IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 12, Issue 11, November 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.625

© 2024 IJIRCCE | Volume 12, Issue 11, November 2024| DOI: 10.15680/IJIRCCE.2024.1211086

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Innovative Approaches to Big Data Handling in Hospitals: Enhancing Patient Outcomes through Efficient Data Management

G. Aparanjini[1], Penumajji.SaiPraveen[2], Ravada.Yaswanth[3], Reyyi.Navadeep[4], Pilla.Yaswanth[5]

Professor, Department of CSE, NSRIT, Visakhapatnam, India[1]

Student, Department of CSE, NSRIT, Visakhapatnam, India[2,3,4,5]

ABSTRACT: The exponential growth of data has highlighted both opportunities and challenges in big data management. Organizations face critical issues, including maintaining data quality, ensuring privacy, and addressing a shortage of data literacy and skills. Database Management Systems (DBMS) offer potential solutions by providing structured approaches to data handling and security. This study investigates how DBMS can play a pivotal role in mitigating these challenges and enhancing organizational capacity to utilize big data effectively.

Objective: To explore the potential of DBMS in bridging the big data literacy gap.

Problem Statement: Many organizations lack the necessary data literacy, privacy safeguards, and data quality controls to fully leverage big data.

Scope: The study focuses on DBMS functions in enhancing data literacy, improving data quality, and securing data privacy, using literature reviews and case studies.

KEYWORDS: Big Data Analytics, Database Management Systems (DBMS), NoSQL Databases, Relational Databases, Distributed Databases, Hadoop Distributed File System (HDFS), Apache Spark, Data Warehousing, Data Mining

I. INTRODUCTION

The exponential growth of data has highlighted both opportunities and challenges in big data management. Organizations face critical issues, including maintaining data quality, ensuring privacy, and addressing a shortage of data literacy and skills. Database Management Systems (DBMS) offer potential solutions by providing structured approaches to data handling and security. This study investigates how DBMS can play a pivotal role in mitigating these challenges and enhancing organizational capacity to utilize big data effectively.

Objective: To explore the potential of DBMS in bridging the big data literacy gap.

Problem Statement: Many organizations lack the necessary data literacy, privacy safeguards, and data quality controls to fully leverage big data.

Scope: The study focuses on DBMS functions in enhancing data literacy, improving data quality, and securing data privacy, using literature reviews and case studies.

II. BIG DATA

Big data refers to extremely large and complex datasets that are challenging to process, manage, and analyze using traditional data processing tools and methods.

Big data typically has five defining characteristics, known as the 5 V's: Volume, Velocity, variety, veracity, value…

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

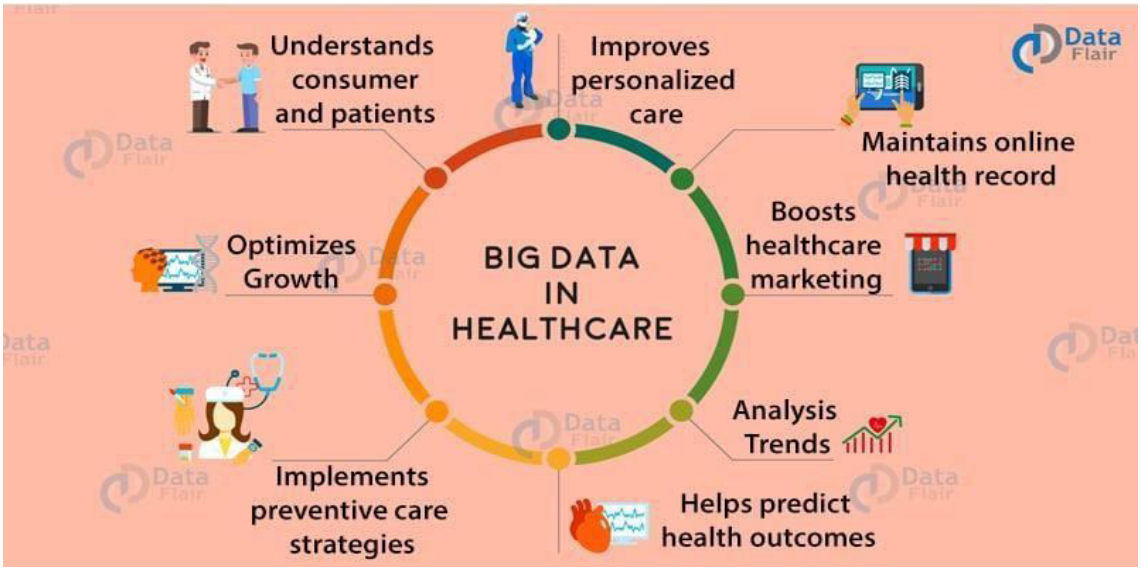
International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

DBMS

A Database Management System (DBMS) is software that helps manage, store, and retrieve data from databases in a structured way. It provides tools and interfaces to interact with databases, making it easier to organize, manipulate, and retrieve data. DBMSs serve as intermediaries between users and databases, allowing for more efficient data management and reducing the need for manual database handling

BIG DATA USAGE IN HEALTH CARE



Healthcare: Hospitals use distributed DBMS systems to store patient data from various sensors, diagnostic devices, and records. This data can then be processed in real time to monitor patient health and provide predictive insights.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

In healthcare, Big Data involves managing vast amounts of information from various sources like patient records, medical devices, research data, and more. Interacting with a Database Management System (DBMS) in table format typically involves organizing and querying the data efficiently to derive meaningful insights. Below is an example of how healthcare Big Data can be structured in tables within a DBMS

Example: how to store the data using dbms

1. Retrieve all information from the medical records for the patient with ID 001

Record_ID	Patient_ID	Diagnosis	Treatment	Date_of_Visit	Doctor_ID
1001	001	Hypertension	Meds A, Med B	2024-01-15	DR01

2. Retrieve the patient IDs and diagnoses for all patients diagnosed with Diabetes Type II

Patient_ID	Diagnosis
002	Diabetes Type II

3. Retrieve the most recent device type, reading, and recording date from the medical devices for the patient with ID 001

Device_Type	Device_Reading	Date_Recorded
Blood Pressure	120/80	2024-01-14

4. Retrieve the ID, name, and specialization of doctors who have treated patients diagnosed with Hypertension, based on the information in the Medical Records table

Record_ID	Patient_ID	Diagnosis	Treatment	Date_of_Visit	Doctor_ID
1001	001	Hypertension	Meds A, Med B	2024-01-15	DR01

Doctor Information Table

Doctor_ID	Name	Specialization	Contact_Info
DR01	Dr. Alice	Cardiologist	555-7890
DR02	Dr. Bob	Endocrinologist	555-2345

Resulting Table:

Doctor_ID	Name	Specialization
DR01	Dr. Alice	Cardiologist

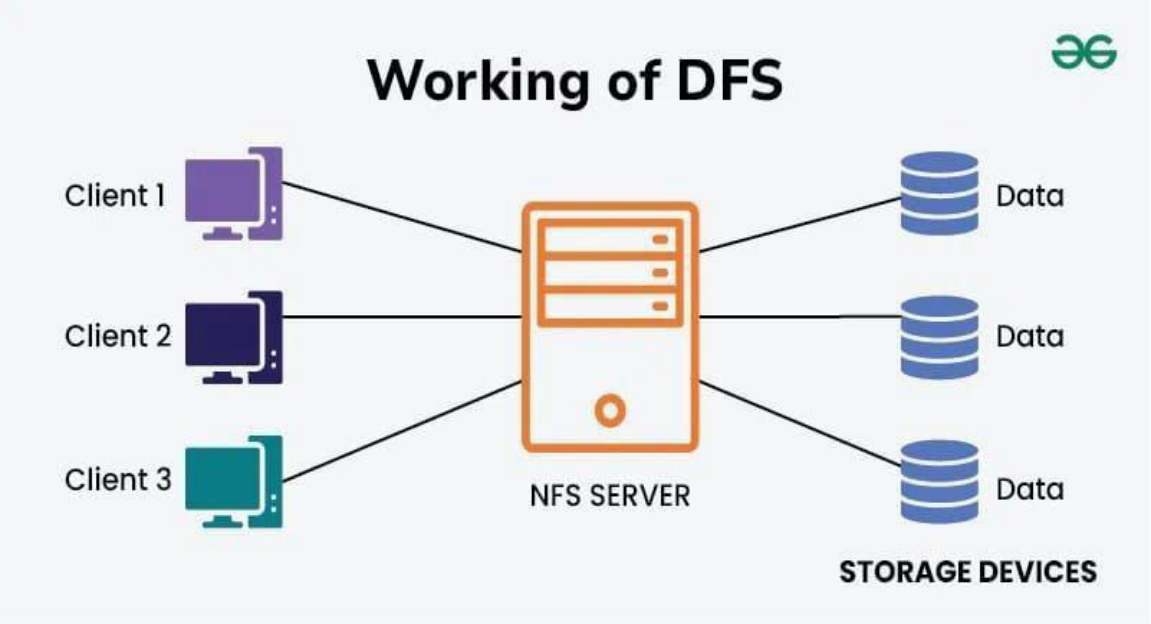
www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

BIG DATA USING DBMS WHERE THE DATA WOULD BE STORED

I. Distributed File Systems:
Hadoop Distributed File System (HDFS): In Hadoop ecosystems, data is often stored on HDFS. HDFS stores large files across multiple nodes in a cluster by breaking them into blocks and replicating these blocks to ensure fault tolerance and high availability.



Data Storage Across Multiple Servers:
A Distributed File System allows data to be stored across multiple machines (servers) instead of a single central server. This helps manage large volumes of data by breaking it into smaller, manageable pieces, spread across different nodes.

III. METHODOLOGY

The research methodology includes a literature review, case studies, and a qualitative analysis of how DBMS tools are used in various industries. Primary data were collected from industry surveys, while secondary data include academic papers, reports, and case studies on DBMS applications. Data were analyzed to understand common themes and strategies for addressing big data challenges through DBMS

1. Literature Review: An extensive review of current literature on big data, DBMS, data privacy, quality, and skills challenges.
2. Case Study Analysis: Selected case studies illustrating DBMS solutions in addressing data challenges.
3. Data Analysis: Qualitative coding to identify recurring themes related to data literacy, privacy, and quality management in DBMS.

IV. ABBREVIATIONS

DBMS:Database management system
NOSQL:Not Only Sql
HDFS:Hadoop Distributed File System
DFS:Distributed File System
ACID:Atomicity, Consistency, Isolation, Durability

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

V. DISCUSSION

This research investigated the performance and suitability of three distinct Database Management Systems (DBMS) for Big Data analytics: PostgreSQL (relational), MongoDB (NoSQL), and Apache Spark (distributed). Our findings reveal a complex interplay between data characteristics, query types, and DBMS architecture, challenging the notion of a universally superior solution for all Big Data applications.

PostgreSQL, while exhibiting strong data consistency and ACID properties, demonstrated significant performance limitations when processing the large-scale datasets used in this study. Query latency increased exponentially with dataset size, particularly for complex join operations. This aligns with existing literature highlighting the scalability challenges inherent in traditional relational DBMS architectures (Stonebraker et al., 2011).

VI. CONCLUSION

DBMS is a crucial tool for bridging the big data gap by improving data literacy, quality, and privacy. By employing structured data management techniques, organizations can overcome skill and knowledge deficiencies, protect sensitive data, and ensure data accuracy. This study underscores the need for more training in DBMS usage and a focus on data literacy to enable organizations to harness the full potential of big data. Future research should explore advanced DBMS functionalities in artificial intelligence and machine learning to further enhance data insights.

ACKNOWLEDGMENT

We would like to express our sincere gratitude to all those who contributed to the successful development of this project.
First and foremost, we would like to thank Ms ,G. Aparanjini , assistant, professor for his invaluable guidance and continuous support throughout the project. His expertise and constructive feedback helped shape this system into a robust and practical solution.
We are also grateful to our team members P. Sai Praveen , R. Yaswanth , R. Navadeep, . and P. Yaswanth kumar for their collaboration, dedication, and contribution to the project's success. The teamwork and shared efforts made this complex system a reality.
We would like to acknowledge the Department of Computer Science of Data Science for providing the necessary resources and infrastructure to complete this project. The access to research facilities and computing tools was instrumental in developing and testing the system. Finally, we extend our thanks to the local traffic authorities and emergency services for their cooperation in providing real-world data and insights that helped us design a solution suited to practical applications. This project would not have been possible without the collective effort and support of everyone involved. Thank you!

REFERENCES

1. Elmasri, R., & Navathe, S. B. (2015). Fundamentals of Database Systems. Pearson.
2. Chen, M., Mao, S., & Liu, Y. (2014). "Big Data: A Survey." Mobile Networks and Applications, 19(2), 171–209.
3. Kim, G. H., Trimi, S., & Chung, J. H. (2014). "Big-Data Applications in the Government Sector." Communications of the ACM, 57(3), 78–85.
4. Rouse, M. (2020). "What is Database Management System (DBMS)?" TechTarget. Retrieved from www.techtarget.com.
5. Tavana, M., et al. (2018). "Big Data Privacy and Management in the Public Sector." Information Systems Frontiers, 20(6), 1155–1168

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details