

# BIG DATA CHALLENGES

By Nasser Thabet & Tariq Rahim Soomro

Presented by:

Uzma Naeem



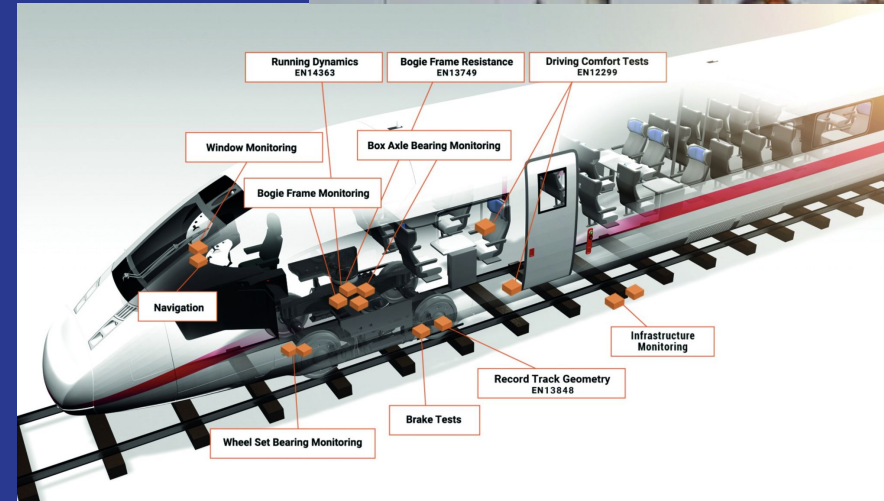
# Introduction to Big Data Challenges

- Brief Overview: Definition of Big Data
- Importance of addressing Big Data challenges
- Two fundamental questions:
  - What is Big Data?
  - What are its characteristics?



# Evolution of Technology & Big Data Era

- Living in the Big Data Era:
  - Rapid evolution of technology
  - Instrumentation & sensing everything around us
  - Interconnectivity of people & things
  - Examples of technology impact: sensors in transportation, intelligence in railways, & weather sensors on roads



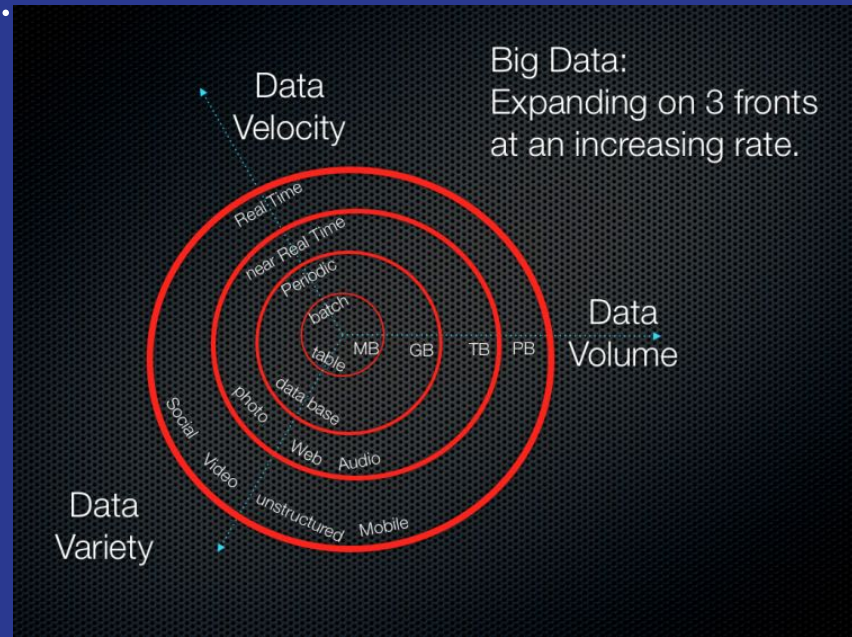
# Three Vs & Advantages of Big Data Technology

- Three Vs & Business Values:

- Initial characterization: Volume, Variety, Velocity

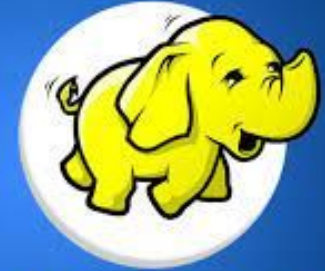
- Advantages of Big Data Technology:

- Cost Reduction,
    - Decision-Making Improvement,
    - Enhancements in Products & Services



# Examples of Big Data Utilization

- Real-world Examples - Case Studies:
  - Caesars' use of Hadoop clusters for real-time decision-making
  - LinkedIn's People You May Know (PYMK) feature as a successful product derived from Big Data

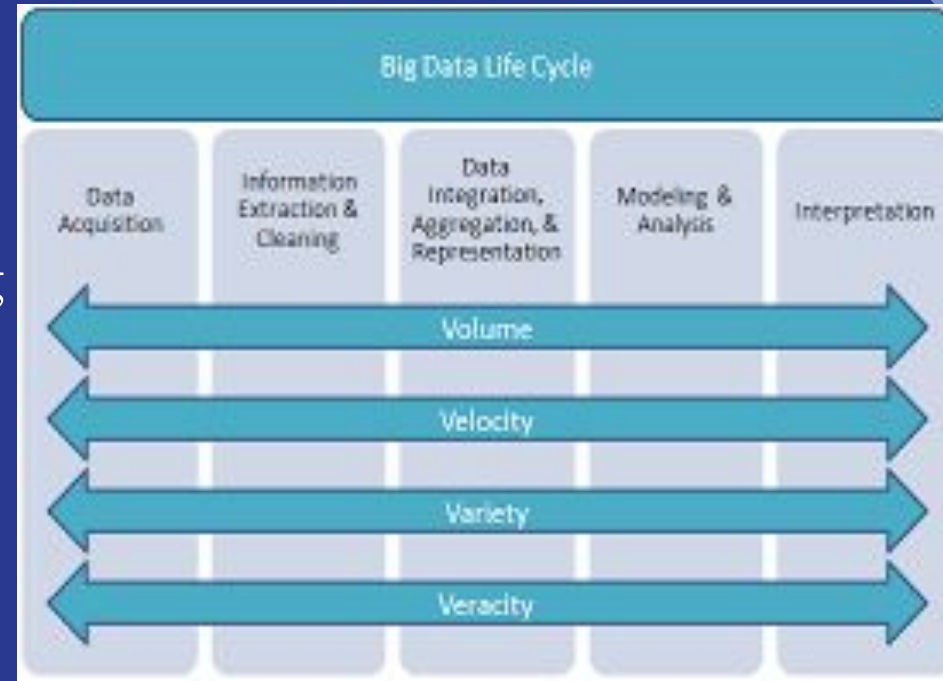


Big Data & Hadoop



# Phases of Big Data Analysis

- Phases of Big Data Analysis:
  - Data Acquisition & Capture,
  - Extraction & Cleaning
  - Data Integration, Aggregation & Representation
  - Query Processing & Modeling
  - Analysis, Interpretation & Presentation





# Challenges of Big Data

**DATA  
CHALLENGES**



**PROCESS  
CHALLENGES**

**MANAGEMENT CHALLENGE**

# DATA CHALLENGES



**Data  
Quality**



  
**KEEP  
CALM  
AND  
TRUST  
THE DATA**

**DATA**  

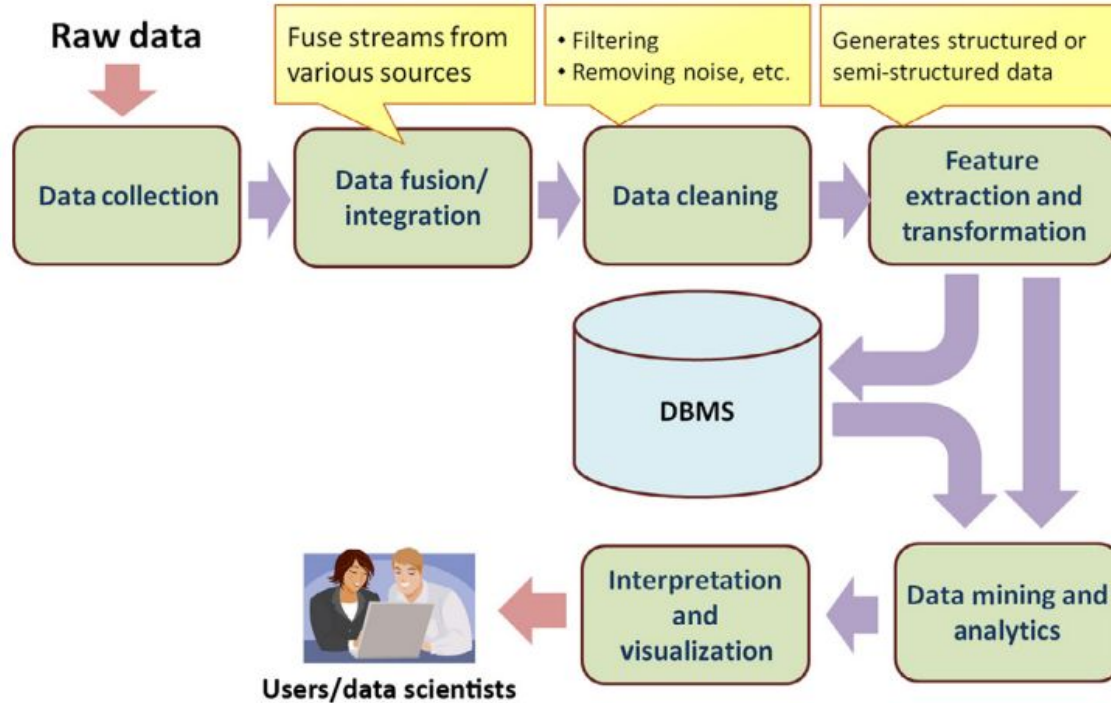
---

**DOGMA**

Dan McClellan & Dan Beecher



# PROCESS CHALLENGES



# MANAGEMENT CHALLENGES



**SECURITY**



**PRIVACY**



**GOVERNANCE**

# Solutions for big data challenges

Layer 1: IT Physical Infrastructure  
(Data Challenges)

Layer 2: Security Infrastructure  
(Management Challenges)

Layer 3: Operational Databases  
(Process Challenges)



# Solutions for big data challenges

Layer 4: Organizing Data Services and  
Tools (Data & Process Challenges)

Layer 5: Analytical Data Warehouses  
(Process Challenges)



# Solutions for big data challenges

Layer 6: Big Data Analytics:  
(Process Challenges)

Layer 7: Big Data Applications  
(Process Challenges)



# Conclusion

- Big data presents challenges beyond just volume, involving diverse data types and a demand for real time analysis
- Technological evolution brings opportunities for cost reduction, improved decision-making, and innovative enhancements, emphasizing the importance of addressing big data challenges so the full potential can be leveraged
- Big Data challenges can be solved by the Big Data Technology Stack (7 Layers)
- Traditional database architectures are not able to process large data sets and need sophisticated systems
- Big data technologies will continue to develop for larger and real-time data processing capabilities