

## Module One

### Data Integration Concepts

- The Need for Data Integration
  - Why We Integrate Data
  - A Projects Perspective
- The Challenges of Data Integration
  - Understanding Data Sources
  - Choosing the Right Data Sources
  - Data Quality
  - Data Availability
- Data Integration Architectures
  - Integration Hub
  - Integration Bus
  - Integration Services
- Data Integration Projects
  - Kinds of Projects
  - Project Activities
- Data Integration Technologies
  - Extract-Transform-Load (ETL)
  - Enterprise Information Integration (EII)
  - Enterprise Application Integration (EAI)
  - Master Data Management (MDM) and More

## Module Two

### Requirements Analysis for Data Integration

- Integration Requirements Concepts
  - Overview
- Source Data Requirements
  - An Overview
  - Kinds of Data Sources
  - Evaluating Data Sources
  - Source Data Analysis and Profiling
  - Choosing Data Sources
- Data Unification Requirements
  - Subject Orientation
  - Entity Consolidation
  - Identity Consolidation
  - Relationship Consolidation
  - Attributes and Values Consolidation
- Data Aggregation and Summary Requirements
  - Levels of Detail
- Data Quality Requirements
  - Data Correctness
  - Timeliness
  - Data Integrity
- Data Capture Requirements
  - Frequency of Data Capture
  - Collecting Historical Data
  - Level of Detail
- Audit, Balance and Control Requirements
  - ABC's of Data Integration
- Metadata Capture Requirements
  - Data About Integration Processes

- Service Level Requirements
  - Meeting Expectations

### **Module Three**

#### Data Integration Functional Design

- Functional Design Concepts
  - Overview
- Source/Target Mapping
  - Mapping Techniques
  - Entity Mapping
  - Data Store Mapping
  - Data Element Mapping
  - The Full Set of Data Elements
- Data Capture Design and Specification
  - An Overview
  - Kinds of Data
  - Push vs. Pull
  - All Data vs. Changed Data
  - Changed Data Detection
  - Data Extraction
  - Data Replication
  - Transaction Logging
  - Messaging
  - Storing Captured Data
- Data Transformation Design and Specification
  - Kinds of Transformations
  - Data Selection and Filtering
  - Conversion and Translation
  - Derivation and Summarization
  - Identifying Transformations
  - Specifying Transformation Logic
- Data Cleansing Design and Specification
  - Detecting Data Quality Defects
  - Repairing Data Quality Defects
  - Quality Metadata and the ABCs of Cleansing
- Identity and Key Management
  - De-Duplication
  - Surrogate Key Assignment
- Design for Integrated Data Delivery
  - Choosing the Right Delivery System
- Data Integration Process Design
  - Requirements – Driven Processing

### **Module Four**

#### Data Integration Technical Design

- Technical Design Concepts
  - Overview
  - Comprehensive Processing Design
- Data Flow Design
  - Moving Data through the Integration Pipeline
  - Data Capture and Data Staging
  - Transformation Processes
  - Transformation Sequence and Dependencies
  - End-to-End Data Flow
- Work Flow Design
  - Extending Data Flow with Events

- Service Level Design
  - Performance and More
- Process Management Design
  - Metadata Capture and Event Logging
  - Balancing and Audits
  - Error and Exception Handling
  - Communication

### **Module Five**

#### Construction, Deployment, and Operation

- Construction, Deployment, & Operations Concepts
  - Overview
- Building Data Integration Systems
  - Tools and Technology
  - Standards, Frameworks, Templates, and Reuse
  - System Management and Data Integration
  - System Testing and Data Integration
- Implementing Data Integration Systems
  - One-Time Data Consolidation
  - Ongoing Data Consolidation
- Operating Data Integration Systems
  - Integration System Operations
  - Customer and User Support
  - Change Management

### **Module Six**

#### Summary and Conclusion

- Best Practices in Data Integration
  - Learned through Experience
- References and Resources
  - For More Information

### **Appendix A**

#### Basis of Course Examples

- Scenario
  - Overview of an Acquisition
- E-Max Systems
  - E-Max HRMS and Payroll
  - E-Max HR and Payroll Data
- PlayNation Systems
  - PlayNation HR and Payroll
  - PlayNation HR and Payroll Data
- E-Max Database
  - Data Elements Listing
- E-Max Flat Files
  - Data Elements Listing
- PlayNation Database Tables
  - Data Elements Listing
- PlayNation Flat Files
  - File Listing

### **Appendix B**

#### Bibliography and References

**Exercises**

- Exercise 1: Integration Options
  - Exercise Instructions
  - Worksheet
- Exercise 2: Data Unification
  - Exercise Instructions
  - Data Descriptions
  - Worksheet 1 of 2
  - Worksheet 2 of 2
- Exercise 3: Identify and Key Management
  - Exercise Instructions
  - Worksheet
- Exercise 4: Data Flow Design
  - Exercise Instructions and Workspace
  - Worksheet