DOI: 10.21276/sjbms

Saudi Journal of Business and Management Studies **ISSN 2415-6663 (Print)**
Scholars Middle East Publishers **ISSN 2415-6671 (Online)**
Dubai, United Arab Emirates
Website: http://scholarsmepub.com/

Big Data Technology for Smart Talent Management System

Sehrish Khan Saddozai[1*], Peng Hui[1], Muhammad Saad Khan[2], Umair Akram[1], Suhail Memon[1], Asma Zubedi[1], Zahid Latif[1]

[1]School of Economics and Management, Beijing University of Posts and Telecommunications, China.
[2]Electrical Engineering Department, Bahauddin Zakariya University, Multan, Pakistan.

***Corresponding Author:**
Sehrish Khan Saddozai
Email: sehrishkhan5259@gmail.com

Abstract Big data is currently a hot topic in HRM and academic research. It is undeniable fact that now a days as more and more data is being accumulated and stored, therefore analyzing and using data is becoming more and more reasonable. Big data technology use in Talent management is very under researched topic and the discussion mostly spins around the usefulness of Dig data in TM field like talent management. In this paper, we have focused on the operational side of TM and how to implement big data for Smart E-HRM in Talent management. In this paper we have proposed a model for Smart TM system and also reviewed that how Big data may be implemented for a Smart TM system and how this system can benefit Talent management in terms of recruitment, promotions and assessment of the employees. In this paper we have proposed a Smart TM system in a telecom organization that will consists of several offices belonging to the same organization and will be linked through a common server and by using big data technology all the employees progress will be analyzed, and by using this data changes will be done immediately in order to solve the crucial issues of talent management. We also performed quantitative analysis about this system with the help of unstructured interviews from employees working in a telecom organization.

Keywords: Big Data, Talent Management, E-HRM, Hadoop

INTRODUCTION

Today we live in era of digital communication where user information and their data are ubiquitous [1] and everybody collects data either it may be consumer related medical related or it may be related to management. But a key question arsis that what should be done with this important data? What will be its importance in terms of Human resource? especially in the field of Talent Management. Some people things that big data is just another E-HRM system but other think of it as a game changer which will take the place of professional HRM. There are many theories on which we can debate on but in the end we see that big data is real and will have a global impact on Talent Management. It is therefore very important to deal with bid data in TM in form of E-HRM and try to find out its usefulness in terms of cost, time saving and efficiency.

In various literature and popular articles and on social media sites there is an intense debate on TM and its practices and challenges that are faced in many organizations [2]. TM is described as a precise approach to identification, recruiting, retention and development of talents for the benefit of an organization [3,4]. In TM basically the organizational needs are mostly highlighted. TM is basically used for the fulfillment of the organizational needs to contribute for the firms profit, performance and sustainability [5, 6].

There is also some hidden risks in using big data for E-HRM in TM as this technology in the future may eliminate HR department and the department like finance department may not need HR department anymore and this new technology may have some implications [7].

The recent developments in web technology and day to day use of internet by the people for daily use has created huge data in form of information which contains billions of information related to companies, industries, academia, movies, products and jobs etc. Due to these technology advancements there is an increase in online web services such as online services for movies as NETFILX [8], online job finding and recommender sites as LinkedIn [9] and online shopping services like eBay and so on. This online service provides us with huge amount of user related data according to user preferences. As a result of this huge amount of data are accumulated over time and creates overloading of data. Due to this overloading of data search engines are not showing promising results, therefore recommender system is the key solution for this overloading of information. Job recommender system for talent management using big data in the past did not work well on big data. So in this paper we have used big data with E-HRM system for Talent Management. We have proposed a model in which by using big data we will

connect several offices belonging to the same telecom organization, thus creating a smart recommender system for Talent Management.

Moreover Big data is now considered as a tool that will influence Talent management in some way or the other and will help in utilizing data in order to find better and suitable candidates for the job. Recommender system is used to find the right candidate for the right job without user intervention [10]. An example of this type of system is amazon.com, ebay.com and LinkedIn have become very popular form the last decade [11,12]. Now days a lot of research has been going on smart TM system for recruiting young talents and also using this system many HR responsibilities can be minimized in terms of employee's management and promotions. Practices especially for the job recommender system, E-recruitment is one of the most popular service used by people to find jobs. In the past these systems used very small amount of data that was mostly unstructured. Now a days we have large amount data that is mostly unstructured and heterogeneous as compared to structured data, which is now slowly but steadily growing at tremendous speed.

In this Paper we have focused on 24 offices of the same organization and proposed a smart E-HRM system linking all 24 offices and we will see how this system will be used for TM practices. This paper is comprised of seven sections; Section.1 is about introduction and Section.2 is about data extraction insight in which we will show how this raw data related to employees will be transformed into wisdom. In the section related to big data we will explore the different stages that are related to collection of different data and visualizing intelligence. In section.3 we will discuss how Smart TM will use big data platform in order to make smart E-HRM decisions for TM. Section.4 consists of data processing techniques that will be used in this TM system. Section.5 consists of our proposed model. Section.6 consists of research methodology. Section.7 consists of data analysis with the help of quantitative analysis. Section.8 consists of hypothesis. Section.9 is about Research Involvement and sample population. Section 10 deals with data analysis results. The paper will conclude in section 11.

RAW DATA ANALYSIS FOR BIG DATA

Data Stages:

In order for E-HRM system to be smart it should contain the necessary data, knowledge and intelligence in order to reach the level of smartness [13]. It consists of four data stages which are all interlink to each other as shown in table below.

Table.1 Raw data analysis

User Data	User Information	Knowledge	Intelligence
Raw data	Unstructured data	Actioned data	Wisdom data

Here the raw data is gathered from the employees from the data sources that are linked to them in the form of RFID cards and biometric identification system. In addition to this system the employees also use online portal in order to update their status and the assigned tasks which they are given. This raw data is then encapsulated and forwarded through different network protocols and this information may be transmitted to the gateway, and from here is it will be sent to the Big data cloud computing platform. Websites now a days are becoming more and more intelligent by predicting our choices [14], same is the case for Smart E-HRM system layer, which will do the crucial tasks like look, listen, adapting, predicting, connecting and correcting on the Big data in order to make smart decision, therefore unlocking an intelligent system.

Big data infrastructure:

As shown in fig.1, it shows different data stages in big data management in order for the system to become intelligent. All these stages are discussed below.

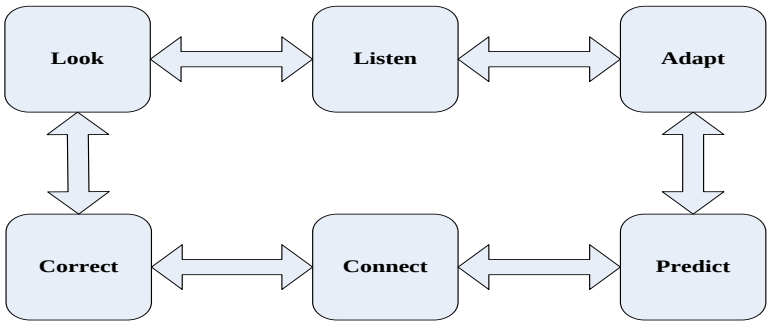

Fig.1 Data stages for Big data

a) Look: Here Look refers to as data which is in the raw form that is collected and indexed by the big data platform in order to access the data smoothly. In smart E-HRM system the data is collected in various forms i.e. human interface which may be RFID cards, biometric system, the data of the tasks given to the employees and their evaluations by their superiors and from the data given on the online web portal where the employees online profiles are already stored in the main server. All these data forms are then stored in a column based data server or cloud like bit map index and tree based index.

b)Listen: Here listen refers to the data in the form of information, here the big data platform analyze the received data, filters it and remove the unwanted data so it can only deal with the useful and valuable data in order to make intelligent decisions about the employees.

c) Adapt: Here Adapt refers to our knowledge data, here the extracted data is analyzed through the criteria of clustering and grouping the received information in order to provide a clear decision to the given entity.

d)Predict: Predicting is the intelligence that our platform will do by mining rules and will compare this data with the past and also with the present data in order to give future values. Analysis such as predictive analysis is used here in order to give correct predictions.

e) Connect: Here connects refers to the wisdom that the platform will give by integration of past and present values to get the required experience of the employees.

f) Correct: Here this data refers to the visualization of the employees data from the previous stages thus giving intelligent decisions based on the employees output.

BIG DATA ANALYSIS AND COMPUTING:

In order for a smart TM system to make intelligent decision regarding Talent Management processes like recruitment, promotions and tasks assignments, the data need to b analyzed in a timely, proactive and fast manner. The employee's data need to be analyzed and compared with the previous data in order to get a clear insight into the employee's performance. In order to analyze this type of data there are two types of computing methods that will be used here i.e. Stream and Batch Computing.

a) Stream Computing: The era of Big data has led to the development of stream computing i.e. online real-time data along with distributed data stream computing [15][16]. Many organizations rely on real-time data streaming in order to spot fundamental issues like trading, system detection and employees progress by evaluating day to day tasks given to them [17][18]. In our proposed system clustered hardware processing of parallel data is adapted to help us process real-time data without any hurdle. Also sophisticated models are employed in order to extract the required real-time data.

5) Hadoop is most commonly used for Batch Computing system. Hadoop is basically a java based open source software framework used for distributed processing of huge datasets across of different distributed nodes. Hadoop framework was developed by Doug Cutting and Mike Cafarella in 2005 [19]. Apache Hadoop was developed to handle multiple servers each with multiple storage and computation capacity. It not detects failure at the application layer but also delivers high availability [20].

Hadoop Architecture: Hadoop basically consists of following two main components [21].

1. HDFS (Hadoop Distributed File System)
2. MapReduce

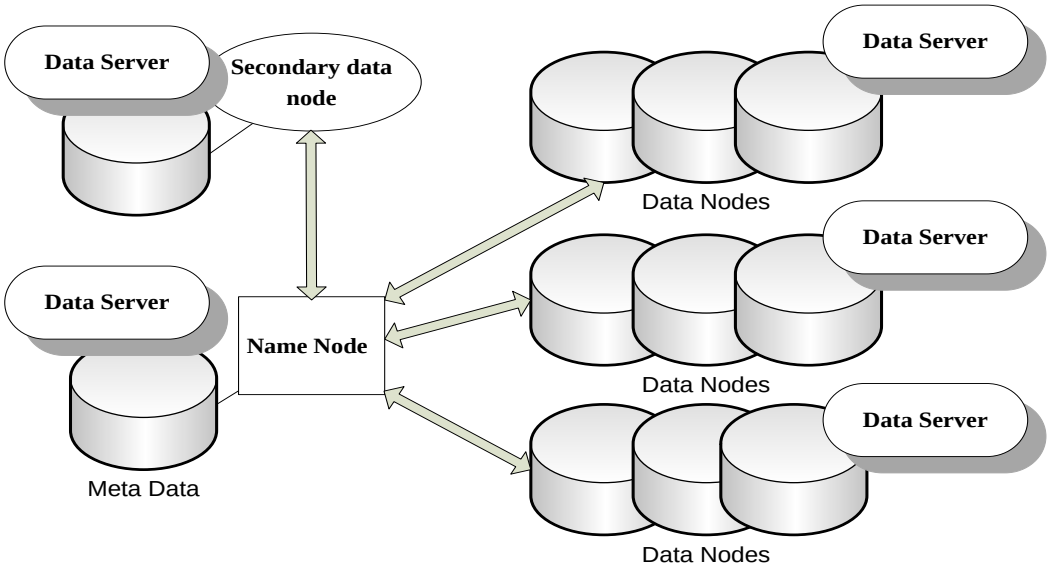

Fig.2 HDFS Structure

1. HDFS: HDFS structure is shown in fig.1. HDFS is basically used for handling files in Big data environment. The data files are first processed and then they are distributed into small blocks of sizes between 64-128 MB and then they are again distributed through nodes. For accuracy, each block is duplicated across different nodes. These nodes serve two main purposes in HDFS cluster, the Name node functions by managing the block distribution, by tracking their location and their operation, and also storing the metadata information.

2. MapReduce: The programming model used for the Hadoop framework is called MapReduce [22]. MapReduce is used for the processing of data for large clusters in parallel manner for the Commodity server. Hadoop works in a very reliable manner by distributing tasks to multiple nodes on the cluster. Thus as the name suggests MapReduce has two phases i.e. to map and to reduce. As the dataset is given to MapReduce, the dataset is already in the form of different data chunks, so this dataset is processed in parallel by the map tasks. MapReduce reduces the tasks on the input by sorting the map tasks output. The file system stores the reduce tasks as well as the input and output of the Map Tasks. MapReduce not only take cares of the schedule tasks but also monitors all take care of the failed tasks if any if any. As we know that nature of the web data is mostly heterogeneous and is in the unstructured form. So in order to deal with this large amount of unstructured data we have proposed a smart TM system in order to deal with the TM issues. This system will evaluate each candidate according to their performance and in the end will give the results according to their efficiency and task completion. Fig 3 shows the proposed smart E-HRM system. The amount of unstructured data is in huge amount as compared to structured one so we get more information from unstructured data. So in order to handle this unstructured data we need to formulate some technique in order to efficiently handle this data.

Proposed Smart TM System:

As we know that nature of the web data is mostly heterogeneous and is in the unstructured form. So in order to deal with this large amount of unstructured data we have proposed a smart TM system in order to deal with the TM issues. This system will evaluate each candidate according to their performance and in the end will give the results according to their efficiency and task completion. Fig. 3 shows the proposed smart TM system. The amount of unstructured data is in huge amount as compared to structured one so we get more information from unstructured data. So in order to handle this unstructured data we need to formulate some technique in order to efficiently handle this data.

RESEARCH METHODOLOGY:

a)Processing of Unstructured data:

There are number of methods for converting unstructured data to structured data and extracting useful information from them, data mining, digital image processing, text analysis and natural text processing are some of the well known data processing techniques. In our proposed method we have used Hadoop to transform unstructured data to give us structured data. Here our employees data is in the form of Unstructured data i.e. employees data base containing employees skills, qualifications, their designated positions, tasks given, their efficiency all are stored in the local server which is also connected to the main server. Hadoop basically processes huge amount of unstructured data though two main techniques which are, HDFS and MapReduce techniques.

b) Stored data Processing:

The data of the employees already stored in the min server is actually noise ad may contain many unwanted, duplicate and irrelevant data. So this data should be analyzed and preprocessed. Canonicalization is one such technique that will be used for the preprocessing of this data. This technique is used for changing information into standard type of information. This type of technique is useful to indentify particular job assignment for a particular person which specialization in the required field. Is also helps to define tasks for person like Software analyst, Software developer, technical engineer by utilizing entity resolution technique by using text analysis [23].

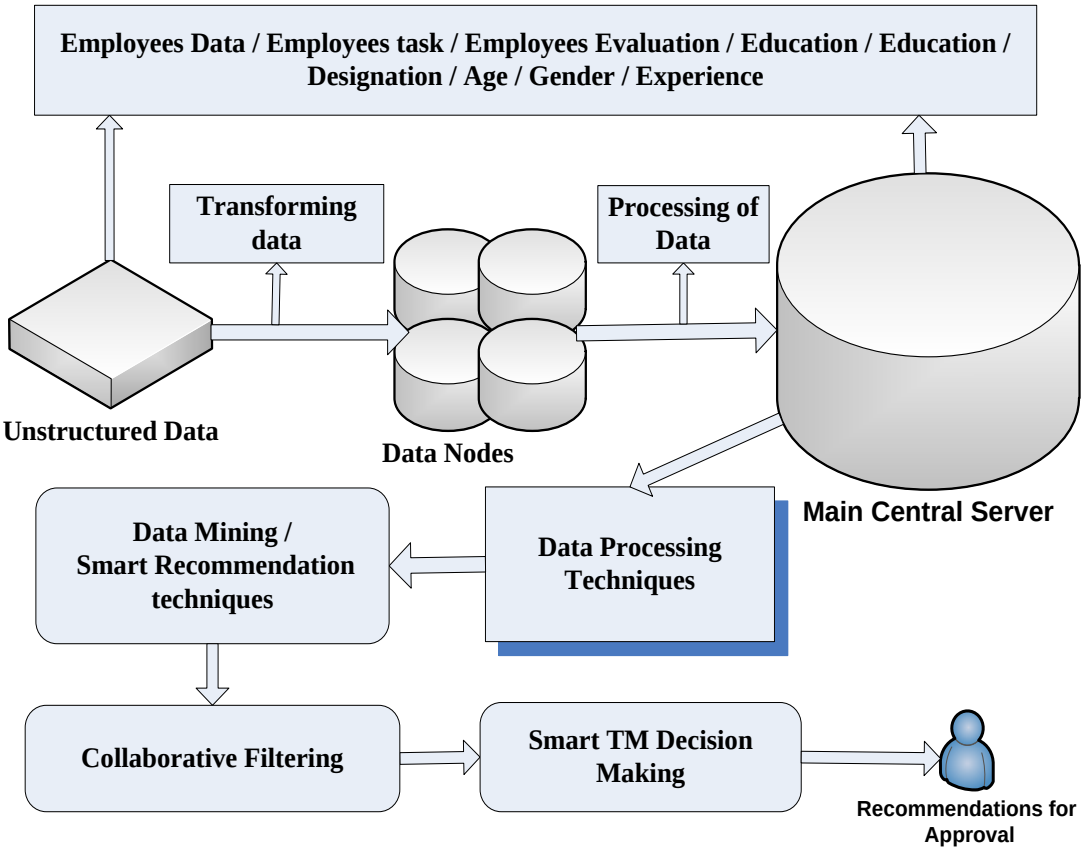

Fig.3. Smart TM system Model

c) Smart TM Data Modeling:

As we know that there are various recommendation based techniques like hybrid base recommendation, collaborating and content based recommendation. But as more and more data accumulates, our smart E-HRM system use data mining technique for processing huge sums of data [24]. The various methods that are employed in data mining are clustering, classification and association etc. In our model we have proposed Reciprocal Recommendation Technique with association and clustering rules for smart decision making process [25].

d)Filtering data for Smart TM Decision:

In this type of filtering the candidates personalized features are accessed. For example like in job recruitment, a candidate only likes to work in office and does not prefer field job. In case of promotion the candidate would like to serve in the same city and also keeping in view about the candidate experience and expertise, keeping in view the candidates progress and his achievements. This process also takes feedback from the all the managers about the employees from an online evaluation forms and same case goes with the managers are also evaluated from the employees. Thus in the end the Smart TM system will give more convincing and relevant results for TM practices.

e) Collaborative Filtering Method:

Smart E-HRM is basically a recommender system for TM, it suggests and predicts the TM decisions [26][27]. Collaborative filtering method is based on user rating. In this method user opinions such as selection and aggregation are used to predict users inclination [28]. Like if a person likes a particular product because of its advantages so the user will also like the similar products with same specialties. So this method user similarity pattern to predict users opinion. The correlation between i and j is represented as similarity value as shown in equation (1), while the

prediction value can be calculated with the help of equation (2), while equation (3) gives the threshold value for baseline prediction for mean μ [28].

$$s(i,j) = \frac{\sum_{u \in U} (C_{u,i} - \overline{R}_u)(C_{u,i} - \overline{R}_u)}{\sqrt{\sum_{u \in U} (C_{u,i} - \overline{R}_u)^2} \sqrt{\sum_{u \in U} (C_{u,i} - \overline{R}_u)^2}} \quad (1)$$

$$p(u,t) = \frac{\sum_{u \in U} s(t,i)(R_{u,i} - b_{u,i})}{\sum_{u \in U} s(t,i)} + b_{u,i} \quad (2)$$

$$b_{u,i} = \mu + b_u + b_i \quad (3)$$

f) Smart TM Decision Making:

According to the required TM job the smart E-HRM will give the final output as smart decision in order to give us the required candidate for the required job, give recommendation for employees promotion, employees assignments and many HRM decision which are made through humans, which may also contain favoritism so these types recommendations will be purely on performance, educational and experience based thus removing human errors and politics.

Based on the smart e-HRM model for TM, we have used dependent and independent variables for our proposed scheme ensuring effective implementation of our work in a large scale organization such as mentioned above.

Data Analysis

a. Independent Variable:

As per literature review, the research regarding implementation of smart e-HRM systems following five independent variables are identified as shown below. These five independent variables / functions are basically defining the main tasks needed to implement this system in any organization using BD environment.

1. Hiring Competent and qualified employees to run this type of system.
2. Making smart e-HRM tasks easier by using big data and related software's.
3. Processing employee's information to make smart decisions and increasing efficiency.
4. Making Stakeholders communication easier by using this smart system for communication.
5. Performance evaluation based on this smart e-HRM system using big data.
6.

b. Dependent Variables:

The research work is Being evaluated by using 2 dimensions i.e. dimension D1, seen as successful and D2 being denoted by unsuccessful in case for the implementation of this system. These two dimensions are incorporated with 4 elements as under.

1. Smart TM system financial benefits
2. Smart TM efficiency in terms of work done
3. Employees and Organization Benefits
4. Employees satisfaction
5. Loyalty to the organization

The above five independent variable have a direct relation to the dependent variables i.e. outcomes in implementing smart e-HRM system for TM.

Hypothesis

On the basis of the above research framework, this study hypothesized that using BD technology for TM for smart e-HRM system is a function of analytical practice based on human intervention of selected elements in HRM. The research hypothesis came up with the fact that quality of practice for each element in smart e-HRM system is closely related to the TM / HRM and correlated to final results in implementing this system. In order to present the hypothesis a conventional framework must be developed for adoption of smart e-HRM system for TM.
The research work depicts hypotheses as;

H1. Assigning this smart system to technical person having technical as well as practical knowledge of BD system, chances of implementation of this system will increase.
H2. Making TM easy and distributed in a rational manner through software's designed especially for the employee's benefits, thus the employees are assigned equal tasks and efficiency will increase.

H3. Using BD technology to process employee's information, the tasks given to the employees according to their qualifications and each employee is given an equal amount of work load thus increasing successful completion of tasks by each employee.

H4. Implementation of BD technology the stakeholder Communication Management become betters, greater are the chances for successful completion of the tasks.

H5. Performance monitoring of the employees by using BD technology for promotion and rewards which in

turns increases competition among the employees and thus making work environment more competitive.

In short, the hypothesis suggests that by implementing BD technology for TM will increase efficiency and reduce cost while not implementing this system reduces efficiency and icreases the costs for the HR department. The study is co-relational and carried out on BD technology and its implementation in TM. Management staff carrying out different TM practices using BD technology in a telecom organization were studied and analyzed. It should be noted that BD technology is mostly implemented in telecom organization as this type of technology is very useful in employees management and also improving day to day tasks like gathering information from all the offices and servers within the organization and evaluating performance of the employees by processing data through data mining.

Research Involvement and sample population:

It was very strictly implemented that the input of data from the respondents must be unbiased, thus allowing the respondents to freely express their views without interfering in their views or implementing our own views thus giving us a clear and unique perspective in to their minds.

The sample population was selection from 5 different offices from all around the country with (BD intensive) professionals, supervising and management projects related to BD technology of varied scope and nature. The analysis was made on the factors given below.

- Large scale BD intensive projects for TM (Heterogeneous)
- Timely Implementation of BD project for TM.
- Building Team for BD implementation project for TM
- Documentations of BD implementation project for TM
- Organization willingness to implement BD technology to TM
- Availability of 4 members during the time of data collection
- The effect on the quality of TM projects is dependent on the implementation of BD implementation on TM

Data Analysis Results:

Data was collected from 5 different offices within the country from 740 employees. This data was then analyzed statistically and then hypothesis testing was carried out towards final results and discussion. SPSS was used to elaborate the impact of BD technology on TM by analyzing the impact of independent variable son the dependent ones. The analysis will result in achieving the conclusion as how to implement BD technology for TM and how to improve it. Significant level variables and their relationship were analyzed. The following table will give details for the independent variables according to analysis.

Table 1: Independent Variables

Hiring qualified and skilled person for the job, (IV-1)	Handling of this system by qualified and skilled professionals in order to make this system successful
Using specialized software's for smart TM system (IV-2)	Specialized software tools for the employees to that each task is made easy and efficient.
Employees information processing for equal work load distribution (IV-3)	Each employee is given equal amount of task and work load as per qualification thus increasing the chances of successful completion of the tasks given to each employee.
Stakeholder communication (IV-4)	Using BD technology the communication between the stakeholder and employees becomes better and more reliable
Performance evaluations smart system(IV-5)	Performance Monitoring through BD technology helps to boosts employee moral by performance based promotions and rewards thus increasing employee's efficiency and create a healthy competitive work environment.
BD technology implementation results Dependent Variable (DV)	These are the results of the outcomes from this smart BD TM system

The research results based on the hypothesis are given below:

H1 = Qualified and Technical Person for the job: In order to implement BD technology for smart TM system, highly qualified and trained individuals should be hired to operate and run this system. Thus from table 6 we see that Pearson coefficient having value of (β) .179 with P < 0.0. According to our Hypothesis of hiring qualified and technical person for the job, results in successful implementation of BD system for smart TM system. Thus both of our variables are correlated with each other.

H2 = Specialized software's for smart E-HRM system for TM: Implementing specialized tools and software's in order to run BD smart TM system and in

turns increasing efficiency of the organization. From table 6 we see that Pearson coefficient having value of (β) .130 with P < 0.0. This indicates that our hypothesis of implementing software's for BD implementation in smart TM system directly affects the efficacy of the organization, thus both variables are correlating with each other.

H3 = Employees information processing for equal work load distribution: Accessing employees information and qualification through BD technology and distributing workload and tasks equally among the employees increases efficiency in completion of tasks in time. Thus From table 6 we see that Pearson coefficient having value of (β) .231 with P < 0.0 shows that work equal work load distribution among employees directly effects the completion of tasks in time and also increases efficiency. Hence both our variables are correlating with each other.

H4 = Stakeholder communication: Stakeholder communication with the employees directly effects the completion of the tasks by the employees thus Pearson coefficient having value of (β) .462 with P < 0.0 is high thus showing that using BD technology as a tool for communication management between the employees the increases tasks completion time by the employees. Thus our variables are correlating with each other.

H5 = Performance monitoring: Performance monitoring system using BD technology shows strong correlation with Pearson Correlation Pearson coefficient having value of (β) .041 with P < 0.0. Thus using BD technology each employee's performance is monitored online and based on employee's performance rather than favoritism each employee is given promotions and rewards thus making a healthy competitive environment free of favoritism and politics. Thus both of the variables are care correlating with each other.

Table 2: Descriptive Statistics

	Mean	**Std. Deviation**	**N**
Results (DV)	3.4215	.8922	740
Right Person (IV-1)	3.3865	1.2144	740
Specialized Software's / Tools (IV-2)	3.2676	.7692	740
Workload (IV-3)	3.9453	1.0172	740
Stakeholder Communication (IV-4)	3.4866	.8495	740
Performance Monitoring (IV-5)	3.2917	.9996	740

Table 3: Model Summary

Model	***R, Project Implementation***	***R^2***	***Adjusted R square***	***Std. Error***	***Change Statistics(R square change)***	***Change Statistics(R square change***
1	.821[a]	.642	.629	0.622	.51472	24.021

Table 4: Model Summary

Model	***Change Statistics(df1)***	***Change Statistics(df2)***	***Change Statistics (Sig. F Change)***
1	5[a]	60	.000

Table 5: ANOVA[a,b]

Model		***Sum of Squares***	***df***	***Mean Square***	***F***	***Sig***
1	Regression	31.142	6	6.812	22.977	000[c]
	Residual	16.051	58	.247		
	Total	47.022	65			

a.b Predictors: (Constant), Performance Monitoring (IV5), Time Lines (IV2), Stockholder Communication (IV3), Workload (IV4), Right Person (IV1)

Table-6. Coefficients[a,b]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	1.221	.342		3.019	.003
	Right Person (IV1)	.179	.088	.284	1.798	.007
	Software Help (IV2)	.130	.101	.051	.466	.000
	Employees information access (IV3)	.231	.072	.033	.401	.008
	Stakeholders and employees communication (IV4)	.462	.116	.576	4.301	.000
	Performance Monitoring (IV5)	.041	.110	.032	.287	.000

As seen from the above results, the effects of BD technology on TM depends on managing stakeholders communication between the management and the employees and then on the workload on the team members. Success or failure of such type of smart TM system depends on selecting the right person for the right job having technical and relevant knowledge about this smart TM system. All these factors make this smart TM system effective in terms of cost and efficiency. Hence we can further evaluate that the independent and dependent variables having different magnitudes correlate with each other positively.

As seen from the table 3, the value of R^2 for this model relationship is about 0.633 with value of f (24.021) and p is <0.00 which means that there is 65.9 % chance for the adoption of bid data technology for this smart TM system. Similarly (IV-1) is slightly correlated with (IV-2), while (IV-3) and (IV-5) having Pearson correlation of less than or equal to 0.4, thus these two are not correlated to (IV-4). However the independent variables having Pearson correlation of greater than or equal to 0.4 at P = 0.0 shows slight correlation but with the exception of (IV1, IV4). This shows that all the independent variables are slightly correlated with each other. It also shows that by selecting all the variables the implementation of this project becomes more realistic. The correlation of all the five variables shows that this research work will also be tested by using regression analysis.

The correlation between IV and DV for the research work shows that the designed model and theoretical frame work are logical and all the IV's are effective. The findings of the model suggests that this research work is qualified to be tested for regression analysis for further confirmation of the impact of BD technology on smart TM system.

CONCLUSIONS and Future Scope

In this paper we have proposed a model for implementing BD for smart Talent Management system in a large scale telecom organization. BD is slowly but gradually building its grasps in telecom organizations worldwide. Thus in order to implement this system in any organization some gradually changes have to be implemented in order o successfully implement this system. The research was based on five independent

variables which were critical in order to implement BD technology in any large scale organization for smart TM system. The use of BD technology for smart TM system is a worthy technology for the enhancement of the traditional HR. In order to meet the strategic goals for any organization, this type of smart TM is necessary as a tool for understanding competence models, knowledge, skills, behavior and personal characterizes of the employees. Smart TM system with the help of BD technology provides many value added services to the organization, including real time tracking of the employees, their current and future needs, growth plans and employees training and sorting out their problems by diminishing the gap between the management and the employees. Implementing this type of smart TM system will improve consistency and remove bias decisions making thus improving the quality and quantity of decision making.

Today's world is now a BD world. A smart TM system in collaboration with BD technology is capable of analyzing employee's information, effortlessly, efficiently and is more personalized in results. This technology is compatible with all users and can be accessed and modified according to the needs of the organization. The proposed BD technology for smart TM system is more personalized in results and is efficient and cost effective. The proposed smart TM is based on Competence Based smart Model Expert system that is capable of solving wide variety of problems in HR domain. The knowledge base obtains the information by interpreting and reading other sources and can be used by the stakeholders for wide range of purposes. The future work will analyze the efficiency and accuracy of this smart TM system for selecting recruiting, training and retention of its employees through this system.

REFERENCES

1. Kraska, T. (2013). Finding the needle in the big data systems haystack. *IEEE Internet Computing*, *17*(1), 84-86.
2. Iles, P., Preece, D., & Chuai, X. (2010). Talent management as a management fashion in HRD: Towards a research agenda. *Human Resource Development International*, *13*(2), 125-145.
3. Collings, D. G., Scullion, H., & Vaiman, V. (2011). European perspectives on talent management. *European Journal of International Management*, *5*(5), 453-462.
4. Tansley, C. (2011). What do we mean by the term "talent" in talent management?. *Industrial and commercial training*, *43*(5), 266-274.
5. Beechler, S., & Woodward, I. C. (2009). The global "war for talent". *Journal of international management*, *15*(3), 273-285.
6. Cappelli, P. (2008). Talent management for the twenty-first century. *Harvard business review*, *86*(3), 74.
7. Huxley, A. F., & Taylor, R. E. (1958). Local activation of striated muscle fibres. *The Journal of physiology*, *144*(3), 426-441.
8. Bennett, J., & Lanning, S. (2007). Netflix: the Netflix prize. In *KDD Cup and Workshop in conjunction with KDD*.
9. Serrador, P., & Pinto, J. K. (2015). Does Agile work?—A quantitative analysis of agile project success. *International Journal of Project Management*, *33*(5), 1040-1051.
10. Ricci, F., Rokach, L., & Shapira, B. (2011). Introduction to recommender systems handbook. In *Recommender systems handbook* (pp. 1-35). springer US.
11. Weinberger, S. E., Hoyt, D. B., Lawrence, H. C., Levin, S., Henley, D. E., Alden, E. R., ... & Hubbard, W. C. (2015). Firearm-Related Injury and Death in the United States: A Call to Action From 8 Health Professional Organizations and the American Bar AssociationFirearm-Related Injury and Death in the United States. *Annals of internal medicine*, *162*(7), 513-516.
12. Wei, K., Huang, J., & Fu, S. (2007, June). A survey of e-commerce recommender systems. In *Service systems and service management, 2007 international conference on* (pp. 1-5). IEEE.
13. Alonso, J., & Dohler, M. (2014). Machine-to-Machine Technologies & Markets-Shift of Industries. *Tutorial, IEEE WCNC*, *6*.
14. Shroff, G. (2013). *The Intelligent Web: Search, smart algorithms, and big data*. OUP Oxford.
15. Chen, X. W., & Lin, X. (2014). Big data deep learning: challenges and perspectives. *IEEE access*, *2*, 514-525.
16. Dayarathna, M., & Suzumura, T. (2013). Automatic optimization of stream programs via source program operator graph transformations. *Distributed and Parallel Databases*, *31*(4), 543-599.
17. Salinas, S., Chen, X., Ji, J., & Li, P. (2016). A tutorial on secure outsourcing of large-scale computations for big data. *IEEE Access*, *4*, 1406-1416.
18. Zeng, X. Q., & Li, G. Z. (2014). Incremental partial least squares analysis of big streaming data. *Pattern Recognition*, *47*(11), 3726-3735.
19. Hadoop, A. (2016). Map/Reduce Tutorial. *URL https://hadoop. apache. org/docs/r1*, *2*. e
20. Yadav, M. K., & Baria, M. J. (2014). Mini-Batch K-Means Clustering Using Map-Reduce in Hadoop. *International Journal of Computer*

Science and Information Technology Research, *2*(2), 336-342.

21. Borthakur, D. (2007). The hadoop distributed file system: Architecture and design. *Hadoop Project Website*, *11*(2007), 21.
22. Shanahan, J. G., & Dai, L. (2015, August). Large scale distributed data science using apache spark. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 2323-2324). ACM.
23. Shin, J., Wu, S., Wang, F., De Sa, C., Zhang, C., & Ré, C. (2015). Incremental knowledge base construction using deepdive. *Proceedings of the VLDB Endowment*, *8*(11), 1310-1321.
24. Dutt, A., Aghabozrgi, S., Ismail, M. A. B., & Mahroeian, H. (2015). Clustering algorithms applied in educational data mining. *International Journal of Information and Electronics Engineering*, *5*(2), 112.
25. Xia, P., Ribeiro, B., Chen, C., Liu, B., & Towsley, D. (2013, August). A study of user behavior on an online dating site. In *Advances in Social Networks Analysis and Mining (ASONAM), 2013 IEEE/ACM International Conference on* (pp. 243-247). IEEE.
26. Raj, N., & Suja Rani, M. S. (2015). An Overview of Content Recommendation Methods. *International Journal of Innovative Research in Computer and Communication Engineering*, *3*(1), 1.
27. Kaklauskas, A. (2015). Passive House model for quantitative and qualitative analyses and its intelligent system. In *Biometric and Intelligent Decision Making Support* (pp. 87-112). Springer International Publishing.
28. Ekstrand, M. D., Kannan, P., Stemper, J. A., Butler, J. T., Konstan, J. A., & Riedl, J. T. (2010, September). Automatically building research reading lists. In *Proceedings of the fourth ACM conference on Recommender systems* (pp. 159-166). ACM.