

7BDIN006W

Big Data Theory and Practice

Lecture 11

Big Data Applications. Enterprise Data Management. Governance and Quality (including ROI) of Data.

UNIVERSITY OF
WESTMINSTER 

Applications of Big Data: Healthcare

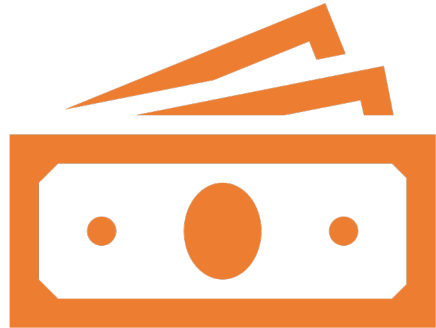


Predictive analytics: Reduced hospital readmission by up to 20% (Source: Healthcare IT News).

Genomic analytics: Reduced cost of genetic sequencing by 99% since 2001 (Source: National Human Genome Research Institute).

Real-time health monitoring: Wearables market is projected to reach \$27 billion by 2023 (Source: CCS Insights).

Applications of Big Data: Finance



Fraud detection: Big Data analytics reduced fraud loss by 10% (Source: McKinsey).

Algorithmic trading: Represents 60-73% of all U.S. equity trading volume (Source: CNBC).

Customer insights: 75% of financial services have enhanced customer experiences using Big Data (Source: Deloitte).

Applications of Big Data: Retail



Personalised recommendations: Increased sales by up to 15% (Source: McKinsey).

Inventory management: Big Data analytics can cut supply chain errors by up to 50% (Source: Accenture).

Customer behavior tracking: The retail data market is expected to reach \$10.4 billion by 2025 (Source: Mordor Intelligence).

Applications of Big Data: Logistics

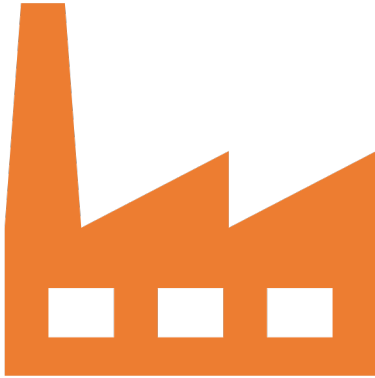


Route optimisation saved up to 20% in fuel costs (Source: American Trucking Associations).

Predictive maintenance: Reduced downtime by 35% (Source: Deloitte).

Real-time tracking: Improved delivery efficiency by 25% (Source: PwC).

Applications of Big Data: Manufacturing



Predictive maintenance: Reduced maintenance costs by 25% (Source: Deloitte).

Production optimization: Increased production efficiency by up to 20% (Source: GE Digital).

Supply chain transparency: Improved end-to-end visibility by 65% (Source: Bain & Company).

Applications of Big Data: Public Sector



Smart cities: \$189.5 billion market by 2023
(Source: Markets and Markets).

Public safety improvements: Crime prediction accuracy increased by 35% (Source: IBM).

Policy making: Data-driven policies have improved public resource management by 30% (Source: McKinsey).

Enterprise Data Management (EDM)

Definition: Systematic approach to collecting, storing, and using data.

Importance: Essential for data quality, accessibility, and security.

Data Management Challenges: Growing data volumes, diversity of data types, and rapid data velocity.

Big Data Tools and Technologies

Hadoop: A framework processing large data sets (75% of businesses using Hadoop report improved performance - Source: Forbes).

Spark: Fast processing for large-scale data (Over 1000 organisations contribute to the Apache Spark project - Source: Databricks).

NoSQL Databases: Handling varied data types (Cassandra, MongoDB - 3x faster performance than traditional databases - Source: MongoDB).

Challenges in Data Management

Volume: 2.5 quintillion bytes of data are created daily, and managing this is a key challenge (Source: IBM).

Variety: Over 80% of organisational data is unstructured, requiring advanced management solutions (Source: Gartner).

Velocity: Data streaming in real-time needs immediate processing (5G will increase data velocity by 10x - Source: Ericsson).

Data Management Strategies

Centralized Data Management: Consolidating data into a single, accessible source.

Data Lakes: Storing raw data in its native format for future use (35% increase in the use of data lakes - Source: Aberdeen Group).

Cloud-Based Solutions: Flexible, scalable data management (Cloud data management market to reach \$46.4 billion by 2025 - Source: Markets and Markets).

Big Data Governance

Definition: The process of managing the availability, usability, integrity, and security of data.

Importance: Essential for ensuring data quality and compliance in the age of Big Data.

Key Components: Data quality, data security, and regulatory compliance.

The Importance of Governance

Ensures Data Integrity: 80% of analysts' time is spent verifying data quality (Source: Harvard Business Review).

Security and Compliance: 47% of organisations consider data security as the main driver for governance (Source: Forbes).

Facilitates Data Sharing: Proper governance encourages safe and efficient data sharing within organisations.

Developing a Governance Framework

Key Components: Policies, standards, procedures, and metrics.

Stakeholder Involvement: Essential for successful implementation.

Regular Audits and Updates: To ensure compliance and adapt to new challenges.

Data Quality

Accuracy: Correctness of data in representing real-world entities.

Consistency: Uniformity of data across different sources.

Completeness: Inclusion of all necessary data elements.

Timeliness: Up-to-dateness of the information.

Ensuring Data Quality in Big Data Projects

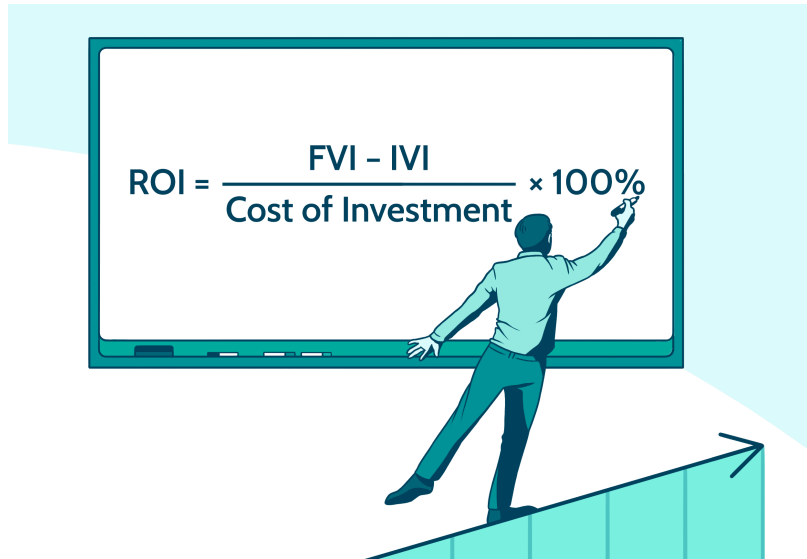
Data Governance Frameworks: Establishing policies and standards.

Data Profiling and Cleansing: Regularly assess and clean data.

Investment in Quality Tools: Utilising advanced tools for data validation and monitoring.

Understanding ROI in Big Data Initiatives

ROI Definition: Return on Investment - a measure of the profitability of investments.



- FVI – Final Value of Investment
- IVI – Initial Value of Investment

Factors Influencing ROI: Data quality, effective analytics, and alignment with business objectives.

Calculating ROI: Cost savings, revenue increase, and efficiency gains.

Measuring ROI of Big Data

Cost Reduction: Savings from process improvements and operational efficiencies.

Revenue Enhancement: Increased sales from data-driven marketing and customer insights.

Efficiency Gains: Time saved and increased productivity from data-driven decision-making.

Strategies for Enhancing Big Data ROI

Aligning with Business Goals: Ensuring Big Data initiatives support overarching business objectives.

Skilled Workforce: Investing in training and hiring data professionals.

Continuous Improvement: Regularly review and adjust Big Data strategies based on performance.