 2014). Big Data and analytics are considered to be important for manufacturing firms and there are examples of companies like Merck using Big Data technologies for developing vaccines faster (Henschen, 2014), Volvo forecasting which component might fail under what circumstances, and Xerox analyzing telemetric data to provide better customer service and reduce costs (Big Data Insight Group, 2012). However, although there exist frameworks for implementation of enterprise resource planning

* Corresponding author. Tel.: +91 33 24678300x157.
*E-mail addresses:* debprotimdutta@gmail.com (D. Dutta), indranil_bose@yahoo.com (I. Bose).

http://dx.doi.org/10.1016/j.ijpe.2014.12.032
0925-5273/

(ERP) projects (Mandal and Gunasekaran, 2003), integrated ERP and supply chain management projects (Bose et al., 2008), logistics information management projects (Kuruppuarachchi et al., 2002), e-Kanban systems for supplier replenishment (MacKerron et al., 2014), we observed that there was a paucity of scholarly frameworks related to Big Data implementation projects. To address this gap, we embarked on this research to develop a framework for implementation of Big Data projects and validate the framework through our observations about a descriptive case study on Big Data implementation at Ramco Cements Limited (RCL).

In the next section we provide a detailed review of the extant research in the area of Big Data and analytics projects. This is followed by the presentation of a new framework that can be used by companies as a roadmap for implementation of Big Data projects. Section 4 of the paper tracks the method and processes by which RCL was able to successfully execute a Big Data and analytics project and accrue several benefits from it. In Section 5, we delve into the possible reasons which contributed to the success of the project at RCL and imbibe lessons from it for any firm attempting to develop similar Big Data based solutions. The limitations of this research are explained in Section 5, along with directions on how to extend this research in future. Section 6 concludes the paper.

## 2. Literature review: implementation of Big Data and analytics projects

Studies have shown that best decisions are taken by business leaders when they are armed with data and tools to gather insights from them (Davenport, 2006). Thus, across all organizations, increasingly, business leaders are taking decisions based on data rather than intuition (Davenport, 2006; Lavalle et al., 2011). A study has shown that organizational factors such as top management support and mandates for analytics, openness to change, new ideas and a strong data oriented culture contribute to the competitive advantage created by the use of analytics (Kiron and Shockley, 2011). There is a growing belief within the business community that an analytics revolution is currently underway which will transform the way organizations run their businesses (Kiron et al., 2013). Companies are using data which has more volume, velocity and variety than any data used so far, and using this Big Data companies are getting a sound understanding of customer behavior and patterns which were unknown to them and more sophisticated quantitative tools are being used to understand the business operations and services being provided to clients (Kiron et al., 2012; Kiron, 2013). Big Data projects primarily pose two challenge classes namely the technological obstacle in managing such huge data and the semantics of finding and meaningfully combining relevant information which makes it a case for multi-disciplinary problem solving (Bizer et al., 2011).

Research has shown that analytics projects can often be considered similar to well defined scientific experiments or clinical trials (Marchand and Peppard, 2013) albeit with a more practical and pragmatic approach (Viaene and Bunder, 2001). Analytics projects are of shorter duration and their discovery cycle aids in better cost control compared to IT projects which are of longer duration and require substantial investment of resources (Marchand and Peppard, 2013). However, we believe successful analytics projects are more difficult to implement compared to their IT counterparts as they involve a significant change in the culture of the organization. Not only do analytical solutions need to be deployed like the projects in IT, but the mindset of employees within the organization (Kiron et al., 2013) needs to change so that they start consuming the analytics or insights being churned out by the deployed tools. This change in mindset within an organization is often difficult to achieve (Ferguson, 2012).

In a survey conducted by the Economist Intelligence Unit, nearly 56% of the respondents belonging to the manufacturing sector reported that their firm had not made much progress in Big Data related projects (The Economist Intelligence Unit, 2013). However, to maintain competitive edge some manufacturing companies are embracing technologies such as advanced analytics and cyber-physical system based approaches to improve efficiency and productivity (Lee et al., 2013). Analytics is being increasingly used to solve problems such as traditional time and space assembly line balancing problems (Chica et al., 2013). Hybrid data mining techniques involving spatial statistics and adaptive resonance theory neural networks have improved yield and reduced costs in the semiconductor industry (Hsu and Chien, 2007) while advanced analytics involving association rule mining has been used to define a good logistics plan for food manufacturing firms (Ting et al., 2014). Thus, the manufacturing sector is increasingly adopting high end analytical techniques. Ebner and Bechtold (2012) have shown that the impact of Big Data analytics on manufacturing and the part of the value chain which is going to be affected by it mostly depends on sub-sectors in which the firm operates.

To the best of our knowledge there is a paucity of an integrated framework supported by a detailed and scholarly case study in the use of Big Data. We propose a framework for implementation of Big Data projects in organizations in the next section. To validate the framework we describe the case study of RCL to understand how this company went about leveraging Big Data using analytics and provided substantially better service to their customers. This case study is viewed using the lens of our proposed framework and lessons learnt from RCL's experience are synthesized.

## 3. A research framework for implementing Big Data projects

There are two major data mining models (Harding et al., 2006) that are popular among analytics professionals: CRISP-DM (Cross industry standard process for data mining) (Wirth and Hipp, 2011) and SEMMA (Sample, Explore, Modify, Model, Assess) (Obenshain, 2004). A study (Azevedo and Santos, 2008) has shown that both CRISP-DM and SEMMA can be considered to be implementation of KDD. However, it is unknown if these frameworks can be applied to projects involving Big Data which are characterized by 3Vs. In the following section we present a framework that can be used as a roadmap for implementing Big Data projects. Fig. 1 presents the detailed framework that is organized into three distinct phases: strategic groundwork, data analytics and implementation.

Strategic groundwork is the phase which lays down the platform for successful implementation of a Big Data project before quantitative data modeling and analytics takes over. It involves identification of the business problem, detailed research and brainstorming regarding the possible solutions, technologies and assessment of their capabilities, conceptualization of the probable solution method to be adopted, formation of project teams and development of a project roadmap with milestones and timelines.

Business problem: A sound understanding of the business process or problem is absolutely critical in order to understand the possible improvements that can be realized from the implementation of Big Data analytics projects and the insights generated from it. Directions from senior management, inputs from various business units who are stakeholders of the project help in understanding the scope of the problem. This is also the right stage to set the expectations of all stakeholders right and dispel any myths about what the Big Data project can achieve for the organization (Kohavi et al., 2002).

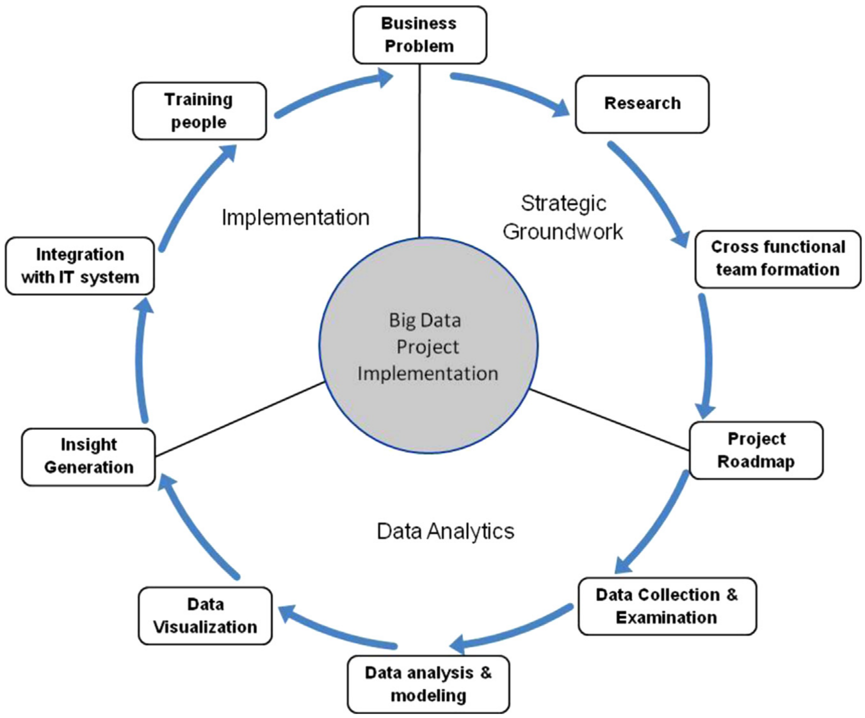


**Fig. 1.** Framework for implementation of Big Data projects in firms.

Research: The purpose of the research is to understand how similar problems have been addressed by other companies and to find out the status of the present IT and analytics infrastructure within the organization which can be leveraged. In addition, research is directed to understand the different standard IT or analytics offerings available in the market, potential vendors, technologies adopted, user experiences, timelines etc. At this stage it is important to avoid the planning fallacy by taking an outside view "which consists in using experience from previous similar ventures" rather than an inside view "focusing on the constituents of the specific planned action" (Flyvbjerg, 2013).

Cross functional team formation: To derive maximum benefit from Big Data projects, a cross functional team needs to be instituted. It should consist of people from various stakeholder business units, IT experts, data modelers and scientists, experts in the field of cognitive science or customer behavior and business decision makers. Inputs from the business units are used by the data modelers to develop generalized predictive models (Schikora and Godfrey, 2003) and obtain trends based on collected data. These are analyzed by experts through the prism of behavioral science to provide critical insights to decision makers. All this is done with the underlying support provided by the IT team.

Project roadmap: Once the problem is identified, and research related to potential solutions is carried out and the cross functional team is constituted, it is imperative that a project roadmap is developed. Ideally this should identify the major activities of the project along with timelines and designated persons assuming each activity and the various milestones. The project map should be flexible enough to incorporate corrections and hence the focus of a manager should be "on project execution and delivery as opposed to adherence to the plan" (Viaene and Bunder, 2001).

Data analytics is the phase which is significantly different for implementation of Big Data projects compared to traditional analytics projects. This phase is technically challenging as it involves modeling of both structured and unstructured data. Efforts are needed for developing innovative data visualization techniques to represent data and discover meaningful insights from it. The key steps are as follows:

Data collection and examination: All types of data that have been historically captured by the organization is examined and analyzed. Various unstructured data such as textual data, social media data, comments from clients, business partners, suppliers etc. which have been sparingly used earlier in any analysis are given significant importance. Often these unstructured data give additional insights that can augment insights generated from structured data. However, the unstructured data may be obtained from multiple social media sources such as Facebook and Twitter pages as well as customer surveys and customer interaction emails. These may contain a lot of irrelevant and duplicate data that will need to be filtered out. The unstructured data may need to go through text tagging and annotation for creation of metadata (e.g., which product, which customer, which location) (Sukumaran and Sureka, 2006). This metadata in turn could serve as a dimension of analysis for the structured data (Baars and Kemper, 2008), thereby helping the firm integrate structured and unstructured data.

Data analysis and modeling: Various quantitative techniques are used on the collected data in order to extract insights. The aim of this step is to identify trends in the data and discover underlying reasons explaining the trend. Often, predictive modeling is used to forecast future trends. The type of quantitative modeling as well as the complexity of computations varies from case to case depending on the type of business problem which is solved at this stage. For structured data analysis, statistical approaches such as regression, discriminant analysis, factor analysis etc. as well as machine learning approaches such as decision trees, neural networks, clustering, support vector machines, rough sets, among others can be used. For unstructured data analysis, text mining can be used for determination of important concepts (Coutinho et al., 2013) and their interrelationships as well as extraction of important terms (Abrams et al., 2013) related to concepts and analysis of sentiments expressed about the determined concepts (Kang and Park, 2014).

Data visualization: This involves developing innovative visual representations of the data which can help in insights generation. While tables are a natural way of representing data, they are not very intuitive for detecting similar and anomalous patterns in data. A more natural way of detecting such patterns is to use visualization techniques. Depending on the business requirement, standard data visualization tools can be adopted with modifications or entirely new visualization tools need to be developed. For analysis of high volume of structured and unstructured data, two types of visualization techniques, namely content terrain maps and wide widgets, are likely to be quite useful. While content terrain maps (e.g., heat maps, thematic maps) show different types of data on a geographical map and allow user interaction with that data for detection of anomalies, wide widgets show "an interactive, visual structure that mirrors some central spine in the information" (Rao, 2003) and allow navigation of graphs or hierarchies that often emerge from unstructured data. Since the visualization used for analysis of Big Data is likely to be complex this phase can include creation of a knowledge base that captures expert visualization capabilities (O'Leary, 2013).

Insights generation: This is the last logical step where the data is analyzed to understand the trends and underlying latent reasons behind those trends. This analysis is useful for uncovering likely future scenarios. This step involves the transformation from analysis of data to insights and actionable business inputs that are of value to the organization (LaValle et al., 2011).

Implementation: This is the last stage of the Big Data project wherein the models, trends and visualization tools are deployed and integrated with the present IT system and thereafter training is provided to the managers to use the newly developed system in their day to day business decision making.

Integration with IT system: A strong IT infrastructure consisting of robust architecture, efficient systems and well laid out processes is an absolute necessity for implementing Big Data projects. The newly developed model needs to work in sync with the existing IT systems; hence it is essential that the integration process is carried out meticulously and exhaustively tested in order to remove any bugs. Success in integration may lead to improved business performance as well (Bose et al., 2008; Mekawie and Elragal, 2013).

Training people: Usually, most employees are not very comfortable to tap into a new source of insights while taking their business decisions. Hence, users need to be "trained in how to use the tools and the data that is available" and to dispel their apprehensions and showcase the benefits of the newly developed system they should be "given access to people who can help them" (Watson and Wixom, 2007). However, in many cases training does not imply automatic acceptance by prospective users. In such situations, some amount of prodding to accept and proactively start using the new system or even disincentivizing the users from using the earlier legacy system is necessary to obtain traction.

An important contribution of the proposed framework is that it combines the change management aspects of an IT project management framework with the data management aspects of an analytics framework. Models that focus on change management facilitators such as organizational requirements, project organization structure, supporting team, training, among others (Kuruppuarachchi et al., 2002) highlight the importance of change management during IT project implementation at an organization. However, they do not delve into how to store, manage and represent data gathered during the implementation of IT projects. On the other hand, the focus of analytics models like CRISP-DM and SEMMA is primarily on data modeling and application of other quantitative techniques on the data to solve business problems. They do not emphasize the change management aspects associated with a major IT project. Unquestionably, the data modeling and analysis is one of the most significant aspects of a typical analytics project. However, there are certain other equally important aspects in Big Data projects such as cross functional team formation, training of people regarding how to use the system etc. At the same time, complicated Big Data projects involve different types of data, significantly different technologies for storage and analysis, innovative visualization tools to extract insights and involvement of cognitive scientists in the team to convert insights generated into meaningful business explanation. Hence, we believe a more integrated framework, like the one shown in Fig. 1 that improves on the data management aspects of analytics frameworks and incorporates change management aspects of IT project implementation frameworks, is required for present day Big Data projects. This generic framework has the potential to be adopted for implementation of such projects by firms of any size and capability, belonging to any industry, and operating in any geography.

### 3.1. Case method

In order to validate the proposed framework we chose to conduct an in-depth descriptive case study of an organization that has implemented a Big Data project since the case method is quite popular in business research (Lee, 1989; Dubé and Paré, 2003). The objectives of this case study included:

- understanding how an organization has carried out such a project;
- identifying the challenges it has faced in the planning and implementation of the project;
- discovering steps it has taken to ensure success in the project; and
- learning about the benefits it has derived from the project.

We chose a single case study since we wanted to understand the details involved in the project and also because there were not too many examples of successful implementation of Big Data projects that we could find and access. One of the authors created a semi-structured questionnaire based on preliminary understanding of the project from secondary research. This was distributed to the interviewees, which included the CIO and another member of his implementation team. The interviewees were requested to provide written or verbal feedback. Some written response was gathered from the interviewees and this helped in creating a logical chain of evidence (Benbasat et al., 1987). This was followed by detailed face to face interviews over multiple hours with the interviewees. The interview session also included live demonstration of the system developed by RCL so that the authors could understand the capabilities of the system. The authors also closely examined materials from the company's website and from other related websites.

## 4. The case of Ramco Cements Limited

Ramco Cements Limited (RCL) is the flagship company of the Ramco Group of Industries, which was founded in 1938. The Ramco Group consists of multiple lines of businesses including

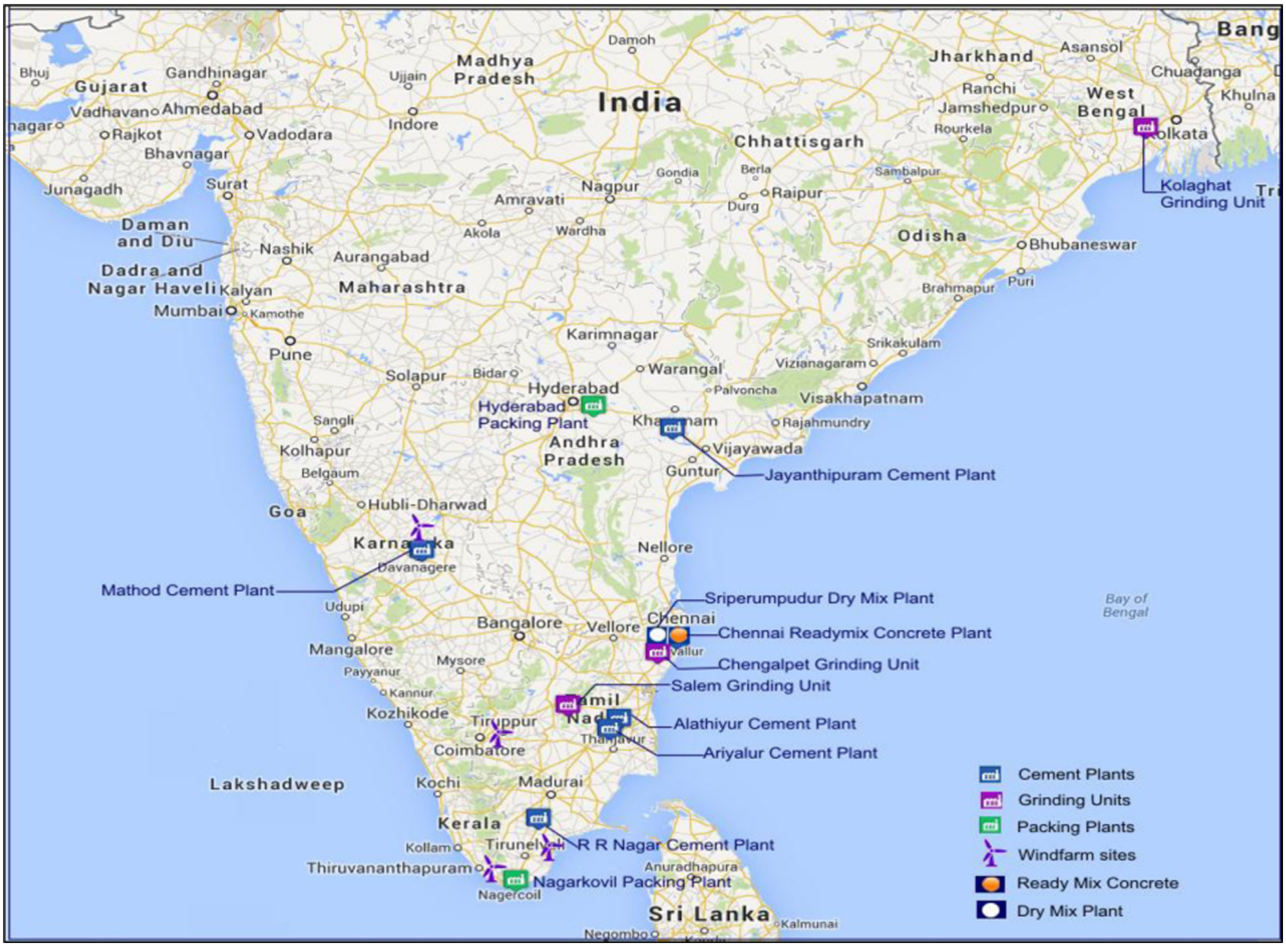


**Fig. 2.** Location map of various operating plants of RCL.

textiles, cement, fiber cement products, software, surgical cotton, and bio-technology. The group grew into a massive organization with annual revenues of US$ 800 million and achieving international recognition for its quality products and services. The Ramco brand enjoys a very strong recognition in southern India and is increasingly looking at expanding its footprint in other parts of India.

RCL was started in 1957 with its first cement manufacturing plant in R R Nagar. RCL primarily produces "Portland cement" and caters to the southern Indian market. Presently, the company has 5 cement plants, 3 grinding units, 2 packing plants, a dry mix plant and a ready mix concrete plant spread all over India with a cement capacity of 16 MTPA. In addition to cement manufacturing it has six captive wind mill sites with a combined capacity of 159 MW. RCL is primarily a trade player with a large dealer network with a very small proportion of its sales to large institutional builders. Fig. 2 shows the locations of manufacturing plants, grinding units and captive wind mills that RCL currently owns. It highlights that most units of RCL are located in southern India, the only exception being the Kolaghat grinding unit located in eastern India.

### 4.1. History of IT at RCL

RCL has always been a technologically savvy company with the adoption of latest techniques and tools. It is a pioneer in the cement industry for introducing latest technologies like fuzzy logic based software systems for process control, pre-calciner technology, Programmable Logic Controllers, Surface Mining Technology, Vertical Mills for Cement Grinding etc. The company launched IT projects at each of its plants in order to help monitor and control costs under the leadership of Mr. Dharmakrishnan, the present CEO. Mr. Dharmakrishnan, who was the CFO then, envisaged the use of IT as a tool for competitive differentiation and greater efficiency and transparency, and encouraged the management team to develop newer and more innovative IT strategies. RCL was one of the first in the cement industry to invest heavily in developing IT infrastructure. Fig. 3 depicts the evolution of IT within RCL. RCL's primary objective was to use IT to bring transparency to the transactions across the organization. Having a centralized server with a single consolidated database, RCL ensured that there were no discrepancies with regard to data and other inputs from various divisions. This helped in making objective and informed decisions.

RCL deployed an ERP solution from Ramco Systems to standardize reporting across plants. However, initial efforts of deploying an ERP solution failed due to a number of issues which included lack of customization of the system to the needs of RCL and inadequate training provided to employees in using ERP data for decision making, among others. To overcome the setback, RCL designated Mr. Dharmakrishnan, then CFO of RCL, as the project champion responsible for leading the entire initiative. In addition, RCL had to undergo a major cultural change in order to promote greater co-ordination and sharing of data between teams and plants. A process oriented approach was initiated that included clear communication about the pitfalls of the system compared to the potential benefits of the new one, thereby slowly exposing

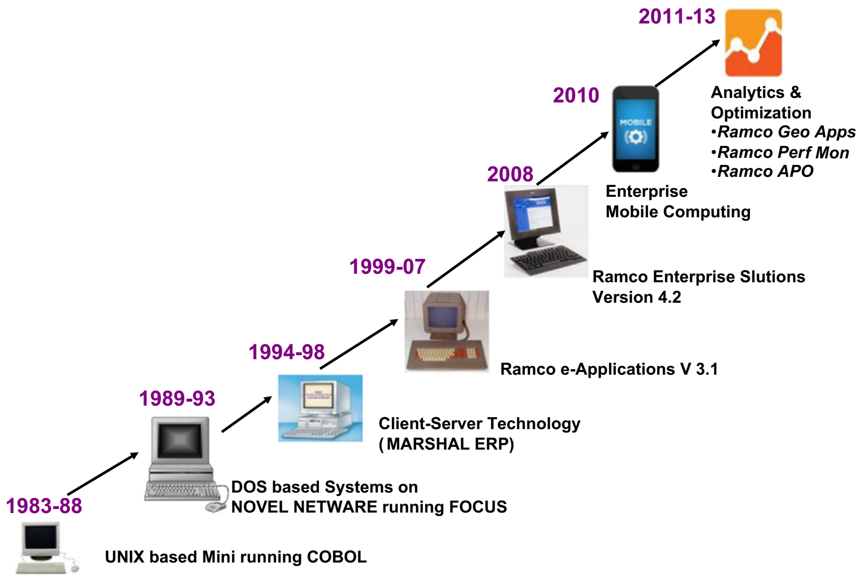


**Fig. 3.** Evolution of IT within RCL.

employees to new ideas and encouraging them to learn and use the system. Consequently, RCL was able to successfully deploy an ERP system in its second attempt.

However, things were not exactly smooth postimplementation. For example, most users started using the ERP system in parallel with the old legacy system and this affected productivity. Senior management along with the Chairman took proactive steps to persuade employees in using the new system while IT teams went through each computer to remove traces of legacy systems physically. Gradually the new ERP system became a part and parcel of the decision making process. The benefits reaped by RCL from the implementation of their ERP system included improved transparency, better customer satisfaction, increased plant efficiency, reduced costs and high profit margins.

Buoyed by the success of the ERP implementation, RCL moved from their decentralized and desktop based ERP 4.0 system to a web based and centralized system. A "Beyond ERP" initiative was launched that included developing an in-house enterprise mobile computing system which enabled processes such as ordering, invoicing, and pack-slipping through PDAs and mobile phones, development of customized MIS reports for mobile devices, deployment of an interactive SMS platform for handling customer queries, enhanced security and compliance features for the entire ERP system, development of an asset tracking system, a method for electronic paperless permission approval process and an interactive voice response system for database management.

### 4.2. Big Data initiative at RCL

Innovation in IT which had become a way of life at RCL propelled the IT team to focus on developing methods and tools to analyze the vast amount of data being captured by their ERP system and obtain meaningful insights that could be used in making more intelligent business decisions. The data being captured at RCL could be broadly categorized into two distinct categories namely: operational data in the form of key process parameters which measured the health, stability and well being of the various processes being run at different plants across RCL and data captured in ERP from different business processes like outbound logistics which included details of customer orders, their past track records, credit history, payments, procurement, status of order shipments etc. The operational data was mostly restricted to decisions taken at the plant and were not taken into account for making strategic business decisions. On the other hand, the data from business processes like outbound logistics captured on a real time basis by the ERP system needed to be analyzed better to extract useful insights such as customer behavior patterns, tracking of key performance indicators (KPIs) for various dealers etc. This data was not similar to traditional data captured in manufacturing organizations. RCL had an elaborate automated system to collect all data electronically as the business process was carried out and stored in the ERP database. Even, non-ERP processes such as weigh bridge, packer, attendance system were integrated with ERP. Thus the errors due to manual entries were minimized. The analytics system picked up from this unified database only and there was no data entry deliberately for reporting purposes to ensure transparency. From a volume perspective the characteristics of the data being captured at RCL are shown in Table 1.

As shown in Table 2, the SCADA and real-time system feeds generated data about various process parameters through different sensors at a massive speed.

Table 3 shows the variety of data being captured by RCL.

#### 4.2.1. Business problem

RCL was running all its business processes on an ERP application developed by Ramco Systems. The system captured voluminous data systematically and stored that in its integrated database. There was an MIS that was built on this database for guiding executive decision making. While the existing, conventional MIS with more than 1200 reports in the form of tabular reports, drill downs etc. gave excellent support to executives, the CEO Mr. Dharmakrishnan felt that there was a need for a system that supplemented this MIS but was tuned more towards key performance monitoring. A strategy session was organized in September 2011 by Mr. Dharmakrishnan with the participation of IT and key management executives where he explained his vision and ideas for the new analytics system that was vastly different from the existing system. The main requirement was to provide a system that listed performance goals and used visualizations to compare actual achievement with respect to such goals. It was meant to bring out critical differences, deviations and geographical significance of the underlying data. It was expected to render well in laptops and tablets and other mobile devices. RCL identified a need for sophisticated logistics optimization as a means to reduce outbound logistics costs which is a substantial component of the total cost in the cement industry. With growth of business, the complexity could not be handled by the present legacy system through a static Excel based solver. A more sophisticated solver was the need of the day. Hence, the business requirement was a comprehensive real-time metric based performance measurement system capable of analyzing performance across multiple levels such as districts, regions, zones, dealers etc. using data from diverse sources, an innovative visual depiction of the analysis which eased the process of insights generation and a sophisticated cost optimization tool with advanced features to reduce the outbound logistics cost for each order. In addition this should be navigable across various browsers and different devices using different operating systems. With these requirements, Mr. Dharmakrishnan advised the IT team to launch the new project and execute it in the shortest possible time and assured his personal support.

*4.2.2. Research*

After understanding the business problem through repeated interviews and interactions with the business representatives, the IT team undertook thorough research. It involved two types of research namely internal as well as external. The internal research centered on existing systems of business reporting, their vulnerabilities, pain areas faced by users and developers, data volumes involved, additional requirements of analytics in terms of information as well as end user devices etc. The idea was to determine how much of existing backend procedures could be re-utilized. The methodology consisted of data collection through one-to-one discussions with key end users, historic data analysis using call logs from help desks, end user behavior analysis through a set of automated processes, user surveys etc. Efforts were made to standardize key performance indicators (KPIs) that could be used for performance assessment and evaluation. The team collected nearly 800 KPIs and short listed about 200 from them and developed calculation procedures for each of them, customizing them as per need.

External research was directed to evaluating whether the standard analytics tools that were available in the market were a good fit for the business requirement on hand. The team visited various vendor sites and went through a number of demonstrations to understand the analytics offerings in the market and the degree to which they could be customized for RCL. Significant efforts were put into the literature study, internet research to collect information on prior experience case studies, and on technologies which could be potentially deployed such as in-memory processing, real-time reporting, Google Maps API, charting tools, SVG Graphics, HTML5 Canvas etc. Another aspect of the research was determination of how to securely render data to the user's devices such as iPads and tablets. There were a lot of technologies available like cloud platforms, VPN, mobile device management etc. in addition to numerous operating systems (Android, Windows, iOS etc.). Since different devices behaved differently, device compatibility was a critical issue. Lastly, the team undertook research on security issues related to the data. The project required that any authorized user from any corner of the globe could access the system on a $24 \times 7$ basis which made the present legacy security system inadequate and vulnerable. Extensive research on technologies like one time passwords, mobile device management, application controls on mobile devices, VPN connectivity for mobile devices was carried out.

*4.2.3. Cross-functional team formation*

The project was entitled “Data Analytics and Optimization” and RCL formed a cross functional team drawing from various functions and different business units apart from the IT team, under the overall guidance of the CEO. In addition, RCL was supported by Ramco Systems in the development process. An implementation team of 12 people was formed with 6 from IT Team, 2 from Ramco Systems and 6 from functional areas like Marketing. Additional

**Table 1**
Volume of data being captured by RCL.

| Description | Data volume | Unit |
|---|---|---|
| ERP database size | 800 | GB |
| Other databases | 300 | GB |
| Assets database | 100 | GB |
| Database growth | 200 | GB/annum |
| **Business processes** | **Transactions** | **Unit** |
| SCADA system | 7 | Billions per day |
| ERP transactions | 4.5 | Millions per annum |
| Enterprise SMS platform | 4 | Millions per annum |
| Enterprise mobile computing | 1 | Millions per annum |
| Sales force application | 1.5 | Millions per annum |
| Cement packer system | 140 | Millions per annum |
| Weigh bridge system | 1.2 | Millions per annum |
| Real time system | 1580 | Millions per annum |
| Near real time server | 16 | Millions per annum |
| Enterprise mail system | 7.3 | Millions per annum |
| Wind mill system | 85 | Millions per annum |
| Gamma metrics | 2 | Millions per annum |
| **Total** | **1842.5** | **Millions per annum** |

**Table 3**
Variety of data being captured by RCL.

| Structured data | Unstructured data |
|---|---|
| ERP database | Salesforce system call details |
| SCADA feeds | Streaming video – plant monitoring |
| Real-time system feeds | Excel sheets |
| Production system logs | Powerpoint presentations |
| Push SMS system | Advertising material including video |
| Mobile computing data | Land documents |
| Tables/reports | Hard copy approvals |
| Charts | Secretarial related docs (Board meetings) |
| Google map representations | Mobile |
| Photographs | Web site |

**Table 2**
Velocity of data being captured by RCL.

| Description | Velocity |
|---|---|
| SCADA system feeds | 250 ms from about 20,000 sensors |
| Real-time system feeds | 30 s feeds for about 2000 parameters |
| Near real-time system feeds | 1 h feeds averaged from above feeds for analytics |
| Plant monitoring | $24 \times 7$ streaming video from about 200 cameras |
| Database table transactions | 10 transactions per second in ERP |
| System batch jobs/ETL | 300 batches per day |
| Land documents | Periodic |
| Geolocation | Periodic |
| Master data | Periodic |

people were allocated for some specific functions such as development of stored procedures. The project team identified a dynamic, technology savvy person within each business unit who was approached and given unofficial responsibility of introducing the new system within his own business unit. This was a clever step as it implied that there was an internal person within each business unit who was pushing the cause of the IT team and helping them overcome the challenges which arose from the business unit.

#### 4.2.4. Project roadmap

The following steps were taken to developing the detailed plan for the project: (a) Developing key project requirements, (b) Creating a roadmap of the project with inputs from critical stakeholders, (c) Dividing responsibilities among the project team, (d) Data modeling to suit the requirements, (e) Developing a list of deliverables for the project, (f) Conducting team workshops to ensure that members were familiar with objectives and deliverables, (g) Creating workable project schedule and outlines all project milestones, (h) Getting approvals and the budget for the project, and (i) Setting a progress reporting system for the team – monthly, weekly or daily reports.

#### 4.2.5. Data collection and examination

The team tried out a number of standard products available in the market, many of which were based on the "cloud" model. Such options were discarded in view of confidentiality of data in addition to problems of interfacing with the existing ERP and BI system. There was a general consensus that the development effort would be larger if off-the-shelf products from the market were used. In the end, the team decided to go for in-house development. However, a lot of inputs and interactions were required from the existing ERP vendor. Moreover, there were security aspects which were too complex for the internal IT team. Hence, it was decided to involve ERP vendor (Ramco Systems) in the project. The role of Ramco Systems was limited to integration with the existing legacy ERP system whereas the RCL team was responsible for developing underlying stored procedures. The optimization project was allocated to the ERP vendor as well. A more technical description is provided in Appendix A.

#### 4.2.6. Data analysis, modeling and visualization

The project contained 3 important components: (a) Ramco GeoApps, (b) Ramco PerfMon, and (c) Ramco APO. The project started in November 2011 with the GeoApps Module. Being a new concept and considering the high amount of field work (in capturing Geo codes) involved, it was planned to be completed in 6 months. The Ramco PerfMon project started in October 2012 and took 6 months to complete. The APO project started in Jan 2013 and is expected to go live from Jan 2014. The work streams, timelines and schedules were identified by the project team in a team workshop in Nov 2011.

Ramco Geoapps: This analytics tool is based on the Google map interface. The idea behind this application was to capture each and every customer's and cement outlet's geolocation and link that data to the critical organizational information residing in the ERP database. With such a set up, data analysis could provide insights to issues related to areas such as business expansion locations, customer off-takes, performance of sales persons, tracking of cement dispatches through wagons, competitor analysis and so on. RCL's other assets such as wind farms and vast tracks of land that were used for manufacturing and mining were also captured and linked to a comprehensive asset management system. Every customer location is mapped out and color coded based on some KPI such as sale volumes, outstandings, growth etc. and on drill down gave details such as location, mobile numbers, history of sales, sales man's visits, sales growth etc. for further analysis. In addition to company's outlets, even the competitor's outlets were mapped. This gives insights on where RCL is weak and where to open its outlets. Fig. 4 shows the GeoApp module where locations of customer are captured and mapped onto a geographic map to understand their spatial significance. The color code is based on their growth. More details about the customers and their order history could be drilled down.

Ramco Perfmon: This was a performance dashboard for executives depicting dealers' KPIs (such as delivery times, sales targets, collections, sales growth etc.) and their achievements. The project used extensive data visualization techniques for interpreting data through graphics, color-codes, geo-mapping etc. paving the way to make complex data into simple, consumable information sets. The system assigned color codes to denote a range of performance. The dark red and orange spots quickly pointed RCL to areas of weakness and anomalies. Various data visualization techniques used in Big Data analytics such as color coding, heat maps, thematic maps were used to depict the findings on KPIs for executives across multiple geographies. Fig. 5 shows the performance of particular areas with respect to a particular KPI using color banding to signify performance levels.

Ramco APO: This was an advanced planning and optimization project primarily designed for outbound logistics optimization on a real-time basis. When an order came in, it could be serviced from multiple sources (e.g., factories, grinding units, godowns) but the logistic costs varied across the sources. Using Ramco APO, every incoming cement order was analyzed with respect to the availability of stock at various sources as well as the logistics cost and other constraints and the best possible source was identified on a real-time basis. The Ramco APO interface consisted of a detailed planner view, an overall macrolevel snapshot of the dispatch operations, and a detailed summary of truck analysis, among others. Fig. 6 shows a snapshot of the factory dispatch dashboard which provided a one stop broad overview of the entire operations pertaining to outbound logistics at any particular time. This type of information was useful for the senior management to gauge the health and status of RCL's outbound logistics operations. Further drill-down allowed a manager to capture relevant details about the trucks out for delivery including their current location, status, expected arrival time, weight being carried, type of truck, ownership details and reasons for delay (if any). Such a system helped in pin pointing the exact reasons which might have led to delay in delivery and was expected to increase efficiency, reduce cost and provide better service to the customers.

All stakeholders across RCL were thoroughly impressed with the new system as it made their job easier. The Google map based KPI performance dashboard helped in improving coordination between the marketing head and the field sales force agent. A regional marketing head explained:

> "*With the introduction of Google map and KPI based analytics system, there are definite improvements in our regular interactions with field salesforce and customer. I like the new KPI dashboard as it is more visual – charts, maps, gauges, diagrams, color coding and so on. The interactive visualizations are worth billions of bytes and it is faster for me to grasp the meaning of many data points that are displayed in color codes, charts and graphs rather than poring over piles of spreadsheets or viewing pages and pages of rows and columns of reports*".

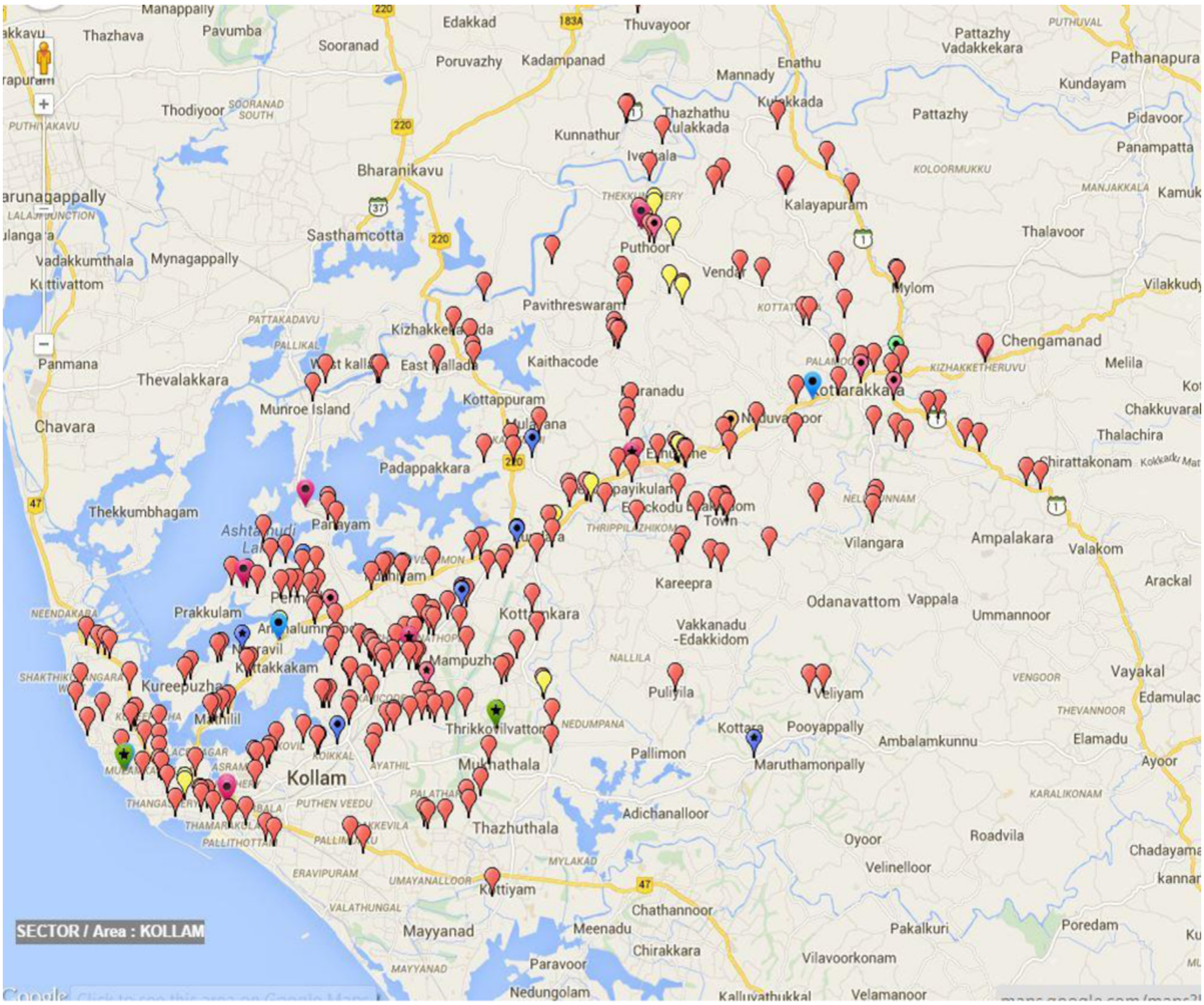


**Fig. 4.** Visual depiction of GeoApp module.

#### *4.2.7. Insights generation*

The project has been a massive success with stupendous benefits. RCL reduced demurrages, penalties, losses and damages arising out of wagon clearance process by up to 70%, and improved consignment clearance time by up to 40%. Since implementing Google maps, the company has successfully captured 20–30% market share in under developed markets by appropriately identifying them and taking a proactive marketing plan to improve sales at those locations. Some relevant KPIs that showed distinct improvement in one year's time in a typical region are as shown in Table 4.

Close monitoring of the weigh bridge data resulted in reduction of truck unloading times. At present 75% of trucks are unloaded within 2 h, whereas earlier only 58% of trucks were unloaded within 2 h. Another important area of improvement is related to accurate determination of weights of cement bags. The Cement Packer is integrated to the system and every bag is monitored for weight deviation. With such arrangement, weight variation in bags came down from 40% to just 2%. This implied a significant recurring cost savings of nearly USD $2 million per annum.

The system has changed the way the decisions are made at RCL. Instead of spending energy on all and sundry, the organization is very focused and informed decisions are made by concentrating on critical items indicated by “Outliers” (Dark reds in the thematic scale). More importantly, the review meetings have become objective as discussions are based on online transparent information, that is authenticated and came directly from the ERP database. The data users are convinced and there are no arguments about the genuineness of the data itself. The time is spent more productively for discussing action plans rather than checking the veracity of the data. The system has also improved the customer satisfaction levels, as seen from the customer dropout rates (53% less than previous year) and the general feedback, mainly due to reduction in delivery times (51% less than previous year) and better response of sales people. The tool had an exception module that alerted management immediately when any performance parameter downgraded in color (say a green moved to an orange, indicating that performance was dipping) and prompted them to take corrective action immediately. There was also an in-built issue tracker that was used to raise queries to concerned sales person (through email/sms) based on the depicted data and to track responses/action taken. For example, suppose a sales manager viewing the dashboard found some discrepancies or significant problems and wanted clarifications from his sales officer or expected some action to be taken, then (s)he could use the system to send email or SMS and also track the action taken by the recipients.

RCL's sales and marketing divisions also benefited from this system. The field sales team now had the capability to easily view information on their mobile devices – information ranging from

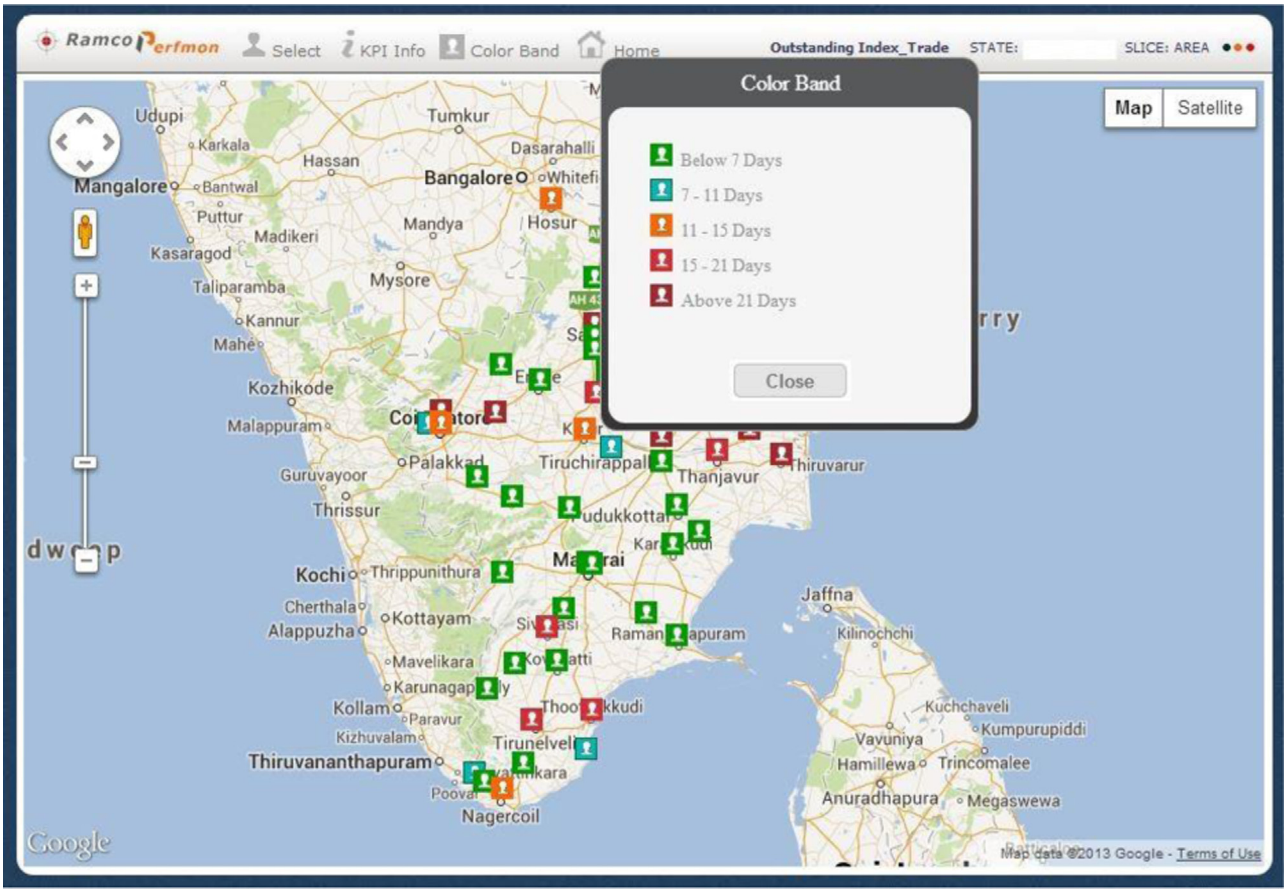


**Fig. 5.** Performance dashboard with geographical significance and color banding.


**Fig. 6.** Ramco APO overall despatch summary dashboard view.

**Table 4**
Benefits achieved from the Big Data project in state of Kerala.

| | |
|---|---|
| Sales growth | 2% (Increased) |
| Order delivery time | −51% (Reduced) |
| New customers | 13% (Increased) |
| Customer attrition | −53% (Reduced) |

competitor distribution networks in their area, to the best and worst performing dealers and sales persons. This not only enabled them to make better sales strategies on-the-go, but also helped management improve their market penetration strategy based on competitiveness in a particular market. They could now locate their cement warehouses and outlets and key customers on Google maps vis-à-vis critical operational information, empowering them to devise effective supply chain strategies and enhance customer service. The Google maps based system allowed them to visualize data at various levels of detail, for example, which regions were experiencing strong growth, which dealers were performing best, or had the most potential for growth. Exception reports and automated SMS messages and alerts generated by the system forewarned the concerned officials about significant deviations so that (s)he could take preventive action.

The stupendous success of the project has been noticed by the senior management. In the words of Mr. Dharmakrishnan, CEO RCL:

> "*Managers who are making important decisions were unable to identify patterns in data without clean data visualisation. Hence we decided that the new system would be simple, colorful, visual, very transparent, natural and easy to adapt. It should help identify outliers, significant patterns, deviations, trends etc. on any devices including mobiles, iPads and tablets*."

The cultural change seen at RCL has also been noticed by its customers who have been impressed with the way RCL operated compared to their peers. A large cement dealer in Chennai stated:

> "*The way the sales people from Ramco Cements interact with us has changed these days. Previously, they used to discuss with us orally or with printouts. Now they come with their tablets and present all information on Google maps and dashboards, which we understand clearly. They explain about growth and de-growth, our historic performances, how to improve sales – all with charts, maps and diagrams very convincingly. I am also guided through future planning for improving business. No other cement company that I deal has such useful analytic system*."

#### 4.2.8. Integration and training

Once the system was developed it was fully integrated with the old legacy IT system. Following successful integration employees were trained in order to reduce their resistance to the new system and to make them well aware of the capabilities of the new system. RCL organized training programs under the title "Data based decision making" and case studies and a management game (Lemonade game) were used to highlight the importance of data in decision making. A demonstration of the new analytic system was provided to the participants, who were later apprenticed in using the new system.

### 4.3. Future planned IT initiatives

The success of the project has given the IT team at RCL a great deal of confidence and have prepared them for further refinements by adopting more innovative methods, and getting into the realm of predictive modeling. A lot of focus in future would be on modeling, forecasting and predictive analysis. Also, with an increase in the number of people using this system RCL had the plan to improve the speed of mobile based computing and make the system more reliable. In future RCL might transform its Data Center to a "Cloud Data Centre" and provide their services on a cloud platform. It also had major plans to extend this analytic system for other KPIs as well (e.g., production, maintenance, mining, inbound logistics, etc.).

## 5. Key lessons learnt from the Big Data project

Big Data projects were inherently complex and required considerable efforts both in terms of technology, manpower and processes to derive the maximum benefit out of it. To quote Mr. Varadarajan, Assistant Vice President IT at RCL:

> "*Big Data analytics is a complex phenomenon. It can't be done by mere installation of powerful hardware like a 16 processor 32 core server with 1 Terabyte RAM. We have to match high performance hardware with appropriate software, database, stored procedures and analytic tools suitable for the purpose. It may require entire rewriting of existing code, procedures and jobs*."

Hence, a Big Data project should be thoroughly conceptualized, researched and planned in detail to ensure smooth execution, successful deployment and adoption by the firm.

In this project, RCL did an impressive job in doing the "strategic groundwork" for the project. As elaborated earlier, the IT team held several rounds of discussion with various business stakeholders and end users of the applications to understand the business requirements from the system. In addition, they undertook extensive research to examine the various possible business solutions and evaluated their compatibility with the present legacy system. It was absolutely imperative that complicated Big Data projects involved a cross functional team that had representation from different business units, stakeholders, IT team and cognitive scientists. With the exception of cognitive scientists, RCL formed a typical cross functional team for this project. They took steps to lay down the detailed project roadmap along with likely timelines and assigned responsibilities for the same. This key step ensured that the status of the project could be easily tracked against the prescribed milestones and corrective actions could be taken in case it was running behind schedule. Regular review meetings and status checks helped in understanding the ground realities.

The data analytics part of a Big Data project usually varied on the basis of business requirements. In this particular initiative at RCL, a lot of focus was on developing sophisticated and meaningful performance dashboards for various dealers, representing them on a Google map and color coding them in order to understand their relative performance. In addition, the focus was on developing innovative visual depiction techniques that could merge the ERP data with the geolocation of various customers in order to provide actionable insights to the decision makers. Modeling could have been built into the entire system when it was being developed. However, with future plans of developing a predictive modeling application in the near future, there were opportunities for significant rework.

There were plenty of difficulties which RCL faced while implementing the system. The biggest challenge was the change in the mindset of its employees with regards to the new system and the openness to learn and embrace it in their day-to-day operations. To ensure participation the company came out with an "Analytics utilization report" that measured how many times an executive accessed the ERP system in a given month. In all review meetings and training sessions, it became mandatory to use only this system (any other media such as excel sheets and power point presentations are prohibited) for reviewing performance and ensure objectivity and transparency A key metric called hit ratio measured the actual hits for the system over a month to the targeted hits per month (60 hits per month). This KPI was monitored and those below 60% were issued an alert. If the hit ratio went below 50%, an exception mail was circulated. If some senior managers were found to be not using the system adequately over a long time, a letter went from the RCL Chairman, gently reminding them on investments made for making better decisions. Two months after the implementation, the average hit ratio increased to nearly 70%. Other post-login user behavior tracking metrics such as number of reports seen, time of seeing every report, queries made, total user session time etc. were available in the system and kept disabled in order to reduce system execution time. They were enabled only for short times on receiving specific requests from managers.

## 6. Limitations and future research

While we proposed a general purpose framework for the implementation of Big Data projects in organizations, we were able to validate this framework through the interpretation of a single case study on RCL. Drawing of inferences based on one single case study could be fraught with its own bias. However, the choice of a single case study was justified in this situation because implementation of Big Data projects in India and in the manufacturing sector is extremely rare and the authors were fortunate to have “access to a situation previously inaccessible to scientific observation” (Yin, 2003) by getting the chance to conduct interviews with RCL executives who were directly involved in this initiative. However, the contribution of the framework can be improved further, if researchers are able to conduct similar in-depth case studies on implementation of Big Data projects in other organizations belonging to different industries and located in different geographies. Such case studies can be compared to the RCL case study in future.

Another limitation of this research was that as per Big Data standards, the volume of data currently being handled by RCL was not as large. However, we needed to look at it from the perspective of the sector which was generating this data. Traditionally, manufacturing organizations have been slow to adopt IT, data capturing and analytics practices. In addition, many manufacturing organizations operated in a B2B environment whereas most of the sectors generating Big Data arose from B2C environment. Hence, if we examined this data using the prism of a manufacturing organization the volume of data being generated by RCL was indeed significantly large. If RCL was able to scale this Big Data project to include process related data (e.g., SCADA data) in future then the volume would increase substantially. If such efforts are made in future, it would be worthwhile to document and compare those initiatives with the present one.

In the project implemented by RCL, the design team did not use any unstructured data analysis as it was not present in the original requirement for the project. However, the need was felt by the design team only after the implementation as unstructured data was emerging in large quantity from the various applications. The unstructured data took the form of market visit remarks by field sales force, customer feedback in the form of qualitative responses to open ended questions in surveys, various emails on different topics of importance, internal blogs, minutes of meetings, etc. In the second phase of the project RCL had plans to include the feature to capture, store and analyze unstructured data by incorporating a text mining and analytics engine. This engine would use natural language based tools, metadata and tags to analyze the unstructured data and present insights. An important direction of future research could focus on understanding the processes, challenges, and impacts of using unstructured data analysis.

## 7. Conclusion

The study of the Big Data initiative at RCL was important from a number of different perspectives. It showed that traditional brick and mortar industries which primarily produced commodities in a B2C or B2B environment could reap immense benefits from Big Data technologies. However, to do so, the firm must develop a comprehensive and detailed roadmap encompassing development of the strategic groundwork, actual data modeling and analysis and finally implementing the developed solution. The new theoretical framework presented earlier was a holistic, easy to understand but detailed roadmap which could be followed by organizations planning to implement Big Data analytics projects in future. We believe this was not only a good theoretical contribution and extension to the existing frameworks but was also a more integrated and implementable framework. Based on this case study we could conclude that the critical success factors for a Big Data project were an omnipresent culture of innovation within the organization, patronage and active involvement of senior management in such initiatives, a step by step logical approach as postulated earlier, an enthusiastic, highly skilled in-house team, a cooperative ERP vendor and availability of sufficient budget and infrastructure. We hope that more companies would be inspired by the success of RCL and would adopt a planned and phased approach as described in this research.

## Acknowledgment

The second author gratefully acknowledges financial support received from the Indian Institute of Management Calcutta in the form of Category II Research Grant (Work order number: 3557/RP: ATBOCOMSUDM) for conducing this research as well as support received from the Indian Institute of Management Calcutta Case Research Center (IIMCCRC) and Ramco Cements Limited.

## Appendix A

The underlying architecture used in the project is shown in Fig. A1. The analytics database accessed various data sources and the results were fed to a Report server. The results could be accessed by internal users like members of the senior management as well as by external users like the sales force using different devices after logging in on their VPN.

An important decision in the project was to provide for real-time data because some analytics applications required data without any time lag. In other words, when the data was captured by the ERP system, it needed to be immediately available for analytics. Consequently, traditional data warehousing techniques were found unsuitable. After meticulous research, the team zeroed into a feature called “Always On” in SQL Server 2012 from Microsoft for this connection between ERP Server and Analytics Server (both using SQL Server 2012). With this feature enabled and properly configured, there was no time lag between the source data and the reported data. Moreover, there was no impact on the performance of the ERP system. Fig. A2 shows the implemented technology wherein the “Always On” feature ensured that the analytics engine received continuous real-time feed of data from the ERP system without slowing it down. As a result the users of

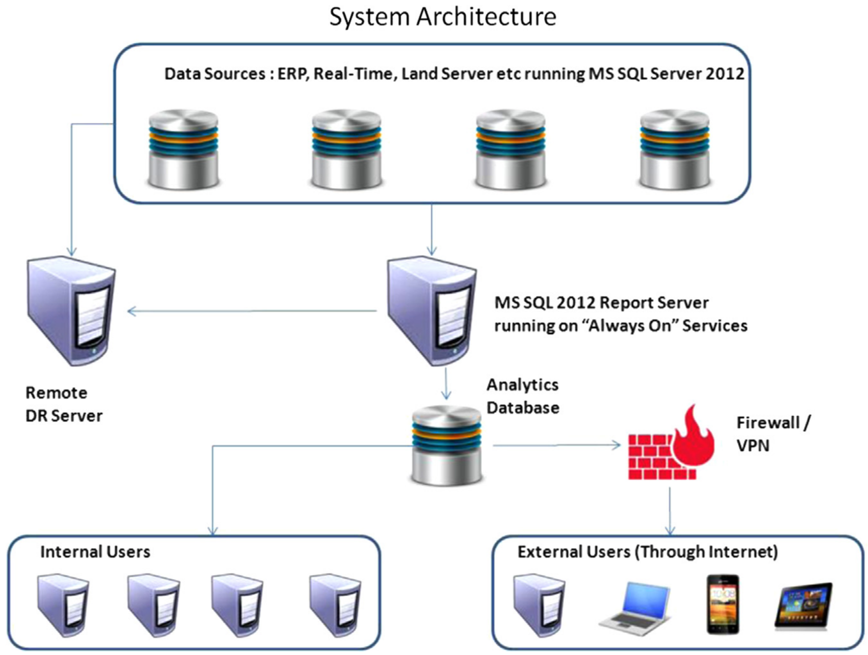


**Fig. A1.** Underlying architecture of the Big Data project of RCL. .

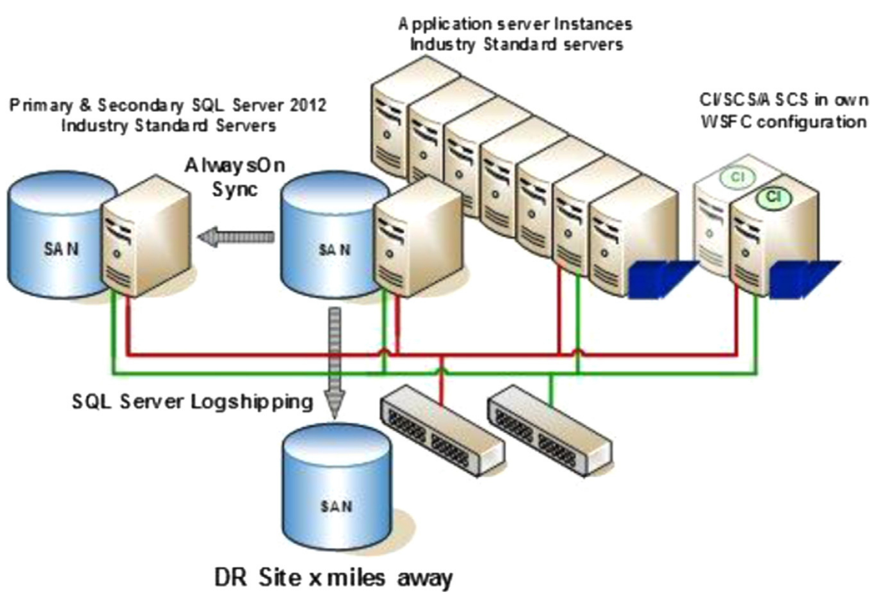


**Fig. A2.** Diagrammatic representation of real time connection between ERP system and analytics engine. .

the system did not have to depend on OLAP cubes of the traditional business intelligence system but could generate live reports at any time. This also allowed users to access the system across all mobile platforms within and outside the corporate local area network (e.g., at RCL godowns that used low bandwidth Internet connections). Furthermore, use of in-memory computation at the back end that resulted in completely rewriting the built-in stored procedures reduced the execution time of the queries significantly.

## References

Abrams, A.S., Jiao, J., Fan, W., Wang, G.A., Zhang, Z., 2013. What's buzzing in the blizzard of buzz? Automotive component isolation in social media postings. Decis. Support Syst. 55 (4), 871–882.

Azevedo, A., Santos, M.F., 2008. KDD, SEMMA and CRISP-DM: a parallel overview. In: Proceedings of the IADIS European Conference on Data Mining, Amsterdam, The Netherlands, pp. 182–185.

Baars, H., Kemper, H.G., 2008. Management support with structured and unstructured data—an integrated business intelligence framework. Inf. Syst. Manag. 25 (2), 132–148.

Benbasat, I., Goldstein, D.K., Mead, M., 1987. The case study research strategy in studies of information systems. MIS Q. 11 (3), 369–385.

Biesdorf, S., Court, D., Willmott, P., 2013. Big data: what's your plan? (Available online). McKinsey Q. (accessed 10.01.14).

Big Data Insight Group, 2012. What businesses can learn from big data and high performance analytics in the manufacturing industry (Available online). Appl. Insight Ser., 1–22.

Bizer, C., Boncz, P., Brodie, M.L., Erling, O., 2011. The meaningful use of big data: four perspectives – four challenges. SIGMOD Record 40 (4), 56–60.

Bose, I., Pal, R., Ye, A., 2008. ERP and SCM systems integration: the case of a valve manufacturer in China. Inf. Manag. 45 (4), 233–241.

Chica, M., Cordóna, Ó., Damas, S., Bautista, J., 2013. A robustness information and visualization model for time and space assembly line balancing under uncertain demand. Int. J. Prod. Econ. 145 (2), 761–772.

Coutinho, F.C., Lang, A., Mitschang, B., 2013. Making social media analysis more efficient through taxonomy supported concept suggestion. Available online: 〈http://www.btw-2013.de/proceedings/Making%20Social%20Media%20Analysis%20more%20efficient%20through%20Taxonomy%20Supported%20Concept%20Suggestion.pdf〉 (accessed 31.05.14.).

Davenport, T., 2006. Competing on analytics. Harv. Bus. Rev. 84 (1), 1–10.

Dubé, L., Paré, G., 2003. Rigor in information systems positivist case research: current practices, trends, and recommendations. MIS Q. 27 (4), 597–636.

Ebner, G., Bechtold, J., 2012. Are manufacturing companies ready to go digital? Capgemini Consulting White Paper. Available online: 〈http://www.capgemini-consulting.com/are-manufacturing-companies-ready-to-go-digital〉 (accessed 10.01.14.).

Ferguson, R.B., 2012. The big deal about a big data culture (and innovation). MIT Sloan Manag. Rev., 1–4.

Flyvbjerg, B., 2013. Quality control and due diligence in project management: getting decisions right by taking the outside view. Int. J. Proj. Manag. 31 (5), 760–774.

Harding, J.A., Shahbaz, M., Srinivas, Kusiak, A., 2006. Data mining in manufacturing: a review. J. Manuf. Sci. Eng. 128, 969–976.

Henschen, D., 2014. Merck optimizes manufacturing with big data analytics. Information Week. Available online: 〈http://www.informationweek.com/strategic-cio/executive-insights-and-innovation/merck-optimizes-manufacturing-with-big-data-analytics/d/d-id/1127901〉 (accessed 31.05.14.).

Hsu, S., Chien, C., 2007. Hybrid data mining approach for pattern extraction from wafer bin map to improve yield in semiconductor manufacturing. Int. J. Prod. Econ. 107 (1), 88–103.

Jacobs, A., 2009. The pathologies of big data. Commun. ACM 52 (8), 36–44.

Kang, D., Park, Y., 2014. Review-based measurement of customer satisfaction in mobile service: sentiment analysis and VIKOR approach. Expert Syst. Appl. 41 (4), 1041–1050.

Kiron, D., 2013. Organizational alignment is key to big data success. MIT Sloan Manag. Rev. 54 (3), 1–6.

Kiron, D., Ferguson, R.B., Prentice, P.K., 2013. From value to vision: reimagining the possible with data analytics. MIT Sloan Manag. Rev., 1–19.

Kiron, D., Prentice, P.K., Ferguson, R.B., 2012. Innovating with analytics. MIT Sloan Manag. Rev. 54 (1), 47–52.

Kiron, D., Shockley, R., 2011. Creating business value with analytics. MIT Sloan Manag. Rev. 53 (1), 57–63.

Kohavi, R., Rothleder, N.J., Simoudis, E., 2002. Emerging trends in business analytics. Commun. ACM 45 (8), 45–48.

Kuruppuarachchi, P.R., Mandal, P., Smith, R., 2002. IT project implementation strategies for effective changes: a critical review. Logist. Inf. Manag. 15 (2), 126–137.

LaValle, S., Lesser, E., Shockley, R., Hopkins, M.S., Kruschwitz, N., 2011. Big data, analytics and the path from insights to value. MIT Sloan Manag. Rev. 52 (2), 21–31.

Lee, A.S., 1989. A scientific methodology for MIS case studies. MIS Q. 13 (1), 33–52.

Lee, J., Lapira, E., Bagheri, B., Kao, H., 2013. Recent advances and trends in predictive manufacturing systems in big data environment. Manuf. Lett. 1 (1), 38–41.

MacKerron, G., Kumar, M., Kumar, V., Esain, A., 2014. Supplier replenishment policy using e-Kanban: a framework for successful implementation. Prod. Plan. Control 25 (2), 161–175.

Madden, S., 2012. From databases to big data. IEEE Internet Comput. 16 (3), 4–6.

Mandal, P., Gunasekaran, A., 2003. Issues in implementing ERP: a case study. Eur. J. Oper. Res. 146, 274–283.

Marchand, D.A., Peppard, J., 2013. Why IT fumbles analytics. Harv. Bus. Rev. 91 (1-2), 1–9.

Miele, S., and Shockley, R., 2013. Analytics: the real-world use of big data. IBM Institute for Business Value. Available online: 〈http://public.dhe.ibm.com/common/ssi/ecm/en/gbe03550usen/GBE03550USEN.PDF〉 (accessed 10.01.14.).

Mekawie, S., Elragal, A., 2013. ERP and SCM integration: the impact on measuring business performance. Int. J. Enterp. Inf. Syst. 9 (2), 106–124.

Obenshain, M.K., 2004. Application of data mining techniques to healthcare data. Infect. Control Hosp. Epidemiol. 25 (8), 690–695.

O'Leary, D.E., 2013. Artificial intelligence and big data. IEEE Intell. Syst. 28 (2), 96–99.

Oxford Economics, 2011. The new digital economy: how it will transform business. Available online: 〈http://www.google.co.in/url?sa=t&rct=j&q=&esrc=s&source=web&cd=3&ved=0CD0QFjAC&url=http%3 A%2 F%2Fwww.businessandleadership.com%2Fdownload%2Ffs%2Fdoc%2Freports%2Fthe-new-digital-economy-1.pdf&ei=DeHUUpuWM66QiQfjs4HIBA&usg=AFQjCNGQKXaQccqoEy8Na0ctStXgaJPH1Q&bvm=bv.59378465,d.aGc〉 (accessed 10.01.14.).

Rao, R., 2003. From unstructured data to actionable intelligence. IEEE IT Prof. 5 (6), 29–35.

Records, R.L., Fisher, Q.K., 2014. Manufacturers connect the dots with big data and analytics (Available online). Comput. Sci. Corp., 1–6 (accessed 31.05.14).

Schikora, P.F., Godfrey, M.R., 2003. Efficacy of end-user neural network and data mining software for predicting complex system performance. Int. J. Prod. Econ. 84 (3), 231–253.

Sukumaran, S., Sureka, A., 2006. Integrating structured and unstructured data using text tagging and annotation. Bus. Intell. J. 11 (2), 8.

The Economist Intelligence Unit, 2013. The hype and the hope: the road to big data adoption in Asia-Pacific. Available online: 〈http://www.hds.com/assets/pdf/the-hype-and-the-hope-summary.pdf〉 (accessed 10.01.14.).

Ting, S.L., Tse, Y.K., Ho, G.T.S., Chung, S.H., Pang, G., 2014. Mining logistics data to assure the quality in a sustainable food supply chain: a case in the red wine industry. Int. J. Prod. Econ. 152, 200–209.

Viaene, S., Bunder, A.V.D., 2001. The secrets to managing business analytics projects. MIT Sloan Manag. Rev. 53 (1), 65–69.

Waller, M.A., Fawcett, S.E., 2013. Data science, predictive analytics, and big data: a revolution that will transform supply chain design and management. J. Bus. Logist. 34 (2), 77–84.

Watson, H.J., Wixom, B.H., 2007. The current state of business intelligence. IEEE Comput. 40 (9), 96–99.

Wirth, R., and Hipp, J., 2011. CRISP-DM: towards a standard process model for data mining. Available online: 〈ftp://public.dhe.ibm.com/software/analytics/spss/documentation/modeler/14.2/en/CRISP_DM.pdf〉 (accessed 10.01.14.).

Yin, R.K., 2003. Case Study Research Design and Methods, third edition. Sage Publications.