International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 05 Issue: 04 | Apr-2018 www.irjet.net p-ISSN: 2395-0072

BIG DATA ANALYTICS: THE ENTERPRISE RESOURCE PLANNING

Mr. DHANANJAYA KUMAR K

Assistant Professor of Computer Science, Yuvaraja's College, University of Mysore, Karnataka, India.

Abstract: *ERP system is the source of planning of Enterprise Resource Organization and it is an integrated application software solution offered by a vendor to support the seamless integration of information or data flows through an organization. It is provided as a package comprising different modules, such as product management, quick billing details, finance or accounting, human resources management, supply product chain and customer information. ERP system process the implementation is mainly the lengthy process and completely more complex resulting in many cases of unsuccessful implementation which have negatively impacted on the performance of an organization's business and up to 90% of implementations did not achieve all the desired benefits. A majority of ERP problems are discovered in the last stage of the ERP life cycle known as the post-implementation phase or the after go-live phase. Much research has been undertaken in relation to the critical success factors of ERP implementation in developed countries whereas research on problems encountered in the ERP post-implementation phase are very limited in developed across the countries. So overcome all the problem we are find the solution which is MapReduce technology in hadoop system.*

Key Sets: ERP, Big Data,Hadoop,MapReduce,HDFS.

1. Introduction

Order processing is a key element of order fulfillment. Order processing systems or facilities are commonly called "distribution centers". Order processing systems, in one form or another, have been a part of doing business for ages, and have developed alongside technology to provide powerful means of capturing, tracking and shipping orders. Order processing is a sequential process involving: Picking, sorting, per-consolidation. Orders can be received from businesses, consumers, or a mix of both, depending on the products. Offers and pricing may be done via catalogs, websites, or broadcast network advertisements. Order processing is the process or work-flow associated with the selling, packing and delivery of the packed items to a shipping carrier

An integrated order management system may encompass these modules: Product information, Inventory available, vendors details and receiving, Marketing, Order entry and customer service, financial processing, order processing. Purchasing system has resulted in a very simple and easily understood 'menu based' approach where the operator can easily see what the choices and options are without having to remember a great deal. This Purchase Order Processing system is used to track purchasing activity in a particular shop or retail organization.

2. Objectives

The scope of the Online Order Processing is to take control of purchasing by improving accuracy, and increase employee productivity and implementing data flow of online purchase order processing for your remote staff using no more than standard browser software. Also retailer can enter and manage purchase orders; track shipments and invoices received. The objective of inventory management is to provide uninterrupted production and sales, at the minimum cost.

3. Retail Management System

The retail management system is business process management software that allows an organization to use a system of integrated applications to manage the business and automate many back office functions related to software and technology, services and human resources. Retail management software typically integrates all facets of an operation including product planning, development, manufacturing, sales and marketing in a single database system, application and interface between user.

Fig1: The Enterprise Resource Planning Model

4. Functional Areas of Retail Management System

An RMS system covers the following functional areas. In many RMS systems these are called and grouped together as *ERP or RMS modules*:

- Finance & Accounting: update General Ledger, Fixed Assets, payables including vouchering, matching and payment, receivables Cash Management and collections, cash management, Financial Consolidation
- Management Accounting: Budgeting, Costing, cost management, activity based costing
- Human resources: Recruiting, training, rostering, payroll, benefits, retirement and pension plans, diversity management, retirement, separation
- Manufacturing: Engineering, bill of materials, work orders, scheduling, capacity, workflow management, quality control, manufacturing process, manufacturing projects, manufacturing flow, product life cycle management
- Order Processing: Order to cash, order entry, credit checking, pricing, available to promise, inventory, shipping, sales analysis and reporting, sales commissioning.
- Supply chain management: Supply chain planning, supplier scheduling, product configurator, order to cash, purchasing, inventory, claim processing, warehousing (receiving, putaway, picking and packing).
- Project management: Project planning, resource planning, project costing, work breakdown structure, billing, time and expense, performance units, activity management
- Customer relationship management: Sales and marketing, commissions, service, customer contact, call center support— CRM systems are not always considered part of ERP systems but rather Business Support systems.
- Data services: Various "self–service" interfaces for customers, suppliers and/or employees

Fig2: The Functional Area of Enterprise Resource Planning

5. What is Big Data?

Big data is a term that describes the large volume of data – both structured and unstructured – that inundates a business on a day-to-day basis. But it's not the amount of data that's important. It's what organizations do with the data that matters. Big data can be analyzed for insights that lead to better decisions and strategic business moves.

While the term "big data" is relatively new, the act of gathering and storing large amounts of information for eventual analysis is ages old. The concept gained momentum in the early 2000s when industry analyst Doug Laney articulated the now-mainstream definition of big data as the Five Vs:

Volume. Organizations collect data from a variety of sources, including business transactions, social media and information from sensor or machine-to-machine data. In the past, storing it would've been a problem – but new technologies (such as Hadoop) have eased the burden.

Velocity. Data streams in at an unprecedented speed and must be dealt with in a timely manner. RFID tags, sensors and smart metering are driving the need to deal with torrents of data in near-real time.

Variety. Data comes in all types of formats – from structured, numeric data in traditional databases to unstructured text documents, email, video, audio, stock ticker data and financial transactions.

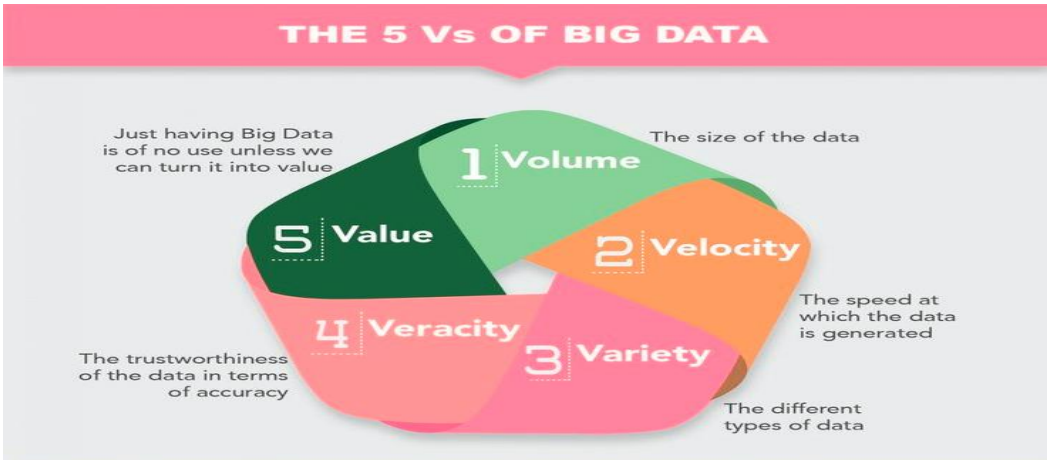


Fig 3: Big Data 5V's Model

Varicity. In addition to the increasing velocities and varieties of data, data flows can be highly inconsistent with periodic peaks. Is something trending in social media? Daily, seasonal and event-triggered peak data loads can be challenging to manage. Even more so with unstructured data.

Value or Complexity. Today's data comes from multiple sources, which makes it difficult to link, match, cleanse and transform data across systems. However, it's necessary to connect and correlate relationships, hierarchies and multiple data linkages or your data can quickly spiral out of control.

6. Why Is Big Data Important?

The importance of big data doesn't revolve around how much data you have, but what you do with it. You can take data from any source and analyze it to find answers that enable 1) cost reductions, 2) time reductions, 3) new product development and optimized offerings, and 4) smart decision making. When you combine big data with high-powered analytics, you can accomplish business-related tasks such as:

- Determining root causes of failures, issues and defects in near-real time.
- Generating coupons at the point of sale based on the customer's buying habits.
- Recalculating entire risk portfolios in minutes.
- Detecting fraudulent behavior before it affects your organization.

7. What is MapReduce?

MapReduce™ is the heart of Apache hadoop It is this programming paradigm that allows for massive scalability across hundreds or thousands of servers in a Hadoop cluster. The MapReduce concept is fairly simple to understand for those who are familiar with clustered scale-out data processing solutions. For people new to this topic, it can be somewhat difficult to grasp, because it's not typically something people have been exposed to previously. If you're new to Hadoop's MapReduce jobs, don't worry: we're going to describe it in a way that gets you up to speed quickly. The term MapReduce actually refers to two separate and distinct tasks that Hadoop programs perform. The first is the map job, which takes a set of data and converts it into another set of data, where individual elements are broken down into tuples (key/value pairs). The reduce job takes the output from a map as input and combines those data tuples into a smaller set of tuples. As the sequence of the name MapReduce implies, the reduce job is always performed after the map job.

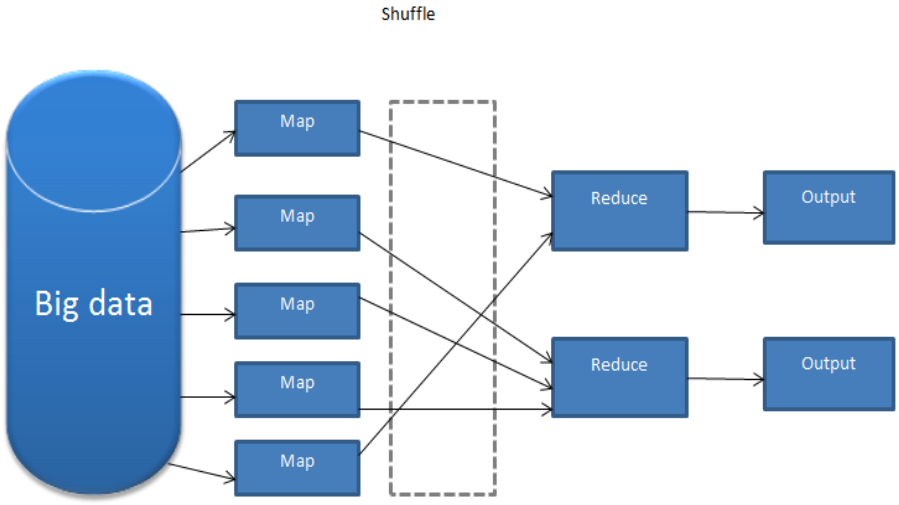


Fig 4: MapReduce Work Flow Architecture

8. What is HDFS?

Apache Hadoop Distributed File System also called as The Hadoop Distributed File System (HDFS) is the primary data storage system used by hadoop applications. It employs a NameNode and DataNode architecture to implement a distributed file system that provides high-performance access to data across highly scalable Hadoop clusters.

9. How HDFS works

HDFS supports the rapid transfer of data between compute nodes. At its outset, it was closely coupled with MapReduce, a programmatic framework for data processing. When HDFS takes in data, it breaks the information down into separate blocks and distributes them to different nodes in a cluster, thus enabling highly efficient parallel processing. Moreover, the Hadoop Distributed File System is specially designed to be highly fault tolerant. The file system replicates, or copies, each piece of data multiple times and distributes the copies to individual nodes, placing at least one copy on a different server rack than the others. As a result, the data on nodes that crash can be found elsewhere within a cluster. This ensures that processing can continue while data is recovered. HDFS uses master/slave architecture. In its initial incarnation, each hadoop cluster consisted of a single Name Node that managed file system operations and supporting Data Nodes that managed data storage on individual compute nodes. The HDFS elements combine to support applications with large data sets. This master node "data chunking" architecture takes as its design guides elements from Google File System (GFS), a proprietary file system outlined in in Google technical papers, as well as IBM's General Parallel File System (GPFS), a format that boosts I/O by striping blocks of data over multiple disks, writing blocks in parallel. While HDFS is not Portable Operating System Interface model-compliant, it echoes POSIX design style in some aspects.

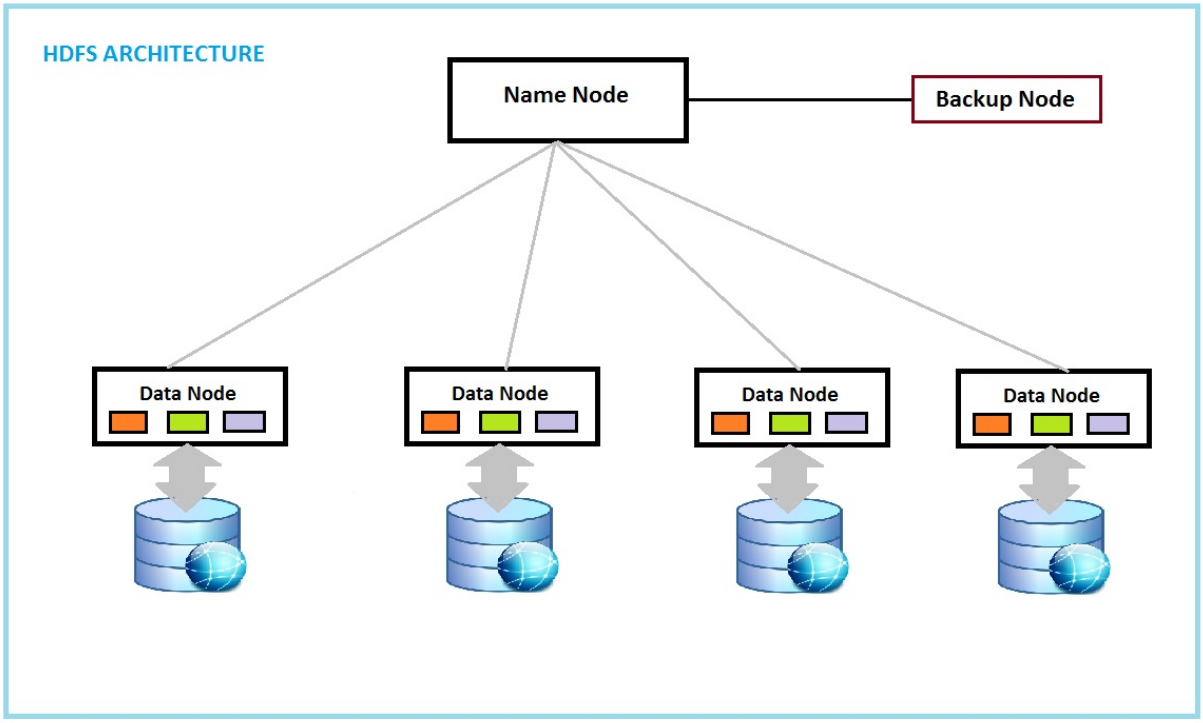


Fig 5: HDFS Architecture

10. System Architecture of RMS System

The system architecture is the process of how the system works overall design which can executed and it is a conceptual design Architecture is a set of structuring principles that enables a system to be comprised of a set of simpler systems each with its own local context that is independent of but not inconsistent with the context of the larger system as a whole. Both definitions focus on system structure. You create an architecture to describe the structure of the system to be built and how that structure

supports the business and service-level requirements. You can define the structure of a system as the mechanisms that the system employs to solve the common problems of the system. A mechanism is a capability that supports the business requirements in a consistent and uniform manner. For example, persistence is a mechanism that should be used consistently throughout the system. This means that any time the system uses persistence, it is handled in the same manner. By defining persistence as an architectural mechanism, you provide a default method of addressing persistence that all designers should follow and implement consistently. The architectural mechanisms, such as persistence, distribution, communication, transaction management, and security are the infrastructure on which you build the system and must be defined in your architecture.

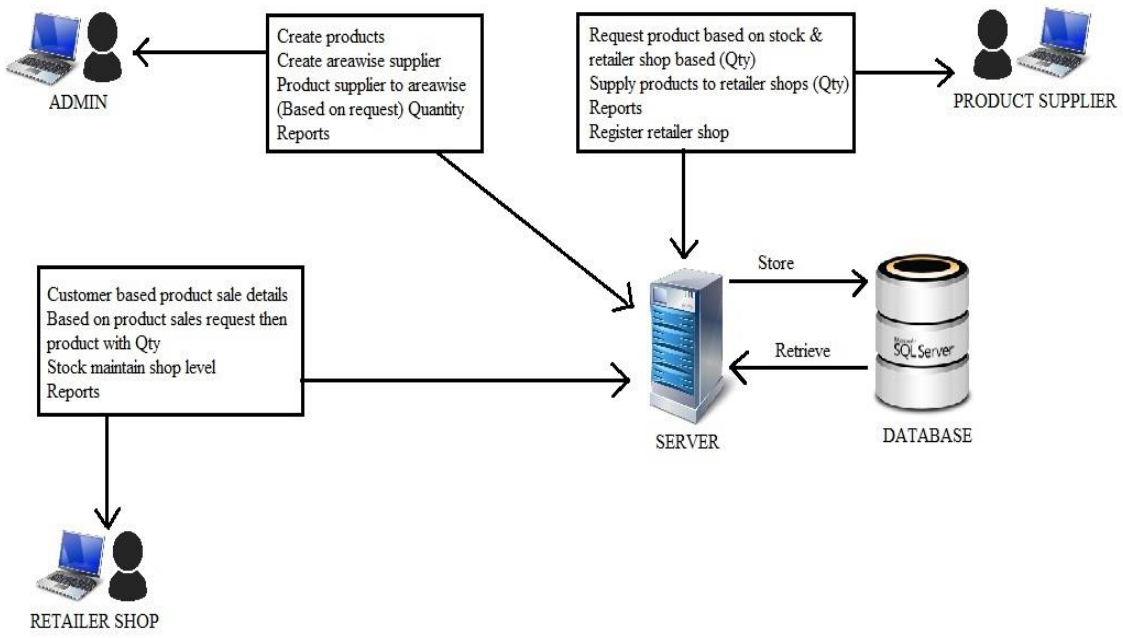
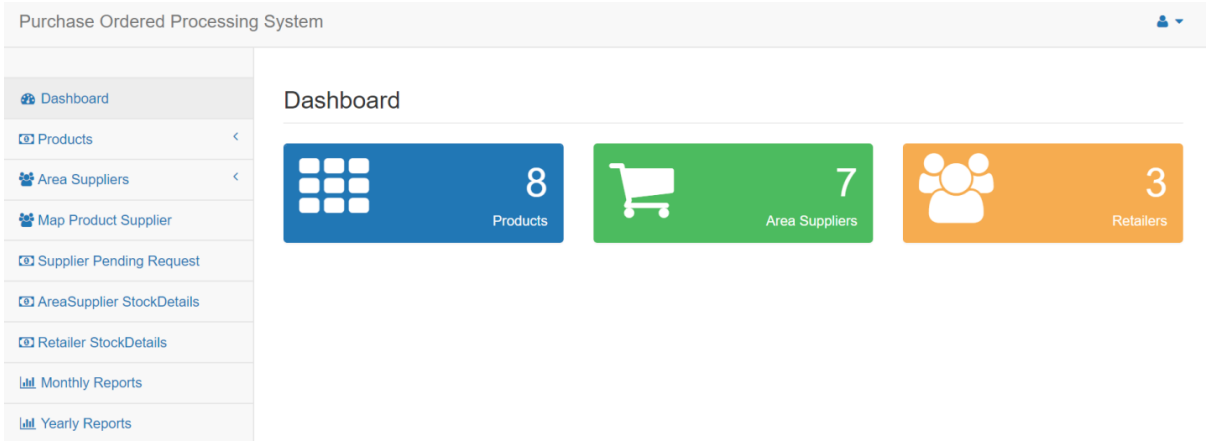
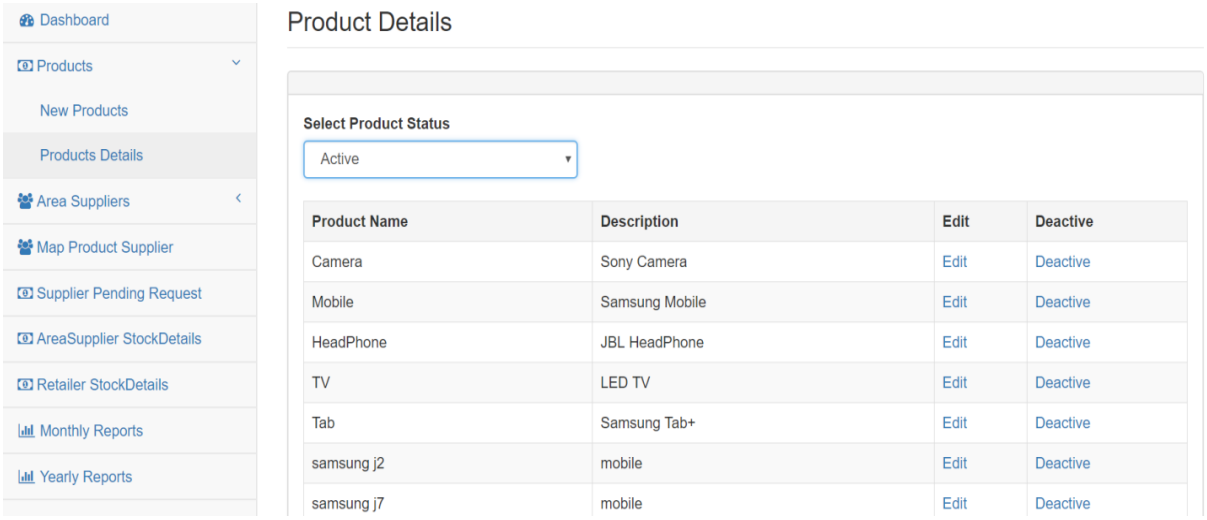


Fig 6: The System Architecture of RMS

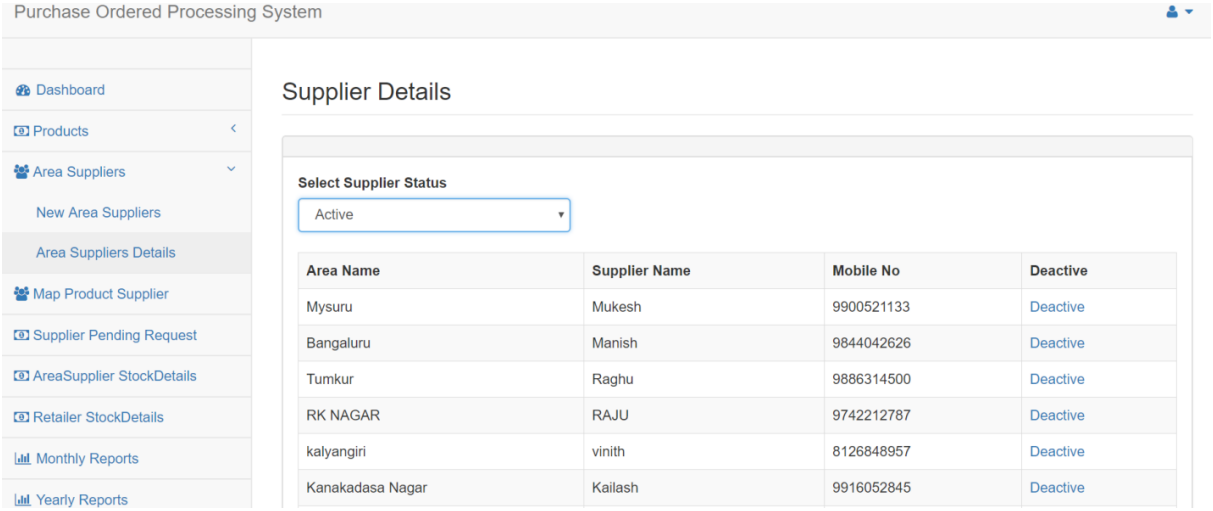
11. Result



Result 1: Product Dashboard



Result 2: Product Details



Result 3: Map Area Supplier

12. Conclusion

This paper helped out day by day the product purchased in online this is aimed to help the automating the purchase order process is more efficient than manual processes since the tasks can be expedited using digital technology, managed more efficiently, and tracked more effectively The Purchase Order Processing System consisting of Indent module, Approval module, Quotation modules, Goods receive or not module and it will analyze easily in big data system.

13. References

Hasselbring, W. (2000). Information system integration. Communications of the ACM,43(6), 32-38.Holland, C. R., & Light, B. (1999). A critical success factors model for ERP implementation. Software, IEEE, 16(3), 30-36.

Abdinnour -Helm, S., Lengnick-Hall, M. L. and Lengnick-Hall, C. A., Pre-implementation attitudes and organizational readiness for implementing an Enterprise Resource Planning system. European Journal of Operational Research2003, 146(2), 258-273.

Adler, P., Interdepartmental interdependence and integration: The caseof the design-marketing interface. Organization Science1995, 6(2),147-167.

Aiken, M. and Hage, J., Organizational interdependence and intra organizational structure. American Sociological Review1968, 33, 912-930.

Akkermans, H. and van Helden, K., Vicious and virtuous cycles in ERP implementation: a case study of interrelations between critical success factors. European Journal of Information Systems 2002, 11, 35-46.

Al-Mashari, M., Al-Mudimigh, A. and Zairi, M., Enterprise resource planning: A taxonomy of critical factors. European Journal of Operational Research 2003, 146(2), 352-364.

Anderson, M. C., Banker, R. D. and Ravindran, S., The new productivity paradox. Communications of the ACM 2003, 46(3), 91-94.

Anderson, J. C., Rungtusanatham, M., Schroeder, R. O. and Devaraj, S., 1995. A Path Analytic Model of a Theory of Quality Management Underlying the Deming Management Method: Preliminary Empirical Findings. Decision Sciences 26(5), 637-658.

Ash, C. G. and Burn, J. M., A strategic framework for the management of ERP enabled e-business change. European Journal of Operational Research2003, 146(2), 374-387.

Bagozzi, R. P., Causal Models in Marketing, 1980(John Wiley and Sons, New York). Barua, A., Kriebel, C. H. and Mukhopadhyay, T., Informationtechnologies and business value: An analytic and empirical investigation. Information Systems Research 1995, 6(1), 3-23.

Elgendy, N.: Big Data Analytics in Support of the Decision Making Process. MSc Thesis, German University in Cairo, p. 164 (2013)

EMC: Data Science and Big Data Analytics. In: EMC Education Services, pp. 1–508 (2012)

He, Y., Lee, R., Huai, Y., Shao, Z., Jain, N., Zhang, X., Xu, Z.: RCFile: A Fast and Space-efficient Data Placement Structure in MapReduce-based Warehouse Systems. In: IEEE International Conference on Data Engineering (ICDE), pp. 1199–1208 (2011)

Herodotou, H., Lim, H., Luo, G., Borisov, N., Dong, L., Cetin, F.B., Babu, S.: Starfish: A Self-tuning System for Big Data Analytics. In: Proceedings of the Conference on Innova-tive Data Systems Research, pp. 261–272 (2011)

Big Data Analytics: A Literature Review Paper (PDF Download Available). Available from: https://www.researchgate.net/publication/264555968_Big_Data_Analytics_A_Literature_Review_Paper [accessed Apr 13 2018].

Big Data Analytics: Available from: http://www.ibm.in.

Enterprise Resource planning: Available from: http://www.google.in/ERP.

Hadoop MapReduce: Available from: http://www.google.in/Apache-Hadoop.

BIOGRAPHIES

Mr. Dhananjaya Kumar K is an Assistant Professor, Department of Computer Science at yuvaraja's college, Karnataka, India. He hold a Master Degree at Computer Science and Engineering from the university of VTU, He has experience on IOT, Hadoop framework and Map-Reduce processes and developed a Big Data and analytics for the needs of his Master Thesis. He has published research paper on Advanced Algorithm and Emergency Management system in elevated toll roads at journals and international conferences and recently published research paper on big data in department of technical education on digital teaching in class room.