Big Data Architecture Concepts

OEM: Arcitura • Duration: 3 Days (24 hrs) • Code: B90.10

COURSE MODULES & TOPICS

Module 1: Fundamental Big Data Science & Analytics

- Understanding Big Data
- Fundamental Big Data Terminology and Concepts
- Big Data Business Drivers and Technology Drivers
- Traditional Enterprise Technologies Related to Big Data
- OLTP, OLAP, ETL and Data Warehouses in relation to Big Data
- Characteristics of Data in Big Data Environments
- Dataset Types in Big Data Environments
- Structured, Unstructured and Semi-Structured Data
- Metadata and Data Veracity
- Fundamental Analysis and Analytics
- Quantitative and Qualitative Analysis
- Machine Learning Types
- Descriptive and Diagnostic Analytics
- Predictive and Prescriptive Analytics
- Business Intelligence and Big Data
- Data Visualization and Big Data
- Big Data Adoption and Planning Considerations

Module 2: Big Data Analysis & Technology Concepts

- Big Data Analysis Lifecycle (from Business Case Evaluation to Data Analysis and Visualization)
- A/B Testing and Correlation
- Regression and Heat Maps
- Time Series Analysis
- Network Analysis and Spatial Data Analysis
- Classification and Clustering
- Filtering, including Collaborative Filtering and Content-based Filtering
- Sentiment Analysis and Text Analytics
- Clusters and Processing Batch and Transactional Workloads
- How Cloud Computing relates to Big Data
- Foundational Big Data Technology Mechanisms

- Big Data Storage Devices and Processing Engines
- Resource Managers, Data Transfer Engines and Query Engines
- Analytics Engines, Workflow Engines and Coordinate Engines

Module 14: Fundamental Big Data Architecture

- Security Engines, Cluster Managers and Data Governance Managers
- Visualization Engines and Productivity Portals
- Machine-Level Data Processing Architectural Models
- Shared-Everything and Shared-Nothing Architectures
- Big Data Analytics Logical Architecture
- Data Sources and Data Acquisition Layers
- Storage, Processing and Batch Layers
- Realtime Processing, including Event Stream and Complex Event Processing
- Enterprise Data Warehouse and Big Data Integration Approaches (including Series and Parallel)
- Poly Source, including Relational, Streaming and Filebased Sources
- Poly Storage, including Automatic Data Replication and Data Size Reduction
- Random Access Storage, including High Volume Binary, Tabular, Linked, Hierarchical and Data Sharding
- Streaming Access Storage, including Streaming Storage and Dataset Decomposition
- Large-Scale Batch Processing, Complex Decomposition and Processing Abstraction
- Poly Sink, including Relational Sink, File-based Sink and Automated Dataset Execution
- Big Data Appliance and Data Virtualization
- Architectural Environments, including ETL
- Analytics Engines and Application Enrichment
- Cloud Computing and Big Data Architectural Considerations
- Cloud Delivery and Deployment Models for Hosting Big Data Solutions

Module 15: Advanced Big Data Architecture

- Enterprise Data Warehouses and Big Data
- Operational Data Store, Data Marts and Analytical Databases
- Big Data Solution Architectural Layers
- Big Data Architecture, Maintenance and Governance
- Big Data Security Architecture
- Series, Parallel, Appliance and Virtualization Approaches
- Big Data and Cloud-based Storage and Data Processing
- Canonical Data and Large-Scale Graph Processing
- Realtime Access Storage and Direct Data Access
- Analytical Sandbox and Confidential Data Storage
- Batch Data Processing and Dataset Denormalization
- Online Data Repository and Big Data Warehouse Architecture
- Operational Data Store and Indirect Data Access
- Integrated Access and Centralized Dataset Governance
- Event Stream Processing and Complex Event Processing
- Fan-in Ingress, Fan-out Ingress and High Velocity Realtime Processing

- Data Egress, Data Visualization and Data Utilization
- Data Wrangling, Data Processing and Data Analysis Processing
- Big Data Solution Design Patterns and Architectural Compound Patterns
- Lambda Architecture, Layers, Characteristics and Benefits

Module 16: Big Data Architecture Lab

- Reading Exercise 16.1: SFI Case Study Background
- Lab Exercise 16.2: Design Big Data Pipeline for SLA Compliance
- Lab Exercise 16.3: Alleviate Customer Dissatisfaction
- Lab Exercise 16.4: Reduce Data Storage Cost
- Reading Exercise 16.5: LOC Case Study Background
- Lab Exercise 16.6: Solution for Intelligent Oil Exploration
- Lab Exercise 16.7: Enhance Oil Well Production
- Lab Exercise 16.8: Reduce Maintenance Costs and Achieve Regulatory Compliance
- Reading Exercise 16.9: TXC Case Study Background
- Lab Exercise 16.10: Identify Fraud and Eliminate Waste
- Lab Exercise 16.11: Prioritize Resource Allocation and Enable Open Data Access
- is authored by a dedicated courseware development team
- has a self-test, accreditation exam and professional certification
- is available via two different eLearning platforms
- undergo a common development process
- are authored to be consistent in quality, structure and style
- share a common vocabulary and symbol notation
- are authored in collaboration with subject matter experts
- About Arcitura
- Instructor-Led Training & Coaching
- eLearning with Arcitura
- Course & Certification Tracks
- Exams & Proctoring
- Digital Accreditations
- Trainer Development
- Partner Program
- Partner Portal
- Privacy Policy
- Candidate Agreement
- Logo Guidelines
- Contact
- Help
- Arcitura on LinkedIn