BIG DATA ANALYTICS IN HR DOMAIN

[1]Praveen Kumar

[1]PhD Scholar,
[1]Faculty of Management Studies,
[1]University of Delhi, Delhi, India.

Abstract : This Big Data Analytics (BDA) is gradually becoming favored by decision makers / managers in all verticals of business entities for data related activities, due to its technological advancements and event based analysis features beside can store huge data of various types. Since almost a decade, it is extensively being used as analytical tools for prescriptive, descriptive and predictive related problems or scenarios, mainly for strategic decision making or forecasting purpose. Corporate world has adopted the BDA technologies to make their processes simpler and smarter. They also utilize BDA ecosystem (sources and storage of data repository) for various domains in large enterprises. Artificial Intelligence (AI) innovation, leads to further inclusion of BDA in domains like Operations, HR, Procurement, Finance and others. It also practiced in sectors like Medical Diagnostic, Banking and Insurance, service sector, E-Governance, Telecommunication and sports etc.

Big Data Analytics include technologies of Big Data and Analytics. Big Data which talks about features like data's volume, velocity, variety, and veracity etc., so it is extensively being used for data storage and streaming of large size data. The analytics is the science of data analysis. Hence the process of Analytics is data intensive and its accuracy comes with high volume of Real time data. The business use case has been used here to draw a method to utilize BDA in today's business environment specifically into Human Resource Management. This paper is about the Big Data Analytics in Human Resource (HR) and its business environment, its frameworks that are in practice by organizations at large. This will help in describing at best, the potential of technologies vis-a-viz maturity of HR processes in Business Eco-system. The purpose is to understand in depth BDA role in various facet of HR Management using various cases of implementation on selective enterprise, which has excelled in HR domain and made it as their core competency. HRIS can also use as Software-as-a-Service (SaaS) to support the function of an enterprise. Beside it is imperative to understand the framework and various BDA tools in market.

IndexTerms - Big Data Analytics, Big Data, Human-Resource Information System (HRIS), Human-Resource Management (HRM).

I. INTRODUCTION

"There is little debate these days about the importance of big data and analytics in supporting the strategic goals of an organization" (Davenport, 2006). With strategic goals are derived from the management prospective to gain certain output, related to development of competitive advantages, profitability, innovation or expansion. In all of these key objectives of strategies, the role of Human Resource is critical and will be the pivot for success. As one of the strategic goals of an organization is to achieve cost benefits and maximum resource utilization by Human Resource Management (HRM), i.e. using HRM to integrate with the strategy and the strategic needs of the firm in order to gain competitive advantage (Wright and Sherman, 1999). The role of Human Resource Management (HRM) Practices on the firm output and its impact on an organization is well understood through various studies (Rogers & Wright, 1998; Becker & Gerhart, 1996; Paauwe and Richardson, 1997). The forecasting for huge data volume in an organization, is being done by many researcher and IT practitioners as depicted in a HBR journal by (Schrage, 2016). "Organizations should expect 10 to 1,000 times more data in the next 12 to 18 months. To make data more valuable, organizations must consider how to define, measure, and assess value creation inside the enterprise and outside." Hence due to characteristics of Big Data, that makes it favorable options out of rest to withstand the volume of data handling, which can be achieved by it. So there is a linkage of information seeker and opinion giver between Big Data and Analytics and Human Resource Management.

Human Resource Management (HRM) can be well concluded with its objectives, by mean of which one can concise its requirement in an organization, these 4 (four) Major Objectives (Management study guide, nd) are : (i) Societal Objectives, (ii) Organisational Objectives (iii) Functional Objectives, (iv) Personal Objectives. All the above objectives have some significance at various levels of organization, and at each stage the need of HR information (data) will increasing many folds, due to recent developments in technological front, sources of data increases (Social Media, Mobile etc.) and so do its size (volume in Terabyte TB or Picobyte PB) and in various modes (structured and unstructured) related to the Human Resource. The amount of data generated for a medium to a large enterprises on yearly basis range from some GB's to few TB. Thus there is a need of Big Data and Analytics in Human Resource Information System (HRIS), it is in use also. Big Data will going to provide the data repository for both structure and un-structure data, due to its versatile feature to create datasets using variety of data (structure, semi-structure or unstructured). Even though enterprises had long understood this demand and had established HRIS (Human Resource Information System) to tackle all the data demands issues.

In traditional HRIS system, the IT into HRM had fulfilled all objectives requirements, by providing vital information from available data in repository, but due to traditional HRIS, which can hold structure data as in form of RDBMS (Relational Data Base Management System) only. It can easily predictable, how knowledge (Delayed and non-real) can be acquired from it. Through Wisdom Hierarchy (Aswathappa, 2005) as depicted in Figure 1, we can identify if the source of information is not correct or partial, this will result into incomplete knowledge and hence wisdom can't be created from it. Traditional HRIS doesn't hold real-time streaming data both structure as well as un-structure, Due to which objectives are partially fulfilled hence realistic decisions can't be taken through it (in complete wisdom).

Figure 1: DIKW hierarchy Model (Aswathappa, 2005)
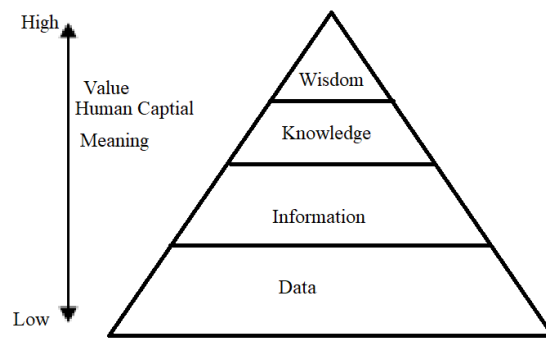
A typical HRIS will consist of the database in structure format that is relational DBMS (RDBMS). Along with application hosted on the Application / Web server (AppServ) to give access to users for populating and retrieving of Data or information as depicted in Figure 2.

Figure 2: Traditional HRIS with App server and RDBMS
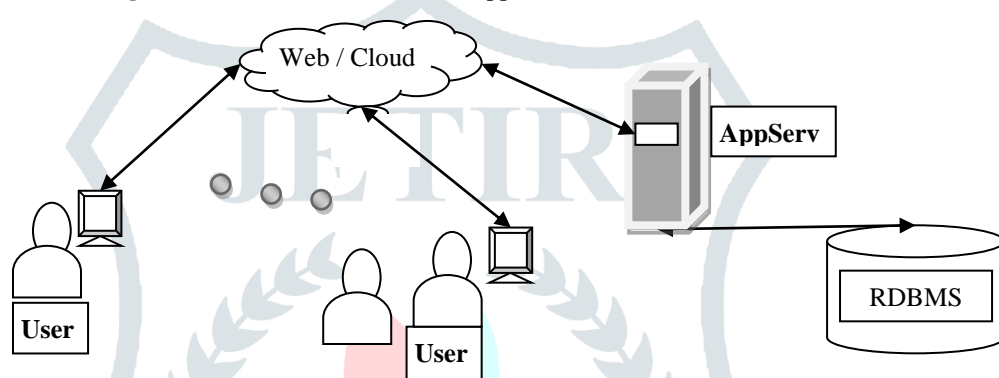
Thus role is either disseminate, populate or calculate the data information input to the system as discussed in Table 1. However with the addition of Big Data Analytics into the HRIS to make it a Big Data HRIS, it would be more real-time data intensive. Analytics when used over it results into information for organization which earlier was reporting only (MIS), Where as in today world more predictive and future oriented, i.e. the biggest opportunity that big data offers to HRM is in predictive analytics around high-volume, large variety, higher data creation from input source and repeatable processes. It is now the time for HRIS or so to say HRM is to stop reporting and start predicting.

Table 1: Comparative on HRIS traditional and Big Data HRIS

	Traditional HRIS	Big Data HRIS
Data type	structured only (RDBMS)	Structured and un-structured both
Data Sources	Inputs, spread sheets, inter-dept. tabular data	Text, voice, social media, and inter-dept tabular data
Analysis Scope	Limited all relational data (Past info)	Extensive internal and external environments
Utilisation	For info dissemination and recording with MIS reports on past performance	Beside the traditional, this can be utilized for the predictive analysis and proactive policies and plans
Examples	ERP modules on HR, Web-Portals, Websites	HR Analytics software's, cloud and on-premises based (SAS, Oracle, and Naukri etc.)

Other benefit that is readily available with Big Data is the role in decision making as "Data-driven is often defined as something that is based on data and facts instead of intuition or personal experience" (Rasmussen & Ulrich, 2015). So the role of analytics on Big Data has a role in arriving at the decision pertaining to various problems (functional, personal or organizational). Hence as depicted below in Figure 3, based on the Figure 1.

The resultant of Big Data and analytics on HRIS will result into a more wise decision on basis of knowledge that the system acquired, with predictive analytics. It create the wisdom that helps in various real-time problem solving. To be discussed through a case study, where in it has been done to understand the role of BDA on HRIS performance, more specifically on the enterprises performance.

Thus the role of data in management objectives can well be understood and so the wisdom to perform the decision making task. Information technology plays a vital role in converting the data to wisdom through Big Data Analytics (BDA), which is an

extension for traditional HRIS. For any organization the principal aim of HRM would be to attract, select, motivate, and retain talented Resources to perform their roles, and in this information technology has its role to play. As very well explained "technology has transformed the way HR processes are currently managed, essentially how organizations collect, store, use, and disseminate information about their HR*"* (Stone et al, 2015).

The sources of data involves in recruitment, career progressed information, training attended, absenteeism records, productivity, reviews, competency profiles and staff satisfaction, along with scanning of one's social media information and activities, as to get Datasets to analyze and gauge employee sentiment and psychology. This would be a detrimental in elimination process of candidates to be recruited or for promotion purpose. It also fulfill the HRM four Objectives as mentioned earlier by:

(i)   **Societal Objectives:** understand the Corporate Social Responsibility (CSR) benefits and also as to how to contribute on social causes, a mode to interact and understand the sentiment through social media (source of data in HRIS)

(ii)   **Organisational Objectives:** Aligning the HR policy in tune with Management Objectives and goals can be achieved and correlated with BDA in HRM

(iii)   **Functional Objectives:** To calibrate the Cost and benefit Analysis and track as not to be more expensive and show management about the ROI (Return on Investment) for each Human Resource employed.

(iv)   **Personal Objectives:** providing the details of career objectives set by organization, along with various information and stats on employee centric.

II. RELEVENCE OF STUDY

This study is important as to identify the linkage between the HRM and BDA, Big data and Analytics is a buzz word in industry for quite long, hence the finding will add value for current and future HR Managers, in adapting the technology with best mode of evaluation based on the framework offered along with it. Besides understanding the benefits that it offer to its enterprises.

III. OBJECTIVES

The objective here is to identify the linkage between BDA into HR domain. As this will add value to further identify the relevance of technology and its impact on the way HR domain works vis-a-viz the technology enabled one. With technology identification of risk on hiring to training Tracking the resource with feedback system and reviews as to identify any deviations in the objectives of resource with that of company.

IV. RESEARCH METHODOLOGY

The research is a descriptive Method with evaluating various research paper and case-studies, that are being carry out by industry experts to understand the framework, identify the linkage, beside understand its approach and objective that are being achieved with it. Here the task of identification of HR analytics based framework with BDA at the core of the system has been done. To give a better and broad understanding of the HRIS, analysis of various case studies of large enterprise has been done as to achieve the target set for the better understanding of its framework and process implementation.

V. DESCRIPTIONS

The details on the basics of technology and HRIS is to be studied in detail as to get the glimpse of these below, including the Big Data Analytics and HRIS and then looking for the comparatives

5.1 BIG DATA CONCEPTS:

The Big Data technologies are the group of technologies that constitute the basic framework for data storage, streaming and retrieval mechanism. This include the technologies for storage, ingestion, API for database link and data analysis, finally data display or Dashboard. There are various frameworks available like Apache Spark, Apache Strom, Apache Hadoop and many others. However as the most usable and acceptable solution across the industry is Apache Hadoop.
Some of the utilization as being found in Indian context are:

(i) UIDAI (Unique Identification Authority of India): utilized the Big Data Hadoop to manage mammoth of data over 125cr 1,247,732,208) (UIDAI. 2019) of profiles with various formats of information in Multi-Structured Data, which means combination of Structured Data, Semi-Structured Data and Un-Structured Data

(ii) Income Tax of India: recent currency reprinting order by Govt lead to black money flush into banking, the Project Insight has been setup there to be used Big Data technologies to track tax evasion. It will help to combine the databases (Big Data) of IT returns, TDS & Bank accounts statements, IT forms from financial institutions of the country. Also, as part of Project Insight, a Compliance Management Centralized Processing Centre (CPC) will handle verification, management of campaign and the generation and follow-up of letters and notices (Prasad, 2019).

5.2 HADOOP FRAMEWORK:

(a) Hadoop is an open-source, a Java-based programming framework (Wikipedia, 2019) that continues the processing of large data sets in a distributed computing environment. It based on the Google File System or GFS. It constitute of:
(b) HDFS (Hadoop Database File System): it provide the storage layer for Hadoop, with distributed data storage environment.
(c) HBase: it is NoSQL and Non-Relational data base, with random, real time, read write access to the database.

(d) Sqoop: it's a tools help to transfer the data between Hadoop and NoSQL. It is also used to export data from relational DB like Oracle or MySQL to Hadoop, or vice versa.

(e) Flume: it is the distributed services for ingestion of data.

(f) MapReduce is much advance to support Mapping of Data and reducing the task by means of programming models. Pig and Hive are based on MapReduce framework and are used as the data flow tools

(g) Impala: it is High performance SQL engine used to access the data from Big Data using MapReduce framework.

(h) Cloudera search: flexible, scalable and robust storage system to perform the search on the database (Big Data HDFS).

As the above tools are just a glimpse of the various tools available in market. To sum up there are various stages that are there in the Big Data framework depicts as shown in Figure 4 below.

Figure 4: Big Data Processing Framework (Dhankhad, 2019)
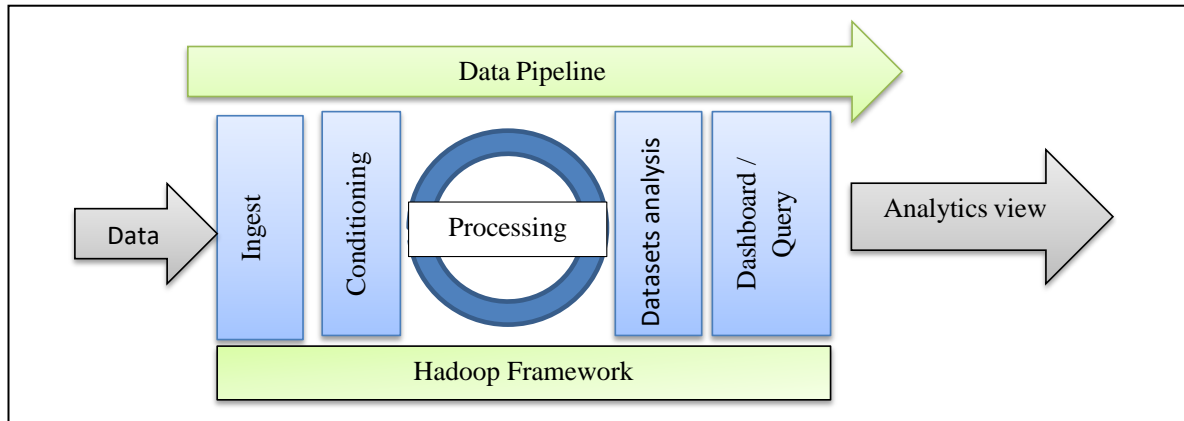
(i)   **Ingested** - where data is ingested or transferred to Hadoop from various resources, such as relational databases system or local files e.g. Sqoop, flume.

(ii)   **Processing** - data is stored and processed e.g. HBase. Spark and MapReduce

(iii)   **Analyzing** - data interpreted by the processing framework e.g. Pig, Hive & Impala.

(iv)   **Access** - the analyzed data can be accessed by users e.g. Hue and Cloudera search.

5.3 BIG DATA ANALYTICS:

The Big Data and Analytics is about the concept of Big Data and analytics. The very concept of BDA can be explained using the framework used in the establishment, where in the eco-system of Big Data be utilized as the base for getting data on employee and related organization activities.

Technology solution provider SAS:

"Big Data Analytics examines large amounts of data to uncover hidden patterns, correlations and other insights. With today's technology, it's possible to analyze your data and get answers from it almost immediately – an effort that's slower and less efficient with more traditional business intelligence solutions." (SAS, nd)

5.4 BIG DATA ANALYTICS FRAMEWORK:

The more to do with the BDA Framework Figure 5, as to understand as to why there is an importance of Data Sources on the establishing of the Eco-system. the only way is to have the right source must be result into creation of datasets (correct) and using these datasets into establishment of data lake or more precisely data warehouse which constitute all the analytics and data repos for the HRIS or precisely for the Enterprise data (MIS or ERP). If the dataset found to be correct and true, which can be represent it using the data visualization over the dashboard. Overall this will provide an overview on the available datasets and also using the analytics tools we can perform the wise decisions and make the HRIS as smart software.

Figure 5: Big Data Analytics Framework
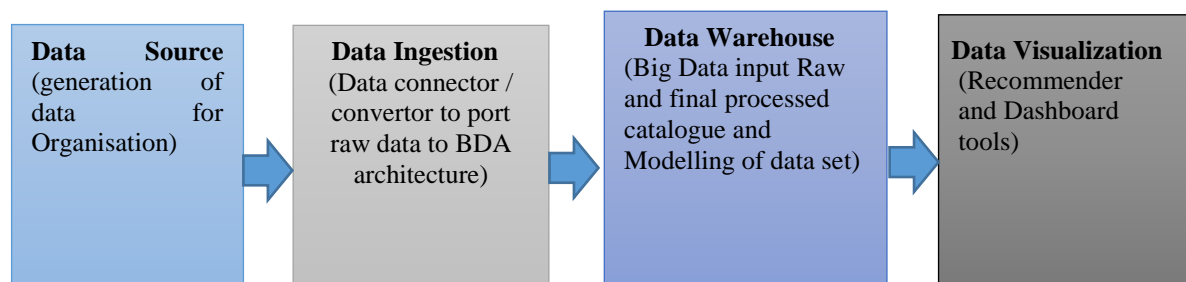
The BDA framework as represented above is to demonstrate the way to utilize the BDA effectively and efficiently. This also give a compliance base for any organization in implementing the solution across the enterprise.

5.5 BIG DATA ANALYTICS TOOLS:

There are various BDA tools available and with various purpose to achieve the targets, there are certain tools marked in the Table 2, with specific purpose assigned for each stage of the BDA frame works. The same can be summarized as under:

Table 2: Big Data Analytics tools

S.No	Stage of BDA	Tools
1	Data Source	CRM, CSV, Logs, NoSQL. Excel, Social Media
2	Data Ingestion	Web, API's, DMS, Sqoop (SQL-Hadoop connector API)
3	Data Warehouse	Data Modelling (Spark/ ML/ R), Data warehouse, Data lake, Redshift (Data warehouse tools)
4	Data Visualization	Tableau, Qilks, SAS View, BI Tools, Data Studio

5.6 CASE STUDIES SUMMARY : HRIS

Summary 1: Implementation of HRIS with BDA components:

Problem statement: The Brokerage firm doesn't have the solution to give its manager any kind of review to its reporting workforce.

Large Brokerage Firm Uses Dashboards to Deliver Key Workforce Insight to Its Managers (Oracle White Paper, 2011).
Size of the firm: 10,000 employee, Industry: Brokerage Firm

Products Implemented: HR analytics, PeopleSoft Entrprise - Human Capital Management v8.9, BI Publisher.

Implementation time: 3 months as compare to 9-12 months for any application development cycle using agile framework based technologies.

GENERAL BENEFITS:

Less time to implement the solution, less time to build the reports ETL matrix and other warehouse data. HR analytics helping the firms to move towards evidence based decision making, eliminating the rumors.

Within 5 odd months the technological advancement leads to getting the dashboard reports meant for the Managers as desired.

With HR analytics dashboard on roll, the accountability of Managers are held in the area as hiring diverse talent, improving productivity and engagement of talent, developing talent, and retaining key talent.

Benefits of ROI: the financial gain that the organization has created using the HR analytics are primarily into cost saving in building the customized in house solution to a pre-built solution, which fits into the process well. Second is the ability to address workforce issues with the insight to action functionality available through the Analytics Framework and analytical workflows.
The bottom line is the lower cost of ownership, as pre-built solution reduced both the IT and Human Resource IS (HRIS) development and maintenance cost. Beside the reduction in the deployment and implementation time.

Benefits of Business Value addition: HR Analytics give Business value by providing information that which workforce segment creates more value and by how business will be impacted by impending them and by how much. Line managers want to know how much money is being lost each day by not having hires in place and how that impacts current projects. It also provide the long term and operational insights on business functioning and the impact it will be having on the customers, employee and profitability by either reducing the turnover or further expanding the business both in long run and short run.

Benefits to HR: The insight on pure HR matters like

(i) Minimized risk of losing key talent: as value creation talents could be protected by understanding on the impact of no availability of talent to having it.

(ii) Managing the internal versus external hiring cost: the cost or tradeoff of hiring talent internally versus external will be evaluated and compared with time and value it would be offering to the enterprise.

(iii) Scaling the salary with value factor: the HR analytics will support in understanding value with respect to salary offered to individual resource that are being hired and deployed, beside actual calculating the project cost.

(iv) Avoid any dissatisfaction in employees: using HR analytics the employee satisfaction can be maintained, results into a satisfied customer and continued revenue generation, there is a notion of a link between employee satisfaction, customer satisfaction, and financial results.

To sum up the HRIS with BDA is more value to its Management for review and results as compare to employee, however it would be benefit in long run to employee too, if they ignore or negate the privacy aspects

Summary I1: Utilisation and Implementation of BDA in traditional HRIS:

Company: MNC (Coca Cola Enterprise- CCE)

Problem Statement: Reduce the centralized HR reporting and analytics service to the business. And establish a HR analytics. (Coca cola enterprise, 2011)

Size of the firm: 5 continent footprint, Industry: Beverages drinks
Products Implemented: in house developed HR analytics over HRIS
Implementation time: 2010 initiated and close in 3 yrs with agile structure of development and implementation.

General Benefits:
The ultimate goal is to have analytics in HRIS, to move towards mature data processing that utilize the descriptive analytics of HR and business data. To tackle the repeated and lengthy task of reporting and reviewing of the team. At various location (5 continents). A centralized dashboard has been developed with automation of the low value information as to reduce the review pressure from the managers. It's a recursive process and hence agile method of updation has been implemented.

Benefit of ROI: Simple descriptive analytics is being used. As to identify the gaps and self-explained to the employee their faults and non-compliance to policy or duties. Hence ROI is there in terms of efficiency of the employee.

Benefits to HR:
Analyse the data and also scoreboard based problem to help in understanding issue with it. Its main benefits are more qualitative:
(i) More staff review can be done with new descriptive analytics enabled HRIS
(ii) Root cause analysis can be done and review work would scoreboard based, instead of details summary readings.
(iii) Employee Satisfaction and work completion will be more visible and calculative.

KEY FINDINGS OF THE STUDY

There exist a clear linkage between technology adaptation and the value addition by it in the process. Hence HRM when being integrated with BDA technology utilization, will results into an eco-system for the employee of an organization and its surrounding, besides the holistic development planning can be done. This will not only benefit the organization (Organization Objectives) but also boost the career graph of employee (Personal Objectives) associated and attached with the system.

IMPLICATION OF THE STUDY

This study has been done to identify the linkage in HRIS and BDA and adaptation of technology in enterprise level has its own benefits and drawbacks. If the benefits are to be highlighted with the ROI then there is a clear case of privacy breech that exist with it. As the involvement of social media data and its sentimental analysis talked about in the BDA integration result into no freedom of thoughts and privacy. Hence implication of the study is both sided and one has to take up the benefits with some losses too. Employee will be more cautious as not to divulge the information or opinion regarding the enterprise.

REFERENCE:

Aswathappa K. (2005). Human Resource and Personnel Management, Text & Cases.(4th e.d). Tata Me Graw-Hill Publishing Company Ltd., P.8

Becker, B., & Gerhart, B. (1996). The impact of human resource management on organizational performance: Progress and prospects. Academy of Management Journal, 39: 779-801.

Davenport, TH. (2006). Competing on Analytics. Harvard Business Review 84(1): 99-107.

Coca - Cola Enterprises (CCE) (2011) Case Study: The Thirst for HR Analytics Grows, https://towardsdatascience.com/a-brief-summary-of-apache-hadoop-a-solution-of-big-data-problem-and-hint-comes-from-google-95fd63b83623

Dhankhad Sahil. (2019, Apr 10). A Brief Summary of Apache Hadoop: A Solution of Big Data Problem and Hint comes from Google. [web log post]. Retrieved from https://towardsdatascience.com/a-brief-summary-of-apache-hadoop-a-solution-of-big-data-problem-and-hint-comes-from-google-95fd63b83623

Management study guide. (nd). HRM definition and objectives, retrieved 10 Feb 2019 https://www.managementstudyguide.com/human-resource-management.htm

Oracle White Paper. (2011). HR Analytics: Driving Return on Human Capital Investment.

Paauwe, J., & Richardson, R. (1997). Introduction to the special issue: Strategic human resource management and performance. International Journal of Human Resource Management, 8: 257-262.

Plessis, Dr Andries J Du. (2015) The use of big data and hris by hr practitioners in new zealand: empirical evidence from a longitudinal study, GBATA paper

Prasad Gireesh Chandra (2019, Apr 03). Income tax department plans deep dive into big data this year, live mint, Retrieved from https://www.livemint.com/news/india/income-tax-department-to-use-big-data-in-full-blast-from-fy20-1554262557273.html

Rasmussen, T. & Ulrich, D. (2015). Learning from practice: how HR analytics avoids being a management fad. Organizational Dynamics. 542: 1-7.

Rogers, E. W., & Wright, P. M. (1998). Measuring organizational performance in strategic human resource management: Problems, prospects, and performance information markets. Human Resource Management Review, 8, 311-331.

SAS. (nd). Big Data Analytics, Accessed 10 February 2018 from http://www.sas.com/en_us/insights/analytics/big-data-analytics.html

Stone D, Deadrick D, Lukaszewski K, Johnson R. (2015). The influence of technology on the future of human resources management. Human Resources Management Review. 2015;25: 216-231

Unique Identification Authority of India (UIDAI). (2019). Accessed on 10 Jan 2019. https://uidai.gov.in/aadhaar_dashboard/

Wikipedia. (2019). Apache Hadoop. Accessed 12 Mar 2019 from https://en.wikipedia.org/wiki/Apache_Hadoop

Wright, P. M., and W. S. Sherman. (1999). Failing to Find Fit in Strategic Human Resource Management: Theoretical and Empirical Problems, in P. M. Wright, L. D. Dyer, J. W. Boudreau and G. T. Milkovich (Eds.), Research in Personnel and Human Resources Management, supplement 4, Stamford, CT: JAI Press, 53-74.