Available online at http://www.journalcra.com

INTERNATIONAL JOURNAL OF CURRENT RESEARCH

International Journal of Current Research
Vol. 13, Issue, 03, pp.16473-16477, March, 2021

DOI: https://doi.org/10.24941/ijcr.40944.03.2021

ISSN: 0975-833X

RESEARCH ARTICLE OPEN ACCESS

BIG DATA: A MEASURE TO RESOLVE COMPLEX TRIBAL ISSUES

[1]*Dr. Vikash Kumar Singh and [2]Ms. Bhartee Singh Baghel

[1]Head, Dept of Computer Science, IGNTU, Amarkantak— 484887 (MP)
[2]Research Scholar, Dept of Computer Science, IGNTU, Amarkantak— 484887 (MP)

ARTICLE INFO

Article History:
Received 10th December, 2020
Received in revised form
24th January, 2021
Accepted 19th February, 2021
Published online 17th March, 2021

Key Words:
Big Data, Advanced Technique,
Tribal Development.

ABSTRACT

The present paper endeavors to analyze the role of big data as a potent measure to resolve complex tribal issues and facilities to promote multifaceted development of tribal communities. Big data technically refers to data sets so large and complex that traditional tools like relational databases are unable to process them in an acceptable time frame or within a reasonable cost range problems occur in sourcing, moving, searching, storing and analyzing the data, but with the right tools these problems can be overcome. Big data with the combination of predictive, analytics use / behavior analytics, or certain other advanced data analytical methods can play vital role with respect to the problems of tribal communities which comprise land alienation, poverty, indebtedness, health, nutrition and education. Big data with advanced tools can process the data related to these problems and reveal meaningful information which will definitely make policy formulation and decision making easier and more accurate.

Citation: Dr. Vikash Kumar Singh and Ms. Bhartee Singh Baghel. *"*Big Data: A measure to resolve complex tribal issues*"*, 2021. *International Journal of Current Research, 13, (03), 16473-16477.*

INTRODUCTION

Big Data and its advanced technique can be applied in diverse areas related to tribal communities which are living in the remote corners of our country. Their life, demographic expansion, data related to their census, sex ratio, mortality rate, life expectancy, tribal groups, their socio-cultural practices, their representation in legal and constitutional bodies at different levels etc can be recorded and analyzed for specific purposes in the policy formulation which will definitely help our policy planners to introduce corrective measures to resolve complex demographic issues of the tribals. Tribal education is another important issue which incorporates number of tribals admitted in primary, middle, secondary and college level; dropout rate; ratio of education among tribal boys and girls; ratio among tribal and non tribal communities; admission in traditional and professional courses; causes of low literacy among tribals; analytics can clearly indicate certain concrete results to resolve the complex education related problems.

Corresponding author:** ***Dr. Vikas Kumar Singh,
Head, Dept of Computer Science, IGNTU, Amarkantak— 484887 (MP).

Tribal land holding and dependency on it is another crucial subject which directly affects the life of the tribal community. Size of the land holding per person, the engagement of the tribal people in a specific land holding, under-employment situation, measures for land consolidation etc all data related to these can be stored and analyzed and thereafter desired results can be attained. Poverty and indebtedness is another acute problem tribals withstand throughout their lives. Data related to tribals living below poverty line, per capita income, indebtedness, sources of income, land holding and income out of these holdings, agricultural products and their salability, forest products and live stocketc if processed through big data tools can reveal meaningful information which can be used in developmental schemes exclusively meant for tribal communities. Health and nutrition is another core area which needs urgent address; mostly tribal communities are suffering from malnutrition and certain infectious diseases which wipe out the entire community. For health related programmes such as immunization, vaccination, free health services and maternity safeguards, family planning etc complete data and its proper analysis is imminently required which is only possible through Big Data and its advanced tools and techniques. The use and adoption of big data within governmental processes allows efficiencies in terms of cost, productivity, and innovation, but does not come without its flaws. Data analysis often requires multiple parts of government (central and local) to work in collaboration and create new and innovative processes to deliver the desired outcome.

Big Data: Conceptualizing the Advanced Tools and Techniques: Big Data is a term for data sets that are so large or complex that traditional data processing application software is inadequate to deal with them. Challenges include capture storage analysis data duration search sharing transfer visualization querying updating and information privacy. The term has been in use since the 1990s, with some giving credit to John Mashey for coining or at least making it popular. Big data usually includes data sets with sizes beyond the ability of commonly used software tools to capture, curate, manage, and process data within a tolerable elapsed time. Big Data philosophy encompasses unstructured, semi-structured and structured data, however the main focus is on unstructured data. Big data "size" is a constantly moving target, as of 2012 ranging from a few dozen terabytes to many petabytes of data. Big data requires a set of techniques and technologies with new forms of integration to reveal insights from datasets that are diverse, complex, and of a massive scale. Big Data is a phrase used to mean a massive volume of both structured and unstructured data that is so large and it is difficult to process using traditional database and software techniques. In most enterprise scenarios the volume of data is too big or it moves too fast or it exceeds current processing capacity. Big Data has the potential to help companies and government departments to improve operations and make faster, more intelligent decisions. This data, when captured, formatted, manipulated, stored, and analyzed can help a company to gain useful insight to increase revenues, get or retain customers, and improve operations. While the term may seem to reference the volume of data, that isn't always the case. The term Big Data, especially when used by vendors, may refer to the technology (which includes tools and processes) that an organization requires handling the large amounts of data and storage facilities. The term is believed to have originated with web search companies which needed to query very large distributed aggregations of loosely-structured data. An example of Big Data might be petabytes (1,024 terabytes) or exabytes (1,024 petabytes) of data consisting of billions to trillions of records of millions of people—all from different sources (e.g. Web, sales, customer contact center, social media, mobile data and so on). The data is typically loosely structured data that is often incomplete and inaccessible.

Big Data: Its Limitations

Prioritizing correlations: Data analysts use big data to tease out correlation: when one variable is linked to another. However, not all these correlations are substantial or meaningful. More specifically, just because 2 variables are correlated or linked doesn't mean that a causative relationship exists between them (i.e. "correlation does not imply causation"). For instance, between 2000 and 2009, the number of divorces in the U.S. state of Maine and the per capita consumption of margarine both similarly decreased. However, margarine and divorce have little to do with each other. A good consultant will help you figure out which correlations mean something to your business and which correlations mean little to your business.

The Wrong Questions: Big data can be used to discern correlations and insights using an endless array of questions. However, it's up to the user to figure out which questions are meaningful. If you end up getting a right answer to the wrong question, you do yourself, your clients, and your business, a costly disservice.

Security: As with many technological endeavors, big data analytics is prone to data breach. The information that you provide a third party could get leaked to customers or competitors.

Transferability: Because much of the data you need analyzed lies behind a firewall or on a private cloud, it takes technical know-how to efficiently get this data to an analytics team. Furthermore, it may be difficult to consistently transfer data to specialists for repeat analysis.

Inconsistency in data collection: Sometimes the tools we use to gather big data sets are imprecise. For example, Google is famous for its tweaks and updates that change the search experience in countless ways; the results of a search on one day will likely be different from those on another day. If you were using Google search to generate data sets, and these data sets changed often, then the correlations you derive would change, too. Ultimately, you need to know how to use big data to your advantage in order for it to be useful. The use of big data analytics is akin to using any other complex and powerful tool. For instance, an electron microscope is a powerful tool, too, but it's useless if you know little about how it works.

BIG DATA versus Relational Database: "Big data" is a buzzword, but it maps to a class of use cases that are typically thought of as involving data mining and analytics. Note that while some big data applications do involve huge amounts of data, many of them do not - "big data" is more about how the data is searched and analyzed versus the size of the data itself. In the context of this question, "big data" is a property of the data and the application, and makes no reference to the software tools used to analyze the data. Note that not all applications involving huge amounts of data are "big data" applications; a crime report mining app used by a police department is a big data app even if the crime report db is only a few hundred megabytes, while the Visa credit card debit and approval system is an Online app - and not a big data app - even if the db involved is large numbers of terabytes. "RDBMS" refers to Relational database engines.

They are often used with big data applications, although some big data apps use NoSql engines, others use both, and many applications don't use db engines at all. The volume of data in the world is increasing at a staggering rate, as industries today are collecting, storing, analyzing, sharing, collaborating more, and requiring more of data from an ever-expanding variety of information sources. Organizations are discovering new methods to leverage big data, generating greater customer and business insight, and improving current operational processes. A powerful and scalable big data integration gateway to securely, swiftly, and reliably move big data between data lakes and analytics applications is a necessity for forward-thinking companies. With the Cleo Integration Suite's big data movement solutions, organizations can seamlessly flow information to varied big data applications, reliably ensuring business value and service continuity. Driven by specialized analytics systems and software, big data analytics can point the way to various business benefits, including new revenue opportunities, more effective marketing, better customer service, improved operational efficiency and competitive advantages over rivals. Big data analytics applications enable data scientist predictive modelers, statisticians and other analytics professionals to analyze growing volumes of structured transaction data, plus other forms of data that are often left untapped by conventional business intelligence (BI) and analytics programs. That encompasses a mix of SEMI STRUCTURED and UNSTURCTURED DATA -- for example, internet clickstream data, web server logs, social media content, text from customer emails and survey responses, mobile-phone call-detail records and machine data captured by sensors connected to the internet of things.

On a broad scale, data analytic technologies and techniques provide a means of analyzing data sets and drawing conclusions about them to help organizations make informed business decisions. BI queries answer basic questions about business operations and performance. Big data analytics is a form of advanced analytic, which involves complex applications with elements such as predictive models, statistical algorithms and what-if analyses powered by high-performance analytics systems. Not only scientific institutes but also companies harvest an amazing amount of data. Traditional database management systems are often unable to cope with this. Suitable tools are lacking in information retrieval on big data. Computer scientists have now developed an approach which enables searching large amounts of data in a fast and efficient way. The beauty of such Big Data applications is that they can process Web-based text, digital images, and online video. They can also glean intelligence from the exploding social media sphere, whether it consists of blogs, chat forums, Twitter trends, or Facebook commentary. Traditional market research generally involves unnatural acts, such as surveys, mall-intercept interviews, and focus groups. Big Data examines what people say about what they have done or will do. That’s in addition to tracking what people are actually doing about everything from crime to weather to shopping to brands. It is only Big Data’s capacity for dealing with vast quantities of real-time unstructured data that makes this possible.

For example, retailers like Wal-Mart and Kohl’s are making use of sales, pricing, and economic data, combined with demographic and weather data, to fine-tune merchandising store by store and anticipate appropriate timing of store sales. Similarly, online data services like eHarmony and Match.com are constantly observing activity on their sites to optimize their matching algorithms to predict who will hit it off with whom. The same logic is being applied to economic forecasting. For example, the number of Google queries about housing and real estate from one quarter to the next turns out to predict more accurately what’s going to happen in the housing market than any team of expert real estate forecasters. Similarly, Google search queries on flu symptoms and treatments reveal weeks in advance what flu-related volumes hospital emergency departments sensors — what GE people like CMO Beth Comstock called “machine whispering” when I talked with her this past summer — are creating a second tsunami of data. Digital sensors on industrial hardware like aircraft engines, electric turbines, automobiles, consumer packaged goods, and shipping crates can communicate “location, movement, vibration, temperature, humidity, and even chemical changes in the air.” As the volume of both human and machine data grows exponentially, so too will organizations’ ability to see the future.

Migration: It doesn't matter if data is being migrated from SQL to NoSQL, from flat files to key-value store, or from XML to an object database, or every permutation of any data store to any other data store. What stays constant is the fact that data migrations are scary and painful. Without the right strategy, a big data migration will leave you with inconsistent data, strange errors, and very angry users.

Backup: Before even considering a massive destructive mutation of your data, you should have working backups. The keyword is "working". Take production dumps of your data, and make sure you can load the same data on a cloned environment. If anything goes wrong when migration day comes along, these backups will be your last line of defense. Backups are also useful for doing practice runs of a migration.

Logging: Create a logger that logs to a separate place from your application logs that's specific for the migration. When the migration is running, the logger should warn on strange data and error on exceptional cases. To keep the log useful, it's important not to flood it with debugging information. Log only the most important details needed for troubleshooting problems: a timestamp, an id reference to the failing record, and a brief description of the failure reason.

Atomicity: Regardless of whether the destination data store supports transactions or not, the migration should always define an invariant for when a record is successfully imported. If this invariant is broken, then whatever has been done to break the invariant should be undone so that your data isn't in some zombie half consistent state.

Idempotence: Not strictly the definition of idempotence, but similar to maintaining consistency, your code should be able to handle re-migrating the same data. If the migration crashes halfway, having this property allows you restart and import again without worrying about weird state issues.

Batch Processing: Having atomicity and idempotence lets your migration be split up into smaller migrations. Instead of migrating a million records in an all-or-nothing migration and crossing your fingers, you can split them up into small 500 record batches. If any single batch fails, you can redo just that single batch, rather than redo the entire migration. This also allows you to balance the migration across more resources like multiprocessors, different servers, and different slave databases.

Validation: After a migration is complete, it's important to be able to validate that everything is still working. This means running your test suite, your integration tests, and also logging in as existing users and clicking around.

Live Migrations: Running a migration with scheduled downtime is hard enough as it is, but in certain applications, a big chunk of downtime is unacceptable. If this is the case, then it's critical to add bookkeeping code that tracks which records has been migrated and which haven't. This allows you to query and incrementally upgrade parts of your system while co-existing with old data and old code.

Plan Ahead: Data migrations will always be a chore. But with the right strategy, at least it'll be one that can be finished, rather than something that drags along and repeatedly slows down your whole team.

Big Data and Health Management: **An Instance of Practicability of Big Data:** Big Data has changed the way we manage, analyze and leverage data in any industry or government department. One of the most promising areas where big data can be applied to make a change is healthcare. Healthcare analytics have the potential to reduce costs of treatment, predict outbreaks of epidemics, avoid preventable diseases and improve the quality of life in general. Average human lifespan is increasing along world population, which poses new challenges to today's treatment delivery methods. Healthcare professionals, just like business entrepreneurs, are capable of collecting massive amounts of data and look for best strategies to use these numbers. The issues such as the need of big data in healthcare, its practical uses, the obstacles to its adoption and finally citing examples from organizations/ government departments which are already getting benefits from its use in healthcare sector.

Why Big Data so Indispensible in Health CARE?: There's a huge need for big data in healthcare as well, due to rising costs in nations like the United States. As a McKinsey report states, "After more than 20 years of steady increases, healthcare expenses now represent 17.6 percent of GDP —nearly $600 billion more than the expected benchmark for a nation of the United States's size and wealth." In other words, healthcare costs are much higher than they should be, and they have been rising for the past 20 years. Clearly, we are in need of some smart, data-driven thinking in this area. And current incentives are changing as well: many insurance companies are switching from fee-for-service plans (which reward using expensive and sometimes unnecessary treatments and treating large amounts of patients quickly) to plans that prioritize patient outcomes. As the authors of the popular Freakonomics books have argued, financial incentives matter – and incentives that prioritize patients health over treating large amounts of patients are a good thing. Why does this matter for big data? In the previous scheme, healthcare providers had no direct incentive to share patient information with one another, which had made it harder to utilize the power of big data. Now that more of them are getting paid based on patient outcomes, they have a financial incentive to share data that can be used to improve the lives of patients while cutting costs for insurance companies. Finally, physician decisions are becoming more and more evidence-based, meaning that they rely on large swathes of research and clinical data as opposed to solely their schooling and professional opinion. As in many other industries, data gathering and management is getting bigger, and professionals need help in the matter. This new treatment attitude means there is a greater demand for big data analytics in healthcare facilities than ever before, and the rise of SaaS business intelligence tools is also answering that need.

Obstacles To Data-Driven Healthcare: One of the biggest hurdles standing in the way to use big data in healthcare is how medical data is spread across many sources governed by different states, hospitals, and administrative departments. Integration of these data sources would require developing a new infrastructure where all data providers collaborate with each other. Equally important is implementing new data analysis tools and strategies. Healthcare needs to catch up with other industries that have already moved from standard regression-based methods to more future-oriented like predictive analytics, machine learning, and graph analytics. However, there are some glorious instances where healthcare doesn't lag behind, such as EHRs So, even if these services are not a cup of tea, you are a potential patient, and so you should care about new healthcare analytics applications. Besides, it's good to take a look around sometimes and see how other industries cope with big data. They can inspire you to adapt and adopt some good ideas.

Instances of Uses of Big data in healthcare sector:

Electronic Health Records (EHRs): It's the most widespread application of big data in healthcare. Every patient has his own digital record which includes demographics, medical history, allergies, laboratory test results etc. Records are shared via secure information systems and are available for healthcare providers from both public and private sector. Every record is comprised of one modifiable file, which means that doctors can implement changes over time with no paperwork and no danger of data replication.

Real-time Alerting: Other examples of big data in healthcare share one crucial functionality – real-time alerting. In hospitals, Clinical Decision Support (CDS) software analyzes medical data on the spot, providing health practitioners with advice as they make prescriptive decisions. However, doctors want patients to stay away from hospitals to avoid costly in-house treatments. Personal analytics devices, already trending as business intelligence buzzwords last year, have the potential to become part of a new healthcare delivery strategy. Wearables will collect patients' health data continuously and send this data to the cloud. Additionally, this information will be accessed to the database on the state of health of the general public, which will allow doctors to compare this data in socioeconomic context and modify the delivery strategies accordingly. Healthcare institutions and care managers will use sophisticated tools to monitor this massive data stream and react every time the results will be disturbing.

Predictive Analytics in Healthcare: We have already recognized predictive analytics as one of the biggest business intelligence trend two years in a row, but the potential applications reach far beyond business and much further in the future. Optum Labs, an US research collaborative, has collected EHRs of over 30 million patients to create a database for predictive analytics tools that will improve the delivery of care. The goal is to help doctors make big data-informed decisions within seconds and improve patients' treatment. This is particularly useful in case of patients with complex medical histories, suffering from multiple conditions. New tools would also be able to predict, for example, who is at risk of diabetes, and thereby be advised to make use of additional screenings or weight management.

Using Health Data For Informed Strategic Planning: The use of big data in healthcare allows for strategic planning thanks to better insights into people's motivations. Care mangers can analyze check-up results among people in different demographic groups and identify what factors discourage people from taking up treatment. University of Florida made use of Google Maps and free public health data to prepare heat maps targeted at multiple issues, such as population growth and chronic diseases. Subsequently, academics compared this data with the availability of medical services in most heated areas. The insights gleaned from this allowed them to review their healthcare delivery strategy and add more care units to most problematic areas

Big Data Just Might Cure Cancer: Another interesting example of the use of big data in healthcare is the Cancer

Moonshot program. Before the end of his second term, President Obama came up with this program that had the goal of accomplishing 10 years worth of progress towards curing cancer in half that time. This bold goal led to a panel which gave recommendations that included some big data use cases. Medical researchers can use large amounts of data on treatment plans and recovery rates of cancer patients in order to find trends and treatments that have the highest rates of success in the real world. For example, researchers can examine tumor samples in biobanks that are linked up with patient treatment records. Using this data, researchers can see things like how certain mutations and cancer proteins interact with different treatments and find trends that will lead to better patient outcomes.

Telemedicine: Telemedicine has been present on the healthcare services market for over 40 years, but only today, with the arrival of online video conferences, smartphones, wireless devices, and wearables, has it been able to come into full bloom. The term refers to delivery of remote clinical services using technology. It is used for primary consultations and initial diagnosis, remote patient monitoring, and medical education for health professionals. Some more specific uses include telesurgery – doctors can perform operations with the use of robots and high-speed real-time data delivery without physically being in the same location with a patient.

Big Data Is Helping To Prevent Unnecessary ER Visits: Saving time, money and energy using big data analytics in healthcare is necessary. What if we told you that over the course of 3 years, one woman visited the ER more than 900 times? That situation is a reality in Oakland, California, where a woman who suffers from mental illness and substance abuse went to a variety of local hospitals on an almost daily basis.

To sum up, the role of Big Data in diverse areas has increased multiple times especially in healthcare, nutrition, education, poverty eradication, land alienation, electioneering campaign, industry management, defense sector, civil amenities etc, and in finalizing social welfare policies. Simultaneously, these advanced tools and techniques can be dexterously applied in resolving complex tribal issues and problems. One instance of the application of Big Data, briefly analyzed in the present paper is healthcare which is beneficial to all either living in the urban areas or rural pockets of India. In other areas too, Big Data is helping positively.

REFERENCES

Judith, Hurwitz et al. 2013. Big Data for Dummies. John Willey& Sons.
Marz, Nathan & James Warren. 2015. Big Data: Principles and Best Practices of Scalable Real-Time Data Systems. Dream-tech Press.
DT Editorial Services. 2015. Big Data: Black Book. Dream-tech Press.
Frank, Bill. 2012. Taming the Big Data Tidal Wave. John Wiley & Sons.
Mayor, Voctor et al., 2013. Big Data: A Revolution That Will Transform How We Live, Work, and Think. (HB) John Murray.
Shankarmani, Radha & M Vijayalakshmi. 2016. Big Data Analytics. 2nd Edition, Wiley India Pvt. Ltd.
Chen, Min et al., 2014. Big Data: Related Technologies, Challenges and Future Prospects. Springer International.
Madsen, Laura B. 2015. Data-Driven Healthcare: How Analytics and BI are Transforming the Industry.wiley India Pvt. Ltd.
http://www.healthcatalyst.com/bigdata-n-healthcare-made-simple
www.datapine.com/blog/bigdata examples-n-healthcare
http://mapr.com/blog/5-bigdata-trends-healthcare-2017
http://www.sas.com>SAS Insights>Big Data
attunelive.com/big-data applications-healthcare

****