World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:8, No:9, 2014

Comparative Analysis of Diverse Collection of Big Data Analytics Tools

S. Vidhya, S. Sarumathi, N. Shanthi

***Abstract*—**Over the past era, there have been a lot of efforts and studies are carried out in growing proficient tools for performing various tasks in big data. Recently big data have gotten a lot of publicity for their good reasons. Due to the large and complex collection of datasets it is difficult to process on traditional data processing applications. This concern turns to be further mandatory for producing various tools in big data. Moreover, the main aim of big data analytics is to utilize the advanced analytic techniques besides very huge, different datasets which contain diverse sizes from terabytes to zettabytes and diverse types such as structured or unstructured and batch or streaming. Big data is useful for data sets where their size or type is away from the capability of traditional relational databases for capturing, managing and processing the data with low-latency. Thus the out coming challenges tend to the occurrence of powerful big data tools. In this survey, a various collection of big data tools are illustrated and also compared with the salient features.

***Keywords*—**Big data, Big data analytics, Business analytics, Data analysis, Data visualization, Data discovery.

I. Introduction

BIG data is avast quantity of data which extracts values by the process of capturing and analysis, this can be possible by innovative architectures and technologies. Nowadays from the platform of traffic management and tracking personal devices such as Mobile phones are useful for position specific data which emerges as novel bases of big data. Mainly the Big data have developed to increase the use of data demanding technologies for it. By using prevailing traditional techniques it is very challenging to achieve effective analysis of the huge size of data. Meanwhile, on the market, big data have become the latest imminent technology, which can serve vast profits to the business organizations. This becomes essential because it contains several issues and challenges related in bringing and adapting, which need to be understood in this technology. The concept of big data deals with the datasets which continues to develop rapidly whereas that becomes tough to handle them by using the current concepts and tools in database management. Data capture, sharing, analytics, search, storage, visualization, etc., is the related difficulties in big data. Many challenges can be forwarded due to the several properties of big data like variety, velocity, variability, volume, value and complexity. Scalability, real-time analytics, unstructured data, fault tolerance, etc., is the several challenges included in huge data management. Obviously the amount of data stored in various sectors can vary in the data stored and created, i.e., images, audio, text information etc., from one industry to another. From the practical perspective, the graphical interface used in the big data analytics tools leads to be more efficient, faster and better decisions which are massively preferred by analysts, business users and researchers [1].

Ms.S.Vidhya, PG Scholar, is with the Department of Information Technology, K. S. Rangasamy College of Technology, Tamil Nadu, India (phone: 9443960666; e-mail: vidhyapsubramani@gmail.com).

Mrs.S.Sarumathi, Associate Professor, is with the Department of Information Technology, K. S. Rangasamy College of Technology, Tamil Nadu, India (phone: 9443321692; e-mail: rishi_saru20@rediffmail.com).

Dr.N.Shanthi, Professor and Dean, is with the Department of Computer Science Engineering, Nandha Engineering College, Tamil Nadu, India (e-mail: shanthimoorthi@yahoo.com).

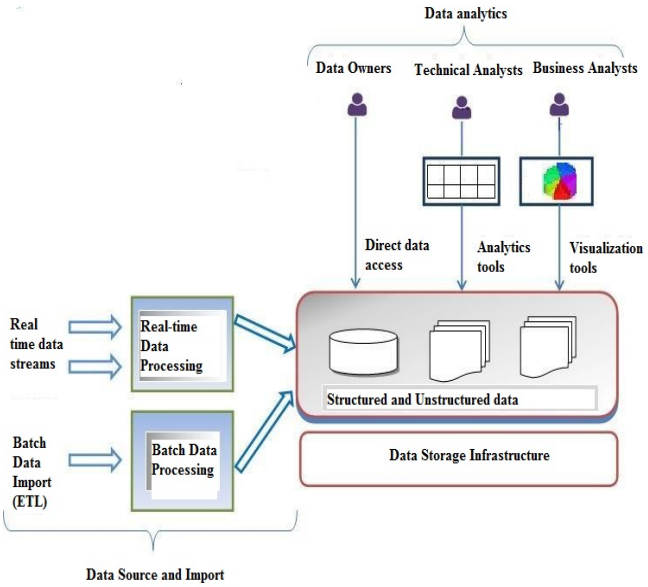

Fig. 1 Big data Architecture

II. Different Big Data Tools

A. Pentaho Business Analytics

Pentaho was actually made as a report generating engine, which later took its form as a software program. Similar to JasperSoft, Pentaho gathers information from new sources by branching into big data. It can be integrated with most of the popular NoSQL databases like MongoDB and Cassandra. Once the connection between the database and Pentaho is established, we can select the information from a database and then make them as reports and views by dragging and dropping the columns. The typical lesson would be the classic sorting and sifting tables to recover out the users who are using a website for a long time. A simple sort by IP address in the log files will show what the heavy users were actually doing.

It also facilitates software for drawing HBase data from

Hadoop clusters and HDFS file data. Pentaho Data Integration, also called as Kettle, is one of the interesting graphical programming interface tools and it has a number of built-in modules for the user to drag, drop and connect them. Pentaho has complete integration with Hadoop and other sources, thus enabling the user to write the code and send it directly to the cluster for execution [2].

Pentaho [3] fetches IT and business users together with firmly coupling data integration with business analytics, thus permitting them to integrate, access, blind, visualize and analyze all data that makes an impact on the business results. Pentaho's open source legacy brings sustained modern innovation, integrated, Embeddable analytics platform, which takes any type of data together, including heavy and diverse data types.

The pluggable, open platform of Pentaho not only produces a flexible analytics resolution for supporting today's existing infrastructures but also adjusts to tomorrow's foreseeable changes. Pentaho Business Analytics offers a range of progressively more advanced analytics starting from reporting to data visualization and predictive analytics in a full platform. Innately open, Pentaho is quick to deploy, easy to use and extremely cost-effective, in a nutshell to be described as a platform built for the future prospects of analytics.

1) Pentaho Data Integration

An ever rising challenge of any organization would be to supervise and pull out value from increasing volumes and varieties of data. Pentaho Data Integration helps the organizations to access data from composite and heterogeneous sources and merges it with existing and diverse data types, thus producing high quality ready-to-analyze information. It has a rich GUI (graphical user interface) and a parallel processing engine that provides high performance ETL (extract, transform and load) capabilities in order to wrap all data integration needs, together with big data.

2) Data Discovery and Exploration

Pentaho Business Analytics provides an up to date, extremely interactive and perceptive web-based interface for business users to access and recover out all data, counting big data. With a variety of increasingly advanced analytics in hand, users are given the provision of creating reports and dashboards, visualizing and analyzing data across multiple dimensions, with no dependence on IT or developers. Pentaho can be installed on-premise, in the Cloud or impeccably implanted into other software applications.

3) Big Data

Pentaho Business Analytics for big data considerably reduces the plan time and complexity to acquire and deploy big data analytics, thus helping the companies to realize the business value of a large bit of diverse data. With the extraction of data and preparation of big and traditional data types in any infrastructure, along with the range of analytics from data discovery to predictive analytics, it is no wonder that Pentaho presents a complete solution for the whole big data life cycle.

4) Embedded Business Analytics

Pentaho Business Analytics has features like open APIs, multi-tenant ready architecture and 100% Java that builds it perfect for ISV and SaaS companies which are facing forward to embed their information-centric applications with the business analytics. The Powered by Pentaho program proposes a realistic approach to the partners by embedding visual and sophisticated analysis, reporting and dashboards the results into solution, in a very meager time span of eight weeks.

5) Predictive Analytics

Aside from data discovery, Pentaho also affords potent, state of- the-art machine learning algorithms and data processing instruments. The meaningful patterns and correlations which are hidden among the standard analysis and reporting can be unveiled by the data scientists and analysts. Sophisticated, highly developed analytics like time series forecasting, helps the user to plan for future outcomes.

6) Features

1. Easy to process high-volume data.
2. Provides flexibility and increased competitor to companies.
3. Performs data mining and predictive analysis.
4. Explore data to insights in three ways.

B. Jaspersoft

The Jaspersoft platform is an open source tool used for generating reports from database columns. In various businesses this software has been already installed, which help to form PDFS from SQL tables, it examines each person at meetings [2]. From JasperReports and iReport, the Jaspersoft Business Intelligence suite can be constructed. To offer a complete family of Business Intelligence (BI) products, the Java reporting library and graphical report designer can be used which is considered as the world's most famous open source. This suite helps to extend the range of their BI requirements, including dash boarding, reporting based on their operational, production and interactive end-user query, data integration and analysis. Products can be obtained either by stand-alone or a part of an integrated suite which make use of the ordinary metadata and shared facilities like security, scheduling, etc. Jaspersoft is constructed on and exposures the complete public API's for allowing customization, extensibility and seamless integration with some other applications for business and developers [4].

The company is skipping to the big data [2], by adding a software layer to attach their report producing software to their corresponding locations in which the big data has been kept. The JasperReport server provides software to utilize the data from various storage platforms which includes MongoDB, Riak, Redis, Cassandra, Neo4j and CouchDB. To enter inside HBase the Hive connector is provided by Jasper Reports which is well-represented by Hadoop. The work seems to be

still in the preliminary state in which several pages in the documentation wiki are blank and not completely integrated tools.

From sources we get the data to make an interactive graphs and tables from the Jaspersoft's Server. The reports may be used as refined interactive tools which help to give an answer from several angles. The refined reports can be used by expanding the Jaspersoft, which has the newer source of data. Jaspersoft provides a more refined method to access the data which is kept in the new location and it doesn't specify any particular new method to view the data.

1) JasperReports - Reporting Library

The world's most popular used open source software is a Java report library [4] which is designed mainly for developers and power-users. The sophisticated pixel-perfect print-ready report rendering or Web can be delivered from any Java application that has been embedded with them. The JasperServer can be imported easily from JasperReports, which is a report designed with a more sophisticated adhoc query, dash boarding requirements, etc., for users. JasperReports consist of relational databases with JDBC wrapped data providers, XML data sources, JavaBeans and Plain Old Java Objects. For supporting legacy or homegrown data sources, the custom data source providers can be easily added or it is used for implementing custom data source logic. The JasperReports consist of fully internationalized calculations, variables and the built-in charting.

2) JasperAnalysis - Data Analysis

The "Slice and Dice" data analysis is offered by Jasper Analysis for business users. In data it can be used to discover patterns, correlations, anomalies and trends to dynamically filter, pivot, chart data, drill in real-time by users. By user and role, the data can be safeguarded by a cell which helps for business users to access only the authorized information.

3) JasperServer – Interactive Report Server

JasperServer is designed mainly for developers and business user which is built on JasperReports. It offers an interactive report server, which helps for delivery of information on the basis of real-time or scheduled. Without burdening any IT or application developers, the contemporary Web 2.0 technologies like DHTML, Ajax, etc., JasperServer provides facilities for users to meet the needs of their own Business Intelligence, through the support for spontaneous reporting, dash boarding and adhoc query. For embedding OEM business intelligence capabilities, JasperServer is considered to be perfect for independent software vendors.

4) JasperETL - Data Integration

The data integration capabilities are offered by JasperETL. By combining data from various sources into a single consistent store, the JasperETL is mainly used which helps in reporting and analysis. Compared to traditional manual coding of data loading and transformation processes, the JasperETL offers their enhancement in scalability, productivity, etc. The purpose of JasperETL is to transfer the data from various normalized transactional data stores, which is done by data improvement /cleaning process and into data marts, operational data stores.

5) Features

1. Capable of real-time analytics.
2. Incorporate all kinds of data.
3. Combine data through data virtualization metadata layer or traditional data warehouse using ETL.
4. Offers visualizations and dashboards within an application.
5. Make design tools for non-designers create visualizations.

C. Splunk

Splunk is a potent platform developed for the analysis of Machine data. It can be defined as the data emitted by machines in great volumes, was rarely used efficiently. But nowadays it is becoming progressively more important in the worlds of technology and business [5].

Splunk is unusual from others, the difference being that it generates index for the data similar to the index generation for a text. It's not precisely the AI routine collection or a report-generating tool, even though it achieves much of that along the way. This indexing is astonishingly flexible, thus making Splunk as a tunable platform for an application and hence making sense out of the log files and sucking them up. Splunk is put on the market with diverse solution packages like Microsoft Exchange server monitoring and Web attack detection. The index is very useful to associate the data in many universal server-side scenarios.

Splunk is designed to obtain text strings and search roughly in the indicator. For example, Splunk get the URLs or IP address in a document and packages them together into a timeline which is built approximately around the time when it detects the data. All other associated fields are used to dig down to the data set. Although it is a simple process, it is very potent if the user is searching for the right type of needle in the data feed. If the user can trace out the right text string, Splunk is very much useful for tracking it. Log files are a vast application for it.

At present, a new Splunk tool named Shep, is used in private beta to provide a bidirectional integration among Hadoop and Splunk, thus permitting the user for data exchange between systems and curing the Splunk data from the Hadoop [2].

1) Making Machine-Generated Data Accessible, Usable and Valuable to Everyone

Splunk Enterprise stands in the leading role in the collection, analysis and visualization of machine data. It affords an integrated way of organizing and extracting real-time insights from huge amounts of machine data, provided from virtually any source. This consists of data from app servers, sensors, websites, social media platforms, business applications, traditional databases, hypervisors and open source data stores. Once we are able to have the data in

Splunk, we can search, monitor, report and analyze it, without considering how unstructured, huge or diverse the data may be. Splunk software furnishes you real-time perception of what took place earlier and what is passing now with new heights of visibility and penetration. This is known as operational intelligence.

a) Powerful Connectivity

Many organizations will have relational data, machine data and other unstructured data in large varieties. Splunk DB Connect is used for real-time connection between one or many relational databases and Splunk Hadoop Connect is used for bi-directional connectivity with Hadoop. Both of these Splunk applications will enable the user to drive more significant insights from all of the data.

b) Enterprise – Scale Big Data

Across several geographies, several data centers and several cloud infrastructures, Splunk software collect and does the indexing for tens of terabytes of data per day. Because the insights from your data are mission-critical, Splunk offers the resilience that is much need of the user for mission critical data with a low-cost, dispersed computing environment.

c) Real-Time Monitoring of the Entire Hadoop Stack

In the end-to-end Hadoop environment, Splunk App for HadoopOps encircles all layers of the supporting infrastructure and gives a real-time monitoring and analysis of the health and performance.

d) Robust Platform for Developing Big Data Applications

Developer panels will discover a complete host of ways to influence Splunk and exploit enterprise technology investments. Inbuilt SDKs for JSON and JavaScript with additional downloadable SDKs for Java, Python, PHP, C# and Ruby make it easy to modify and wielding the power of Splunk.

e) Proven Results

7,000 enterprise customers have seen the proven results of Splunk Enterprise and they use Splunk to enhance service levels, decrease operations costs, alleviate security risks, facilitate compliance, improve DevOps collaboration and build new product and service offerings. Splunk customers classically achieve an ROI (return on investment) measured in weeks or months, sometimes even well before the software is installed into production environment.

2) What Makes Splunk Unique

Splunk Enterprise provides end-to-end, integrated, real-time solution for machine data along with the below core capabilities:

- Widespread collection and indexing of machine data, from nearly any source
- Potent search processing language (SPL™) to explore and analyze real-time as well as historical data
- Provides real-time observation of patterns and thresholds and sends real-time alerts if specific conditions arise
- Great reporting and analysis
- Has custom dashboards and views for diverse roles
- Flexibility and degree of commodity hardware
- Security and access controls based on Granular role-model
- Provides assistance for multi-tenancy and stretchy, distributed deployments
- Connection to other data stores, including real-time integration with relational databases and bidirectional connectivity with Hadoop
- Strong, flexible platform for big data applications [6].

3) Features

1. Develop perceptions from Big Data with speed and simplicity.
2. Suitable for major Hadoop distributions, including first-generation MapReduce and YARN.
3. SplunkHadoop Connect supports bi-directional integration.
4. Splunk Enterprise stands in the leading role in the collection, analysis and visualization of machine data.

D. Tableau

The Tableau is an American Software Company with its headquarters located in Seattle. It manufactures a variety of interactive data visualization products built in business intelligence [7]. Tableau software implemented Hadoop years ago and it uses Hive for query structuring. Then it attempts with great effort to cache information in the memory in order to make the tool more interactive. Among the other tools that are developed to create reports offline, Tableau needs to give an interactive mechanism so that the user can dig the data as much as possible. Caching assists work with few latency of a Hadoop cluster. The well-polished software helps the user to feel more interested to work. The ability to reslice data across various graphs gives an artistic effect to the software [2]. The different products developed by Tableau are listed below:

1) Tableau Desktop (Business Analytics Anyone Can Use)

Tableau Desktop is one of the visualization tools that make the user to see the data in a very different number of ways with each slicing. The user has the freedom to even mix data and inspect it in a different perspective. The tool is optimized to provide all data columns and permits the user to shuffle them before embedding it into graphical templates [2].Tableau Desktop stands on breakthrough technology from Stanford University, which lets the user drag & drop for data analysis.

It is very simple to use that a user with Excel knowledge can easily learn it and can get extra results with meager effort. Also, it is 10 – 100x faster than any other existing solutions [7].

2) Tableau Server

Tableau Server is an application focused on business intelligence that offers a browser-based analytics. It is considered to be one of the best alternatives to the traditional business intelligence software that maintain slow pace. The

feature that makes Tableau as outstanding is that it is designed for everyone to use. It requires no scripting and hence gives the advantage of anyone becoming an analytics expert. The user can grow the deployment as per the requirement, he can conduct training online free of cost and fetch answers in few minutes.

3) Tableau Online

Tableau Online being a hosted version of Tableau Server makes business analytics much easier. The user can share dashboards with a very big network of clients and partners in a matter of minutes. It gives real-time, interactive data views that allow people to respond to their own questions either through a tablet or a web browser in a protected and hosted environment.

Tableau Online can extend as per the user requirements and it is easy to use wherever the user needs it. There is no need of purchasing, managing or setting up any infrastructure.

4) Tableau Public

Tableau Public is used to build incredible interactive visuals and thus it is very useful for anyone who desires to tell stories on the web. It is provided as a service that lets the user to be up and running the whole night. Apart from visuals creation, it also makes those visuals to be published quickly, with no help from the programmers or IT.

The Premium version is offered for organizations who desire to improve their websites with more interactive data visualizations. The other premium features include higher data size, ability to hide underlying data.

5) Why Should We Use Tableau?

Few reasons for choosing tableau are

a) It Is Very Easy to Use:

It requires no programming knowledge of any sort. All the user needs is some data, tableau for report generation and a creative mind. It helps the user to build enchanting visuals with its drag and drop feature and VizQL.

b) VizQL

VizQL is a visual query language that interprets drag-and-drop actions into data queries and then conveys that data in a visual manner. VizQL summarizes all the query complexities and gives remarkable gains to look and realize the data. The net result is an innate user experience that allows the people to answer at their thinking speed. VizQL is considered to be a representation of foundational advancement in analyzing and visualizing the data.

The basic invention is a patented query language which converts the user actions into a database query and gives the graphical output. Rapid analytics and visualization have become a reality due to VizQL. It is so easy that a user with no training can look and comprehend data faster and easier and that makes VizQL different from its fellow languages.

6) Advanced In-Memory Technology- The Data Engine

The biggest advantage of the tableau is its memory technology, which operates ever increasing heap amounts of data, in a very good speed among all other existing databases and data silos. Ability to run on ordinary computers and the capacity to leverage the entire memory hierarchy to L1 cache from disk are the major success factors of the tableau. Adhoc analysis in seconds and no fixed data model provides the power to everyone who uses tableau [7].

7) Features

1. Few clicks are used to access any data.
2. Additional data sources can be easily layered.
3. From disk to L1 cache it influences the complete memory hierarchy.
4. For mobile it has swipe, tab and touch functionality.
5. It removes memory silos from In-memory analytics database.

E. Karmasphere

Like many other big data tools that did not emerge as reported tools, Karmasphere Studio was originally developed as a set of plug-ins for Eclipse. The specialized IDE is apt for creating and running Hadoop jobs easily. While configuring a Hadoop job with Karmasphere Studio, you can notice that they show partial results along the way of configuration. The magnificent feature of Karmasphere Studio is that it shows the test data at each step while setting up the workflow, thus making the user to understand the outlook of the temporary data as it is being analyzed and reduced.

Karmasphere also has a tool called Karmasphere Analyst, which is developed to ease the procedure of plowing through all data in a Hadoop cluster. It has many features for creating a good Hadoop job, like subroutines which uncompressed Zipped logs files. Later it pulls them together and creates an output table for perusing by promoting the Hive calls [2].

For huge unstructured and structured data sets on Amazon Elastic MapReduce, Karmasphere gives a high productivity elucidation. The flexibility and scalability of Amazon Elastic MapReduce along with the user-friendly Karmasphere desktop tools leads to the development of several Apache Hadoop-based applications to produce insights from the user data. Launching of new or existing Amazon Elastic MapReduce job into the Karmasphere Analyst or Karmasphere Studio desktop tools is available at hourly pricing and no long-term commitments or upfront fees.

We can run Karmasphere Analytics with Amazon Elastic MapReduce in two licensing models – “License Included” and “Bring-Your-Own-License (BYOL)”. The Karmasphere software will be certified by AWS in the "License Included" model and hence no need for the users to separately purchase Karmasphere licenses. If the user already has Karmasphere licenses, we can take the "BYOL" model to launch Amazon Elastic MapReduce job flows with Karmasphere Analytics.

1) Karmasphere Analyst

Karmasphere Analyst is a visual, desktop workspace used by analysts and data professionals for exploring Big Data and interacting with them on Amazon Elastic MapReduce. It offers visual tools to utilize SQL, or similar languages, to create ad-hoc queries and interrelate with the outcome. The workspace gives access to unstructured and structured data situated on Amazon Elastic MapReduce job flows, Amazon S3, or local file systems. It provides intuitive analytics via graphical wizards and SQL. It also permits users to distribute results to databases, files, and other applications like Tableau or Microsoft Excel.

2) Karmasphere Studio

Karmasphere Studio is a plug-in developed for the Eclipse IDE. It presents a recognizable graphical environment for running the complete life-cycle. It is used for designing Hadoop applications on Amazon Elastic MapReduce, includes prototyping, testing, debugging, optimizing, and deploying etc., on those applications. Karmasphere Studio increases the productivity of developers, by simplifying the development of MapReduce jobs on Amazon Elastic MapReduce and thus saving time and effort. It is instinctive, visual interface makes it easy to use for the beginners to the experienced users [8].

3) *Features*

1. It contains structured dashboard.
2. It has SQL data explorer.
3. A Hadoop algorithm has 250-plus pre-packaged.
4. Karmasphere has SPSS, SAS and R Analytic Models.
5. Self-service analytics have dynamic data lenses.

III. Summarization of Big Data Tools

The following Table I demonstrates the comparative aspects of the diverse tools in big data based on compatible data sources and its operating system. The main objective of this comparison is not to criticize which is the best tool in big data, but to demonstrate its usage and to create alertness in various fields.

TABLE II
Comparison of Big Data Tools

Name of the Big Data Tools	Mode of Software	Types of Data	Data Sources	Database Support	Operating System
PENTAHO	Open Source	Structured and Semi-structured Data	Hadoop, NoSQL and analytic database	MongoDB and Cassandra HBase	Windows, Linux, OSX
JASPERSOFT	Commercial and Open Source	Structured and Unstructured data	JDBC, Delimited text, Positional text, LD IF, XML	Mongo DB, Cassandra, Redis, Riak, CouchDB Neo4j, HBase	OS Independent
SPLUNK	Commercial	Unstructured data Time-series textual Machine data	Files, the network scripted outputs	Relational IBM Database 2, SAP, Sybase	Windows XP, Vista, 7 and 8
TABLEAU	Commercial	Structured and Unstructured data	Database, Cubes, Hadoop Cloud	MySQL, Microsoft SQL Server, Oracle, EMC, GreenPlum	MS Windows 8.1, Vista or Server 2012 R2, 2012, 2008 or 2003
KARMASPHERE	Commercial and Open Source	Structured, Semi-Structured and Unstructured data	Web logs, Mobile devices, and Sensors	Base HDFS file data	Red Hat/CentOS/ Ubuntu Linux

IV. Conclusion

In this paper, more than a few big data tools were elucidated along with their features of several tasks. Big data provide vastly effective supporting processes for collection of data sets which is too complex and large. This mandatory requirement gives the way for developing many tools in big data research. Whereas these tools are generated both in real time and also in very large scale which comes from sensors, web, networks, audio/video, etc. Thus the aim of this survey is to enhance the knowledge in big data tools and their applications applied in various companies. It also provides obliging services for readers, researches, business users and analysts to make enhanced and quicker decisions using data which will promote for development and innovation in the future.

References

[1] http://www.slideshare.net/HarshMishra3/harsh-big-data-seminar-report
[2] http://www.infoworld.com/d/business-intelligence/7-top-tools-taming-big-data-191131.
[3] http://www.pentaho.com/sites/default/files/uploads/resources/pentaho_ba_solution_brief.pdf.
[4] http://www.asiasoft.hk/english/sfw/Jaspersoft-Business-Intelligence-Suite-ds-EN_AS.pdf
[5] https://www.splunk.com/web_assets/v5/book/Exploring_Splunk.pdf
[6] http://www.splunk.com/web_assets/pdfs/secure/Splunk_for_BigData.pdf
[7] http://casci.umd.edu/wp-content/uploads/2013/12/Tableau-Tutorial.pdf
[8] http://aws.amazon.com/elasticmapreduce/karmasphere/

Ms.S.Vidhya holds a B.Tech degree in Information Technology from N.P.R College of Engineering and Technology, affiliated to Anna University of Technology,Tiruchirappalli, Tamil Nadu, India in 2013. Now she is an M. Tech student of Information Technology department in K. S. Rangasamy College of Technology. She has published 2 international journalsand presented three papers in the National level technical symposium. Her Research interests include Data Mining, Web Mining, Wireless Networks.

Mrs. S. Sarumathi received B.E. degree in Electronics and Communication Engineering from Madras University, Madras, Tamil Nadu, India in 1994 and the M.E. Degree in Computer Science and Engineering from K. S. Rangasamy College of Technology, Namakkal Tamil Nadu, India in 2007. She is doing her Ph.D. programme under the area Data Mining in Anna University, Chennai. She has a teaching experience of about 15 years. At present she is working as Associate professor in the Information Technology department at K. S. Rangasamy College of technology. She has published 5 reputed International Journals and two National journals. And also she has presented papers in three International conferences and four national Conferences. She has received many cash awards for producing cent percent results in university examination. She is a life member of ISTE.

Dr. N. Shanthi received the B.E. degree in Computer Science and Engineering from Bharathiyar University, Coimbatore, Tamil Nadu, India in 1994 and the M.E. degree in Computer Science and Engineering from Government College of Technology, Coimbatore, Tamil Nadu, and India in 2001. She has completed the Ph.D. degree in Periyar University, Salem in offline handwritten Tamil Character recognition. She worked as a HOD in the department of Information Technology, at K. S. Rangasamy College of Technology, Tamil Nadu, India since 1994 to 2013, and currently working as a Professor &Dean in the department of Computer Science and Engineering at Nandha Engineering College Erode. She has published 29 papers in the reputed International journals and 9 papers in the National and International conferences. She has published 2 books. She is supervising 14 research scholars under Anna University, Chennai. She acts as the reviewer for 4 international journals. Her current research interest includes Document Analysis, Optical Character Recognition, and Pattern Recognition and Network security. She is a life member of ISTE.